Streptomyces specialized metabolites for agricultural applications

by ©Jody-Ann Clarke

A Thesis submitted to the

School of Graduate Studies

in partial fulfillment of the requirements for the degree of

Masters of Science

Department of Biology

Memorial University of Newfoundland

December 2021

St. John's, Newfoundland and Labrador

ABSTRACT

Streptomyces biosynthesize a wide variety of specialized metabolites, including volatile organic compounds (VOCs). *Streptomyces* VOCs are diverse and used in different capacities, including competition for resources and intra- and inter-species communication. And so, VOCs may be exploited in farming as inhibitors to plant pests and pathogens. The demand for safer alternatives to synthetic chemicals in agricultural practices makes these naturally occurring compounds ideal for applications. In this study, 133 *Streptomyces* isolates obtained from different locations in St. John's, Newfoundland, were phylogenetically typed using the *rpoB* housekeeping gene, and different strains were selected for fermentation and pooled culture VOC screening. Six of the phylogenetically different isolates (JAC25, JAC45, JAC60, JAC74C, JAC81, and JAC95) were subjected to replicate VOC analysis and phytotoxic effects on plants. These isolates were shortlisted based upon the number and types of VOC they produced, i.e., various VOCs with published bioactivities, including insect repellency making them good targets for pest management in crop farming. Ninety-five bacterially derived compounds were identified among the six isolates in the replicate VOC analysis. One isolate (JAC60) significantly inhibited the development of plant seedlings suggesting its capacity to make a possible natural herbicidal compound.

GENERAL SUMMARY

Streptomyces bacteria use VOCs in nature for intra- and inter-species communication. Sometimes these VOCs may be found in beneficial symbiotic relationships with plants where plant predators are warded off. Therefore, these compounds may be exploited for agricultural applications, wherein *Streptomyces* VOCs may effectively improve plant growth, productivity, and disease resistance. This thesis looked at the genetic relatedness of a collection of soil *Streptomyces* obtained in St. John's, Newfoundland. Genetically different *Streptomyces* were selected for VOC analysis to identify compounds that could be potential pest management agents. The results showed that VOCs are diverse among the *Streptomyces* isolates and that *Streptomyces* may not always be plant-friendly. Future experiments, including insect repellency testing, are encouraged before using *Streptomyces* and derived products directly on plants.

ACKNOWLEDGMENTS

My supervisor, Dr. Kapil Tahlan, identified, designed, and secured funding for this project. I would also like to thank Dr. Hiller for funding the study. I thank Dr. Tahlan and members of my supervisory committee, Dr. Dawn Bignell and Dr. Kirk Hiller, for their helpful comments and suggestions about experimental design and thesis writing. I thank Dr. Bignell for providing the protocol for doing the phytotoxicity test. I conducted all of the benchwork except for the processing of the volatile organic compound, which was performed by Dr. Sean McCann performed. Regarding the data analyses, I would like to thank Dr. McCann for his assistance with the GC-MS analysis and the R codes development and for providing advice and insight in interpreting the results of these analyses. I want to thank Dr. Zhenlong Cheng for teaching me lab techniques, Gustavo Adolfo Diaz Cruz for photography techniques used for the qualitative assessment of the phytotoxicity test, and Arshad Ali Shaikh for helpful comments and suggestions during the writing of this thesis.

Table of Contents

ABSTRACT…….……...……ii

GENERAL SUMMARY……iii

ACKNOWLEDGEMENTS……..……..……iv

Table of Contents…..……...…….….……..v

List of Tables……..…….……..viii

List of Figures……..…….…….….….ix

List of Abbreviations and Symbols……..…….…..…….…….……x

List of Appendices……..…….……xi

Chapter 1: Introduction…….1

1.1: The genus *Streptomyces*……1

1.2: Bacterial gene sequences for molecular identification……..2

1.2.1: The 16S rRNA gene…….……3

1.2.2: Protein-coding gene targets…….……7

1.2.3: Multi-locus sequence typing and analysis……..11

1.2.4: Whole-genome sequencing…….....13

1.3: *Streptomyces* volatile organic compounds (VOCs)…….……....15

1.3.1: Role of VOCs in *Streptomyces*……..……18

1.3.2: Biosynthesis of *Streptomyces* VOCs…….……19

1.3.3: Potential applications of *Streptomyces* VOCs in agriculture……..…25

1.4: The main objectives of this thesis…….…30

Chapter 2: Materials and Methods…….….32

2.1: General equipment and materials…….…32

2.2: *Streptomyces* procedures…….32

2.2.1: Maintenance and stock preparation…….……..32

2.2.2: Reviving of *Streptomyces* isolates from the strain collection……...…..33

2.2.2.1: Liquid cultures for mycelial stocks and chromosomal DNA extraction…….…….…..33

2.2.2.2: *Streptomyces* cultures for spore stock preparations…..…….33
2.2.3: Chromosomal DNA extraction…….……..…34
2.2.4: Liquid cultures for bioassays and VOC collection……..…..……..35
2.2.4.1: Liquid cultures for pooled media VOC collection.……36
2.2.4.2: Liquid cultures for replicate VOC collection…….36
2.2.5: Media growth profile of *Streptomyces* isolates……..…….37
2.2.6: Antibiotic susceptibility profiling of *Streptomyces* isolates…….……..37
2.2.7: DNA assay for *Streptomyces* growth in fermentation media…….….37
2.3: DNA procedures……..39
2.3.1: Polymerase chain reaction and agarose gel electrophoresis conditions……..39
2.3.2: Primer design for sequencing the *rpoB* gene amplicon……...……41
2.3.3: Phylogeny for determining relatedness of *Streptomyces* isolates…..….42
2.3.4: 16S rDNA identification of *Streptomyces* isolates…….…43
2.3.5: AntiSMASH analysis of JAC60 possible species IDs……...…….44
2.4: Bioactivity screening of culture supernatants…….……..44
2.5: VOC collection and GC-MS analysis procedures……46
2.5.1: VOC collection, including apparatus set-up and cleaning…….…….46
2.5.2: VOC collection trap processing…….....47
2.5.3: Identification of VOCs…….…..48
2.6: Phytopathogenicity screening of select *Streptomyces* isolates……49
2.6.1: Phytotoxicity assessment of *Streptomyces* isolates…….....50
Chapter 3: Results and Discussion……..52
3.1: The collection of *Streptomyces* isolates is phylogenetically diverse……52
3.2: The *Streptomyces* isolates had various phenotypes……..55
3.3: VOC analysis……...61
3.3.1: Screening of VOC production by select *Streptomyces* isolates……..…61
3.3.2: Six isolates selected for replicate VOC analysis…….……79

3.3.3: VOC production profiles for replicate cultures of the six selected *Streptomyces* isolates……..…..93

3.3.4: There was no direct relationship between *Streptomyces* growth and VOC production……...….109

3.4: Bioassay typing of selected *Streptomyces* isolates……112

3.4.1: Antibiotic bioactivities varied for phylogenetically different isolates……...112

3.4.2: Phytopathogenicity testing of select *Streptomyces* isolates…….116

3.5: *Streptomyces* isolates identification required more than the 16S rDNA in some cases…….…..121

3.6: Concluding remarks……...……126

References……...128

Appendices…….….152

List of Tables

Table 1.2.1.1: Bacteria specific 16S universal PCR primers……..……..5

Table 1.2.2.1: Some common bacterial protein-coding genes used in species delineation……9

Table 1.3.3.1: Some VOCs produced by *Streptomyces* and relevant bioactivities.……...…26

Table 2.3.1.1: Oligonucleotide primers used for PCR amplification and DNA sequencing in the current study……..……..…….40

Table 2.4.1: Indicator organisms used for bioactivity assessment of culture supernatants……..……..45

Table 3.2.1: Characteristics of *Streptomyces* isolates from this study……...……..56

Table 3.3.1.1: Bacterially derived compounds detected in pooled culture VOC analysis…….…….63

Table 3.3.1.2: The numbers of bacterially derived VOCs detected in the pooled cultures……...…….75

Table 3.3.1.3: The number of VOCs detected from each *Streptomyces* isolate in pooled culture VOC analysis……..…….77

Table 3.3.2.1: VOCs identified with high confidence in the pooled culture screen and their published bioactivities……...……80

Table 3.3.2.2: VOCs detected in the pooled culture screen for the six *Streptomyces* isolates selected for replicate VOC analysis….……87

Table 3.3.3.1: Bacterially derived compounds identified in replicate VOC analysis……...……96

Table 3.4.2.1: Table 3.4.2.1: Radish seedlings height after treatment with the *Streptomyces* isolates and *S. scabies* 87-22 supernatants from the SFM and OBB media, respectively.……..……..……119

Table 3.5.1: Molecular identification of select *Streptomyces* isolates……...……123

List of Figures

Figure 1.2.1.1: Relative location of the 16S rDNA variable regions denoted V1- V9……..6

Figure 1.3.1: Representative structures of the major classes of bacterial VOCs……...16

Figure 3.1.1: Phylogenetic comparison of *Streptomyces* isolates in this study…….54

Figure 3.3.1.1: Example view (~9- 23 min) of pooled culture GC-MS chromatograms for six *Streptomyces* isolates……...…..62

Figure 3.3.2.1: Number of VOCs detected in the pooled culture screens for the six *Streptomyces* isolates selected for replicate VOC analysis……..…..86

Figure 3.3.3.1: Example (~10- 22 min) of GC-MS chromatogram for JAC81 replicate cultures (x3) fermented in SFM medium……...……..94

Figure 3.3.3.2: The numbers of VOCs produced by the *Streptomyces* isolates in replicate culture analysis…….……95

Figure 3.3.4.1: Growth of the *Streptomyces* isolates fermented in SFM, YMS, and the synthetic media…….……111

Figure 3.4.1.1: Comparative single-culture supernatant bioassay against six indicator organisms…….……113

Figure 3.4.1.2: Comparative replicate supernatant bioassay against six indicator organisms……...……115

Figure 3.4.2.1: Phytotoxic effects of *Streptomyces* isolates spores on radish seedlings…….……117

Figure 3.4.2.2: Phytotoxic effects of *Streptomyces* isolates supernatants on radish seedlings……120

List of Abbreviations and Symbols

The following abbreviations and acronyms are used throughout this thesis. Acronyms common throughout the life sciences are not included in this list.

Abbreviation	**Meaning**
µM	Micro-meter
BGC	Biosynthetic gene cluster
BLAST	Basic local alignment search tool
Bp	Base pairs
CO_2	Carbon dioxide
CPS	Counts per second
dH_2O	Distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
GC-MS	Gas chromatography-mass spectrometry
HCl	Hydrochloric acid
KCl	Potassium chloride
KH_2PO_4	Potassium phosphate monobasic
MHA	Muller Hinton agar
Min	Minute
MLSA	Multi-locus sequence analysis
MLST	Multi-locus sequencing type
mm	Millimeter
Na_2HPO_4	Sodium phosphate dibasic
NaCl	Sodium chloride
NIST	National Institute of Standards and Technology
nm	Nanometer
nM	Nano-molar
nt	nucleotide
OTU	Operational taxonomic unit
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
RI	Kovats retention index
Rpm	Rotation per minute
rpoB	RNA polymerase beta subunit
rRNA	Ribosomal RNA
RT	Retention time
SFM	soy flour mannitol
sec	second
ST	sequence type
VOC	volatile organic compound
WGSA	Whole-genome sequence analysis
YMS	yeast extract-malt extract-starch
YPD	yeast extract-peptone-dextrose
ZOI	Zone of inhibition

List of Appendices

A.2.1: Media preparations……..……..……..…….152

A.2.3.2.1: *Streptomyces* reference sequences used to make consensus sequence for *rpoB* primer design……..154

A.2.3.2.2: *Streptomyces* reference sequences for *rpoB* primer design……..159

A.2.3.3.1: NCBI reference *Streptomyces* and isolates *rpoB* sequences compiled……...228

A.2.3.3.2: *Streptomyces* isolates *rpoB* sequences compiled…….…….290

A.2.5.3.1a: Integration R-Code for GC-MS before GC column cut……..…….312

A.2.5.3.1b: Integration R-Code for GC-MS after GC column cut……320

A.3.1.1: Phylogenetic comparison of *Streptomyces* isolates to *Streptomyces* reference sequences……328

A.3.5.1: AntiSMASH predicted BGCs in the genome of *S. griseofuscus*…….……..329

A.3.5.2: AntiSMASH predicted BGCs in the genome of *S. murinus*……..…331

Chapter 1: Introduction

1.1: The genus *Streptomyces*

Streptomyces is the largest genus of Actinobacteria and of the family Streptomycetaceae. The genus comprises sporulating Gram-positive bacteria with filamentous shape and high GC-content in their genomes (Dworkin, 2006). They are ubiquitous soil dwellers occupying various habitats, including the rhizosphere and sometimes as symbionts of plants and animals. *Streptomyces* species life cycles include forming vegetative mycelia under favorable conditions, i.e., when nutrients and water are available. When nutrients become limited, aerial hyphae develop from vegetative mycelia and mature into spores (Chater, 2006). *Streptomyces* spores can remain dormant for long periods and germinate when conditions become favorable once again, and then the cycle repeats. Notably, specialized metabolite production, including antibiotics, occurs in a growth-phase-dependent manner (Demain & Fang, 1995). Specialized metabolic products are not essential for the survival and growth of the producing organism under laboratory conditions; but instead, they enhance the ability of the producer to thrive in its natural habitat (Katz & Baltz, 2016). These molecules offer a selective advantage to compete with other microorganisms for space and nutrients. Specialized metabolite production is theorized to coincide with or slightly precede the development of aerial hyphae in surface-grown cultures. In liquid-grown cultures, it generally occurs in the stationary phase (Bibb, 2005). This event is frequently assumed to result from nutrient limitation, and specialized metabolite biosynthesis during this phase of the life cycle is speculated to be a mechanism by which *Streptomyces* ward off competitors by keeping nutrients and other resources away

from the species (Chater, 2006). Many of these specialized metabolites can be isolated and used to inhibit susceptible species outside of the natural context in which they are used.

Streptomyces are recognized as industrially-relevant because they produce an array of beneficial natural products with uses in medicine and agriculture. Specialized metabolites produced by *Streptomyces* are used in clinical settings and far exceed the use of chemically synthesized ones (Heul et al., 2018). Widely used molecules derived from *Streptomyces* include antifungal, anti-cancer, immunosuppressant, antibacterial, and anthelminthic agents (Barka et al., 2015). Genome mining techniques have revealed that even a single *Streptomyces* species, selected at random, contains an impressive number of biosynthetic gene clusters (BGCs) responsible for producing diverse specialized metabolites (Baltz, 2017; Blin et al., 2019). This diversity of BGCs in *Streptomyces* makes this group of bacteria ideal for the exploitation of novel compounds. When determining *Streptomyces* isolates for specialized metabolite analysis, molecular identification is a valuable tool discussed subsequently.

1.2: Bacterial gene sequences for molecular identification

Several methods have been used to identify bacteria based on phenotypic properties, such as staining, biochemical testing, and molecular methods that employ DNA sequencing. Of the two, modern techniques often use DNA sequencing, as it provides higher resolution and reduces ambiguities often associated with phenotypic tests (Tang et al., 1998). However, even though DNA sequencing is accepted as the gold standard, it has its problems. There is no consensus when selecting DNA sequences to identify bacteria, but some gene sequences are used more often than others. Sequence-based analysis is

usually done with several criteria in mind. First, when selecting genes for identification, one should consider targets that may serve as molecular clocks for phylogeny (Reller et al., 2007). Second, selected genes should have variable and hypervariable regions flanked by conserved segments common to all bacteria. Conserved regions are necessary for universal primers to bind during PCR amplification and sequencing the selected DNA sequence. The variable and hypervariable regions contain nucleotide sequences for species identification (Reller et al., 2007). PCR amplified and sequenced DNA is then compared to reference sequences available in public and proprietary databases to determine relatedness. Relatedness is expressed as a percent identity between two sequences. Depending on the DNA sequence, a specified difference in percent identity is acceptable before classifying a new species (Reller et al., 2007). With these things in mind, one or more gene sequences are selected that will best classify unknown bacterial species into taxonomic groups depending on the desired level of species discrimination required. Methods of DNA identification include using the 16S rRNA gene, protein-coding genes, multilocus, and whole-genome sequencing, which are briefly discussed.

1.2.1: The 16S rRNA gene

The gene most often sequenced for bacterial identification encodes the 16S rRNA (16S rDNA), approximately 1500 bp in size (Johnson et al., 2019). It has been extensively used and has readily available reference databases for comparison, so it is helpful to include it in species identification. Although widely used in the molecular identification of bacteria, there is no consensus regarding which regions of the 16S rDNA are appropriate for species-level resolution (Pollock et al., 2018). The gene consists of regions with different degrees

of conservation and variation, where more conserved sequences have served to create universal 16S rDNA primers for PCR (Table 1.2.1.1). There are nine hypervariable regions (Figure 1.2.1.1), based on the *Escherichia coli* numbering system, which have been identified to be helpful in species identification (Baker et al., 2003). Within the genus *Streptomyces*, the highly variable gamma and alpha regions are thought to be better targets to differentiate between species (Kataoka et al., 1997; Mehling et al., 1995; Song et al., 2004). Select hypervariable regions in the 16S rDNA are typically PCR amplified, sequenced, and compared to reference sequences made available through accessible databases. These include the National Center for Biotechnology Information (NCBI) nucleotide database (Agarwala et al., 2018), which comprises DNA and RNA sequences, and the SILVA rRNA database project (Quast et al., 2013), which has curated ribosomal RNA data including that for 16S rDNA.

Different species of bacteria may contain varying copy numbers of 16S rDNA. When multiple copies of 16S rDNA are present, they may exist as variants accounting for differences in sequences within the same species (Case et al., 2007). Because of these sequence differences within copy variants, species classification using the 16S rDNA sometimes poses interpretation problems and limits the resolution achievable for different species. In addition, the 16S ribosomal subunits are present in all prokaryotes, including the domain archaea, as the subunits are responsible for the conserved function of protein translation, which makes the gene sequence relatively conserved at the genus level, which can obscure inter- and intra-species identification (Fox et al., 1992). Other challenges associated with using 16S rDNA include the lack of a standard where different

Table 1.2.1.1: Bacteria specific 16S universal PCR primers.

Name	Sequence 5'–3'	Length (nt)	Reference
8F/ 27F	AGAGTTTGATCCTGGCTCAG	20	(Reysenbach et al., 1994; Shi et al., 2019)
9F	GAGTTTGATCCTGGCTCAG	19	(Farrelly et al., 1995; Hansen et al., 1998; Inerney et al., 1995)
63F	CAGGCCTAACACATGCAAGTC	21	(Marchesi et al., 1998)
341F	CCTACGGGAGGCAGCAG	17	(Teske et al., 1996)
515F	GTGCCAGCAGCCGCGGTAA	19	(Hansen et al., 1998)
516F	TGCCAGCAGCCGCGGTA	17	(Nagashima et al., 2003; Reysenbach et al., 1994)
806R	GGACTACHVGGGTWTCTAAT	20	(Apprill et al., 2015; Caporaso et al., 2011)
907R	CCCCGTCAATTCCTTTGAGTTT	22	(Teske et al., 1996)
1408R	TGACGGGCGGTGTGTACAAGGC	22	(Hansen et al., 1998; Marchesi et al., 1998)
1492R	ACGGCTACCTTGTTACGACT	20	(Reysenbach et al., 1994; Shi et al., 2019)

Figure 1.2.1.1: Relative location of the 16S rDNA variable regions denoted **V1- V9**. Variable regions are illustrated based on their relative size to each other. The figure was adapted from information made available by Baker et al. (2003).

hypervariable regions, whether alone or in combination, are used to identify a single species or to categorize unidentified bacteria into operational taxonomic units (OTUs). In the case of OTUs, a cut-off percent for sequence identity is used to cluster bacteria with the assumption that they are the same species. However, grouping different species into OTUs increases the potential for false decreased diversity due to the high percent identity of the 16S rDNA sequences among closely related species. Therefore, closely related species with similar 16S rDNA sequences, even within the variable regions, can be mistakenly labelled as the same species when one is trying to identify below the genus level. For instance, the "*Bacillus cereus* group" of bacteria, including *Bacillus cereus, Bacillus thuringiensis,* and *Bacillus anthracis,* cannot be adequately differentiated using their 16S rRNA gene sequences and are distinguished by the plasmids that they carry (Rasko et al., 2005). Likewise, bacterial strains may be more difficult to distinguish based on the 16S rDNA alone. For example, ten *Streptomyces* strains with 16S rDNA identical to the type strain of *Streptomyces cyaneofuscatus* were distinguished into five strains when other analyses including phenotypic, microscopic, and genetic comparisons were used (Antony-Babu et al., 2017). Such analyses suggest that higher resolution of the *Streptomyces* strains is sometimes required; and that 16S rDNA sequence comparison alone will not suffice while studying members of this and other genera.

1.2.2: Protein-coding gene targets

There does not appear to be a universal gene sequence in the literature with sufficient discriminatory power for identifying bacteria below the genus level. The domain bacteria consist of many species with genomes evolving at different rates. Therefore, DNA sequences that provide the best solution for a specific molecular identification need must be selected. Ideally, such targets should include universally conserved and ubiquitous genes for the molecular identification of bacteria across taxa. Protein-coding genes have been chosen recently for such analysis, yet there is no consensus on which ones to use. One study analyzed 143 single-copy genes present in more than 95% of 132 taxonomically diverse bacterial genomes, 39 of which had two conserved regions, for primer design, separated by at least 900 nt (Santos & Ochman, 2004). Ten of these universally conserved bacterial genes are listed in Table 1.2.2.1, among other common protein-coding genes used for the taxonomical analysis of bacteria. Other studies apply a similar approach for identifying phylogenetic markers to divide bacteria into subgroups (Gil et al., 2004; Wu et al., 2013; Zeigler, 2003). A recent study identified 92 bacterial single-copy genes present in more than 95% of the complete genome sequences from 1,429 organisms belonging to 28 phyla (Na et al., 2018). Further analysis of these 92 genes revealed that the majority (n = 67) encodes proteins with roles in cellular processes, including translation, ribosomal structure, and biogenesis.

Many protein-coding genes exist as a single copy, making them easier to work with than the multi-copy and multi-variant 16S rDNA. They can also evolve faster than rDNA genes because of substitutions at the third position of codons that do not change the amino acid sequence of the encoded protein (Han et al., 2012). And because of the faster rate of

Table 1.2.2.1: Some common bacterial protein-coding genes used in species delineation.

Gene	Gene product	Reference
fusA	elongation factor G	(Holmes et al., 2004; Santos & Ochman, 2004)
gyrB	DNA gyrase β subunit	(Han et al., 2012; Holmes et al., 2004; Santos & Ochman, 2004)
ileS	Isoleucine-tRNA ligase	(Santos & Ochman, 2004)
lepA	elongation factor 4	(Qin et al., 2006; Santos & Ochman, 2004)
leuS	leucine tRNA synthetase	(Ribeiro et al., 2015; Santos & Ochman, 2004)
pyrG	cytidine triphosphate synthetase	(Santos & Ochman, 2004)
recA	recombinase A	(Han et al., 2012; Holmes et al., 2004; Santos & Ochman, 2004; Thompson et al., 2005)
recG	ATP-dependent DNA helicase	(Santos & Ochman, 2004)
rplB	ribosomal protein L2	(Mekadim et al., 2018; Santos & Ochman, 2004)
rpoB	RNA polymerase β subunit	(Devulder et al., 2005; Holmes et al., 2004; B. J. Kim et al., 2004; Santos & Ochman, 2004)
atpD	ATP synthase β subunit	(Han et al., 2012)
cpn60	chaperonin 60	(Miñana-Galbis et al., 2009)
hsp65	heat-shock protein 65	(Devulder et al., 2005)
pyrH	Uridylate kinase	(Thompson et al., 2005)
rpoA	RNA polymerase alpha subunit	(Thompson et al., 2005)
trpB	tryptophan synthase β subunit	(Han et al., 2012)
sod	metalloenzyme superoxide dismutase	(Devulder et al., 2005)

evolution, protein-coding genes tend to have less conserved regions when compared across taxonomic groups. Consequently, universal primers for such genes might not be readily available for gene amplification and sequencing for subsequent species identification. Therefore, when a higher level of discrimination is required, details regarding the family and genus of bacteria being analyzed need to be considered for target selection. Due to differential evolution rates, some protein-coding genes are good candidates for intra- and inter-species delineation at taxonomic levels where the 16S rDNA is not sufficiently discriminatory. For example, the *cpn60* gene (also known as *hsp60* or *groEL*) is found in both bacteria and eukaryotes and displays more variability in sequences of inter- and intra-species compared to 16S rDNA, even though they display similar phylogenetic tree topologies (Links et al., 2012). The *Bacillus subtilis* group comprises eight closely related species and strains indistinguishable by 16S rDNA sequence (98.1–99.8 % identity), but they can be distinguished using the *gyrB* gene (75.4–95.0% identity) (Wang et al., 2007). In another example, *Mycobacterium gastri* and *Mycobacterium kansasii,* indistinguishable based on 16S rDNA sequence (Rogall et al., 1990), were successfully typed with 96.2% identity using the *recA* gene (Blackwood et al., 2000). For even greater species and phylogenetic resolution, more than one protein-coding gene sequences can be used (Guo et al., 2008), where the different sequences from the same organism are concatenated, creating sequences that result in more robust phylogenetic trees.

Of the more widely used protein-coding genes used for bacterial identification, the *rpoB* provides better delineation when compared to the 16S rDNA (Khamis et al., 2005; Kim et al., 2004; Vos et al., 2012). The percent cut-off used for *rpoB* gene sequence-based

delineation depends on the portion of the *rpoB* gene used, and for accurate annotation, a hypervariable region (between 2300 and 3300 nt) is recommended for analysis (Adékambi et al., 2009). In such cases, a 300–600 bp fragment of this region requires 94–95% identity, and a 600– 825 bp fragment requires 96–97% identity to be considered the same species (Adékambi et al., 2009). However, when classifying a new bacterial species, it is proposed to sequence the entire *rpoB* gene and use a cut-off of 97.7% identity (Adékambi et al., 2009) and classify a new species strain, using a delineation cut-off of 98.2% identity (Adekambi et al., 2008). The *rpoB* gene was selected along with the 16S rDNA in this study to identify *Streptomyces* isolates.

1.2.3: Multi-locus sequence typing and analysis

For intra-specific level classification, delineation methods include multilocus sequence typing (MLST), which uses allelic mismatches from a small number (usually seven) of housekeeping genes (Enright & Spratt, 1999; Maiden et al., 1998), typically protein-coding like those discussed above (Table 1.2.2.1). PCR amplified DNA from gene sequenced alleles are then assigned arbitrary numbers or codes with a new or unique allelic sequence getting a different number. For each isolate, the combination of assigned codes from each locus defines its sequence type (ST), where isolates with identical sequences are classified as the same ST. Online MLST reference databases for diverse organisms are available, e.g., PubMLST.org (Jolley et al., 2018), which can be used for comparison. Sometimes, more than seven loci are used in the analysis, which might not be necessary, as shown in a study that used six and eleven loci in 107 *Neisseria meningitidis*is isolates to obtain similar levels of discrimination (Maiden et al., 1998). However, MLST is limited in

application to within species and does not work for mixed samples of bacteria from different genera, which would result in allelic mismatches that create false STs (Gevers et al., 2005).

In comparison, multilocus sequencing analysis (MLSA) can use a set of genes present in different species within a genus or genera within a family (Glaeser & Kämpfer, 2015). Unlike MLST, which uses allele profiles to differentiate isolates within a species, MLSA allows for the examination of mixed genera. Gene sequences are generally concatenated, and a phylogenetic tree is prepared to measure relatedness among bacterial isolates. Bacterial MLSA uses 400-500 bp partial sequences of housekeeping genes (Jolley & Maiden, 2014), and as few as three genes (*rpoA*, *recA*, and *pyrH*) have been demonstrated to build more robust phylogenetic trees instead of any one of the single genes in the case of *Vibrio* isolates (Thompson et al., 2005). In a separate report, the sequences of four genes (16S rDNA, *hsp65*, *rpoB*, and *sod*) from 81 *Mycobacterium* isolates provided good results, except that six members of the *M. tuberculosis* group were not resolved using this method (Devulder et al., 2005). In a study of *Streptomyces* with identical 16S rDNA sequences, further delineation was possible using five protein-coding genes, *atpD, gyrB, recA, rpoB,* and *trpD* (Antony-Babu et al., 2017). As many as 93 (Chevrette, Carlson, et al., 2019) and 155 (Alam et al., 2010) genes from genome sequences have been used in MLSA to resolve *Streptomyces* and other closely related species. When using large numbers of genes for multilocus phylogeny, sequencing the genomes is less time-consuming than PCR amplifying and sequencing each gene individually and dramatically streamlines the process.

1.2.4: Whole-genome sequencing

All of the methods described above only examine a small portion of the bacterial genome and miss genetic variations caused by horizontal gene transfer, including the gain and loss of BGCs. It is not uncommon for bacteria to undergo mutation as a means to counter antibiotic susceptibility or for them to pick up resistance genes from the environment (Munita & Arias, 2016). These events pose challenges while examining pathogenic microbes in a clinical setting. For example, multidrug resistance *Klebsiella pneumoniae* isolates classified as a single lineage (ST258) based on MLST, partitioned into two distinct clades following whole-genome sequencing (DeLeo et al., 2014). Analysis of the two ST258 isolates revealed inconsistencies in the number and types of resistant genes present in each strain, indicating the importance of using higher resolution techniques for characterizing pathogenic microbes under clinical settings. These kinds of dilemmas usher in the sequencing of whole genomes to enable the correct sorting of bacterial isolates. It is also important to follow the acquisition of genetic material such as complete BGCs in *Streptomyces*, which are important events that can lead to the production of beneficial specialized metabolites in the new host. Whole-genome sequencing allows for further examination of species with identical single genes sequences or close MLSA percent identity. Information in whole genomes includes the core genome, composing genes common to all studied strains; the dispensable genome, which is shared by a subset of the strains and provides information on diversity and environmental adaptations; and the strain-specific genes that offer insights at below species level (Tian et al., 2016; Vernikos et al., 2015). For example, a comparison of a *Streptomyces* isolate and a reference species

genomes showed 100% identity of their 16S rDNA and 99.92% identity using MLSA, but genome analysis between the two revealed that one of them contained 11 unique genes and is missing 98 genes present in the other (Heinsch et al., 2019). Here, we see that a single gene or a multigene approach misses the variability that a whole-genome method provides.

The increasing affordability of whole-genome sequencing and the availability of bioinformatics tools make it easier to conduct molecular sequencing analyses. Useful for genome studies, the WGSA database (www.wgsa.net) allows the characterization of species genomes, multilocus STs, resistance profiles, and phylogenetic placement among the existing species based on core genes (Balloux et al., 2018). At the time of writing, the WGSA database comprises 43,767 genomes. The Genomes OnLine Database (GOLD) (Mukherjee et al., 2021) has information on 374,599 organisms, with Actinobacteria genome projects accounting for 8.7% of those. The NCBI genome database (Agarwala et al., 2018), which comprises 249,357 prokaryotes genomes, may also be used. The *Streptomyces* specific database, StrepDB (http://strepdb.streptomyces.org.uk/), provides annotated reference genomes of *Streptomyces*. As of July 2021, there were 671 annotated *Streptomyces* species in the List of Prokaryotic Names with Standing in Nomenclature (LPSN) database (Parte et al., 2020), which provides comprehensive information on the nomenclature of prokaryotes.

As it is now recognized that the 16S rDNA alone is insufficient to identify some bacteria to the species level, other genes can be helpful when used in combination with it. However, it is recommended to include the 16S rDNA because it has been extensively used for a longer time and has readily available reference databases for comparison. It is helpful

to use it in complementation with at least one protein-coding gene. In the current study, *rpoB* along with 16S rDNA was used for the molecular identification of select *Streptomyces* isolates. Further testing, including morphology and metabolomics, encourages a more accurate resolution of unknown or unnamed species. Furthermore, it is understood that bacteria can acquire genetic material from the environment, which may play a part in the classification below the species level. Cognisant of these factors, a *rpoB* gene phylogeny was used to determine relatedness and create OTUs for the *Streptomyces* isolates from Newfoundland; and representatives from each OTU were chosen for specialized metabolite analysis.

1.3: *Streptomyces* volatile organic compounds (VOCs)

Generally known for the numerous antibiotics they produce, *Streptomyces* also synthesize an array of VOCs (Citron et al., 2015; Citron, Rabe, et al., 2012). VOCs are described as small odorous molecules (containing up to 20 carbon atoms) with low molecular mass (100–500 Daltons), high vapor pressure, low boiling point, and also containing a lipophilic moiety (Schmidt et al., 2015). These characteristics aid the volatility and diffusion of such molecules to mediate interactions among species in the environment. Bacterial VOCs are grouped into major chemical classes, including hydrocarbons (alkanes, alkenes, alkynes), alcohols, aldehydes, ketones, esters, and terpenes (Figure 1.3.1) (Schulz & Dickschat, 2007). Some VOC classes and members have been more extensively studied than others. VOCs are produced by plants (Vivaldo et al., 2017), animals including humans (Curran et al., 2005; Ni et al., 2012), and fungi (Schnurer et al., 1999). VOCs are often exploited for their aroma in the food, cosmetic, chemical, and pharmaceutical industries

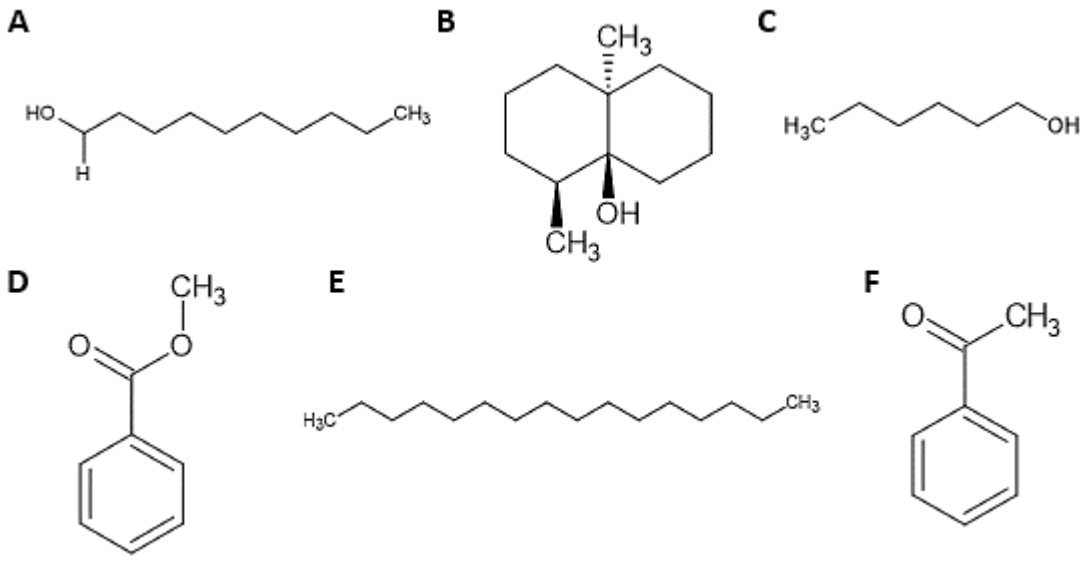


Figure 1.3.1: Representative structures of the major classes of bacterial VOCs **A.** Decanal (aldehyde) **B.** Geosmin (terpene) **C.** 1-Hexanol (alcohol) **D.** Methyl benzoate (ester) **E.** Hexadecane (hydrocarbon) **F.** Acetophenone (ketone)

and are frequently manufactured synthetically for these purposes. However, newer approaches may see microbes as commercial sources for such relevant compounds (Gupta et al., 2015). In addition, VOCs are essential to the sustenance of ecosystems as they regulate a range of functions and behaviours related to the success of a species in its habitat.

VOCs act as information-carrying signals or pheromones from one individual to another, whether intra- or inter-species. This may take on different outcomes in bacteria, including competition and cooperation, as VOCs may be antagonistic or take on supportive roles. For instance, 2,3 butanediol and acetoin are required for full virulence in the plant pathogen, *Pectobacterium carotovorum*. Similarly, the same compounds can increase the production of virulence factors in clinically relevant *Pseudomonas aeruginosa* (Audrain et al., 2015). In some cases, VOCs can modify the behaviours of other bacteria like ammonia, trimethylamine, hydrogen sulfide, nitric oxide, and 2-amino-acetophenone, which can alter biofilm formation or affect the motility of bacteria (Audrain et al., 2015; Raza et al., 2016). Altering motility contains a species and prevents it from acquiring nutrients, thus affecting its survival. Other organisms may exploit bacterial VOCs such as protists, substantial bacterial predators and use volatile signals to locate edible colonies (Schulz-bohm et al., 2017). Bacterial terpene synthase mutants could not mobilize protists (Schulz-bohm et al., 2017), demonstrating that VOCs are directly involved in protist-bacteria interactions.

Streptomyces are sources of valuable specialized metabolites for medicinal use (Emerson de Lima Procópio et al., 2012), suggesting that they might also be good candidates for studying VOCs for agricultural applications. The role of *Streptomyces* VOCs is subsequently discussed, followed by their biosynthesis and potential uses in agriculture.

1.3.1: Role of VOCs in *Streptomyces*

VOCs from *Streptomyces* play essential roles in proliferation, communication, defense, and competition for resources against other microbes. *Streptomyces venezuelae* is shown to perform exploration, a type of motile growth not universal to the genus (Jones et al., 2019), where endogenously produced trimethylamine (TMA, a VOC) increases environmental pH leading to reduced iron availability. In response to TMA, *S. venezuelae* upregulates siderophore biosynthesis and secretion to chelate iron for uptake, which reduces available iron for other bacteria and fungi in the immediate vicinity. Glucose depletion due to growth and a rise in pH due to TMA causes *S. venezuelae* to produce non-branching vegetative hyphae and rapidly cross solid surfaces (Jones et al., 2017). Interestingly, explorer cells can release air-borne VOCs to communicate this exploratory behavior at a distance with other species members to induce their exploratory growth. This behavior is presumably done to escape nutrient-poor conditions and occupy new niches.

Certain VOCs also mediate *Streptomyces* spores dispersal and thus increase their chances of proliferation. The odorous VOCs, geosmin and 2-methylisoborneol, can be sensed by other organisms and may attract or repel them. For example, the two VOCs attract the springtail arthropod (*Folsomia candida*) to induce electrophysiological responses in their antennae (Becher et al., 2020). Upon sensing these VOCs, springtails are attracted to *Streptomyces* colonies where they feed, and spores are then dispersed via the fecal route and by adhering to their hydrophobic cuticles. Because of the nature of VOCs, they are quickly diffused through air and soil to cover relatively short distances. Therefore, VOCs with antimicrobial activities have the added advantage of inhibiting competitors at

a distance. Ammonia is an example of one such naturally occurring chemical released in considerable amounts by some soil *Streptomyces*, which has been shown to inhibit *E. coli* ASD19 in mixed cultures under laboratory conditions (Avalos et al., 2019). Catabolism of the amino acid glycine is required to produce ammonia in *Streptomyces* species, but glycine defective mutants could not synthesize inhibiting amounts of ammonia even when they were grown on increased amounts of the amino acid (Avalos et al., 2019).

Symbiotic relationships between *Streptomyces* and higher animals are not uncommon. For example, the beewolf digger wasp (*Philanthus* spp., Hymenoptera, Crabronidae) incorporates *Streptomyces* into the larval cocoon, protecting larvae against pathogens (Kroiss et al., 2010). It was found that the *Streptomyces* species produce a cocktail of nine antibiotic substances that inhibited pathogenic fungi and bacteria. In addition, VOCs from microbes might account for communication signals that occur in insects and larger animals. Dodecanol/dodecanal are synthesized by soil *Streptomyces* (Citron, Rabe, et al., 2012; Dias et al., 2017), whereas dodecanol was isolated from the codling moth's (*Cydia pomonella*) pheromone gland (Lofstedt & Bengtsson, 1988) and dodecanal from the male ring-tailed lemurs' antebrachial secretions, which mediates attraction of females during mating season (Shirasu et al., 2020). *Streptomyces* were not linked as producers of dodecanol/dodecanal in the codling moths and ring-tailed lemurs, but soil Streptomyces synthesize both molecul*es*. Therefore, these molecules could be of microbial origin. To further understand the roles of *Streptomyces* VOCs and, by extension, their potential application, it is important to understand their biosynthesis.

1.3.2: Biosynthesis of *Streptomyces* VOCs

Understanding the biosynthesis of different VOCs will allow for the genetic manipulation of producers to prepare over-producing strains. Specialized metabolism is not essential to bacteria, and thus, the genes involved are dispensable. As a result, bacterial genome evolution is a dynamic process aided by horizontal gene transfer. In these events, bacteria acquire genetic material through intra- and inter-species transfer (Thomas & Nielsen, 2005). It is hypothesized that the host maintains beneficial DNA sequences for its survival and proliferation in its habitat (González-Torres et al., 2019). Hence, horizontal gene transfer has a role in shaping the metabolome of bacteria. Acquisition and maintenance of DNA that is nonessential to the growth (under laboratory conditions) of a species is often reflected in the *Streptomyces* genus. This group of bacteria is frequently noted for the diversity and number of BGCs present in a given *Streptomyces* genome (Baltz, 2017). Said BGCs will likely function in a beneficial way to the species survival in the face of competitors or sub-optimal environmental conditions. Furthermore, these BGCs may be passed from parent to offspring through vertical transmission, where they are maintained in the species because of often unknown selective pressures.

Because bacteria constantly gain and lose DNA, nonessential BGCs are not necessarily fixed within a population. Therefore, general molecular methods of identifying and classifying *Streptomyces* are not enough when investigating specialized metabolite biosynthesis. If these general methods alone are used, it would be easy to overlook a *Streptomyces* isolate harboring BGCs for novel specialized metabolites. For example, phylogenetic analysis of 287 *Streptomyces* with 100% identity of partial 16S rDNA sequences contained an average of ten unique BGC families each (Chevrette, Carlos-

shanley, et al., 2019). Another study revealed that terpene synthase genes that originated in bacteria were successfully transferred to fungi (Jia et al., 2019), indicating some compatibility in genetic material between the two microbial domains. Hence, bacteria may acquire similar genes from eukaryotes. Eukaryotes, including insects and plants, produce VOCs, whereas it is not uncommon to have symbiotic relations between bacteria and eukaryotic hosts (Crawford & Clardy, 2011; Kim et al., 2019). Therefore, it is possible that bacteria might be responsible for producing some VOCs in their symbiotic hosts. For example, insect pheromones such as methyl 6-methylsalicylate, methyl 6-ethylsalicylate, mellein, pyrrole-2-carboxylate, conophthorin, and chalcogran are among some isolated from actinomycetes (Citron, Rabe, et al., 2012). Further studies are warranted to investigate the role of such symbiotic relations and VOC and genes involved in the process.

The amount and composition of VOCs produced by microorganisms can vary according to culture conditions and substrate availability (Garbeva et al., 2014). Other important factors influencing the production of VOCs are the physiological state of the producing microorganism, oxygen availability, moisture, temperature, and pH (Insam & Seewald, 2010). As seen with terpenes that comprise the vast majority of naturally produced VOCs, a small number of substrates are used to synthesize a mixture of molecules by cyclases with similar backbones. Monoterpenes are generated from geranyl diphosphate, diterpenes from geranylgeranyl diphosphate, and sesquiterpenes from farnesyl diphosphate (Dickschat, 2016), where sesquiterpene synthases can sometimes also use geranyl diphosphate as a substrate (Dickschat et al., 2014). Diterpene cyclases are further classified

into type I or type II enzymes based on their mechanism of action (Dickschat, 2016) but, both classes have similar substrate specificities.

Many VOCs are theorized to be side products of metabolic pathways formed during primary metabolism or from various intermediates during specialized metabolism (Korpi et al., 2009). Studies of Actinomyces terpenes biosynthetic pathways reveal that it is common for even highly specific terpene cyclases to produce mixtures of compounds along with their dominant product (Rabe et al., 2013). For instance, the well-studied geosmin biosynthetic pathway has at least 15 side products or deduced derivatives (Rabe et al., 2013). Often, genes encoding cytochrome P450s and methyltransferases are part of operons formed by terpene cyclases and are speculated to catalyze a series of oxidations and methylations of the parent sesquiterpene synthase product, resulting in mixtures of end products (Hu et al., 2011).

Sometimes VOCs are precursors in antibiotic biosynthetic pathways, as shown for the *epi*-isozizaene synthase product identified in *Streptomyces coelicolor* (Zhao et al., 2008). The biosynthesis of the sesquiterpene antibiotic albaflavenone is elicited by two-gene containing cluster (Lin & Cane, 2009). The first gene product catalyzes the cyclization of farnesyl diphosphate to the sesquiterpene hydrocarbon, epi-isozizaene and the second carries out the oxidation of *epi*-isozizaene in a two-step process, leading to albaflavenone. Furthermore, homologs of *epi*-isozizaene synthase and albaflavenone genes are translationally coupled in *Streptomyces avermitilis* (Takamatsu, Lin, et al., 2011), although *epi*-isozizaene or any of its oxidized derivatives, albaflavenols and albaflavenone, were not identified under culture conditions used.

β-Cedrene is a sesquiterpene isolated from *Streptomyces arenae, Streptomyces citreus* (Rabe et al., 2013), and soil *Streptomyces* species present in the rhizosphere of disease suppressive soil (Cordovez et al., 2015). Synthesis of β-Cedrene in actinomycetes is thought to only occur in species capable of synthesizing *epi*-isozizaene, (Rabe et al., 2013), and is proposed to be a side product of the antibiotic albaflavenone biosynthetic pathway. In addition, the change of one amino acid at position 198 of the *epi*-isozizaene synthase converts it to a β-cedrene synthase (Li et al., 2014).

The pentalenolactone antibiotic BGC in *S. avermitilis* contains all of the necessary genes to produce the antibiotic (Quaderer et al., 2006). In the pathway, the sesquiterpene pentalenene is converted to pentalenolactone via a proposed series of intermediates, including 1-deoxypentalenic acid, pentalen-13-ol, and pentalen-13-al whereas, pentalenic acid is a shunt product (Takamatsu, Xu, et al., 2011). The pentalenolactone BGC does not contain any specific gene responsible for the formation of pentalenic acid. It was shown that a cytochrome P450 gene (*sav7469*) located outside the BGC in *S. avermitilis* is responsible for converting 1-deoxypentalenic acid to pentalenic acid (Takamatsu, Xu, et al., 2011). Hence, it is apparent that genes outside of major BGCs encode enzymes that use intermediates from dominant pathways. These are a few cases linking volatile shunt products to biosynthetic pathways of specialized metabolites.

Among the first isolated VOCs from S*treptomyces* are the terpenes geosmin (Gerber & Brunswick, 1965) and 2-methylisoborneol (Gerber, 1969), where the latter is also a widely distributed terpene reported in other actinomycetes, myxobacteria, and cyanobacteria (Citron et al., 2015). The two VOCs are responsible for the characteristic

earthy odor of soil and cause contamination of drinking water and edible fish. 2-methylyisoborneol biosynthesis occurs in a two-step pathway, where geranyl diphosphate is methylated to 2-methyl- geranyl diphosphate by an S-adenosylmethionine-dependent methyltransferase, followed by a catalyzation to 2-methylisoborneol by a terpene synthase (Martín-Sánchez et al., 2019). In *S. coelicolor*, two adjacent genes (*sco7700* and *sco7701*) elicit 2-methylisoborneol biosynthesis (Wang & Cane, 2008), whereas geosmin is produced from farnesyl diphosphate by the geosmin synthase encoded by gene *sco6073* (*cyc2*) in *S. coelicolor* (Gust et al., 2003) and *sav2163* (*geoA*) in *S. avermitilis* (Cane et al., 2006).

Geosmin synthase is a bifunctional enzyme that also catalyzed the formation of another terpene, germacrene D, using the same substrate as for geosmin (Jiang et al., 2007). The enzyme has two active sites that perform the Mg^{2+}-dependent cyclization, where the N-terminal site produces germacradienol and germacrene D from farnesyl diphosphate, and the C-terminal one converts germacradienol to geosmin. Germacrene D is a sesquiterpene isolated from *S. citreus* (Pollak & Berger, 1996), *Streptomyces fimicarius* BWL-H1 (Xing et al., 2018), and various other *Streptomyces* (Citron et al., 2015; Rabe et al., 2013).

The complex network of pathways leading to specialized metabolites reveals the diversity of molecules that arise from the transformation of intermediate products. The terpene cogeijerene is a proposed side product of the geosmin pathway (Citron, Gleitzmann, et al., 2012), where the intermediate (10R)-8,10-dimethyl-8-octalin is oxidized by an unknown enzyme. Among the lesser-studied systems is 1,8-cineole synthase, encoded by *cnsA* in *Streptomyces clavuligerus*, which catalyzes the conversion of geranyl diphosphate to eucalyptol (Rinkel et al., 2016).

Although the terpenes encompass the most produced microbial VOCs from the specialized metabolic class, other VOCs are synthesized using different enzyme families. These families include the cytochrome P450 enzymes that synthesize and metabolize various compounds, including alkanes, alcohols, aldehydes, and acids (Kelly & Kelly, 2013; Nie et al., 2014; Scheller et al., 1998). In other instances, the production of aldehydes, short-chain alkanes, fatty esters, and fatty alcohols like dodecanal and tridecane occurs via the Acyl-CoA reduction pathway by the fatty acyl-CoA/ACP reductases (Choi & Lee, 2013; Lehtinen et al., 2018). Therefore, it is essential to know the biosynthesis for specialized metabolites, especially those produced via multiple pathways. This knowledge is necessary when considering genetic manipulation methods to produce bacterial VOCs for large-scale use in industry and agriculture.

1.3.3: Potential applications of *Streptomyces* VOCs in agriculture

Annual losses in crop production due to pests affect both farmers and consumers. Therefore, finding sustainable solutions that reduce synthetic chemical use is paramount for eco-friendly agriculture. Among the notable green solutions is bioprospecting, where natural products are encouraged as alternatives to toxic chemicals (Kanchiswamy et al., 2015). VOCs are one example of safe biocontrol agents that could improve crop health and yield. Although few studies demonstrate the potential of VOCs in agriculture, the possible uses are promising and deserve attention.

In general, the bioactivities of *Streptomyces* VOCs and their applications are not well studied, but some recent reports linking VOCs to ecological processes (e.g., inhibition of microbes) have revealed some potential applications (Table 1.3.3.1). VOCs collected

Table 1.3.3.1: Some VOCs produced by *Streptomyces* and relevant bioactivities.

VOC	Function	Bacteria
Aldehydes		
Decanal	Beetle (*Scolytus schevyrewi*) repellent (Zhu et al., 2020); Bactericidal (K. Liu et al., 2012)	Various *Streptomyces* (Citron et al., 2015; Citron, Rabe, et al., 2012; Groenhagen et al., 2014)
Nonanal	Beetle (*Scolytus schevyrewi*) repellent (Zhu et al., 2020)	*Streptomyces sp.* FORM5 (Groenhagen et al., 2014)
Dodecanal	In male lemur sex pheromone (Shirasu et al., 2020)	Various *Streptomyces* (Citron, Rabe, et al., 2012)
Ketones		
Acetophenone	Tsetse fly repellent (Olaide et al., 2019); Kissing bug (*Rhodnius prolixus*) repellent (Franco et al., 2018); Western pine beetle repellent (Erbilgin et al., 2007)	Various *Streptomyces* (Citron et al., 2015; P. Xu et al., 2014)
Alcohols		
1-hexanol	Beetle (*Scolytus schevyrewi*) attractant (Zhu et al., 2020)	Various *Streptomyces* (Citron, Rabe, et al., 2012)
Benzyl alcohol	Head lice repellent (Toloza et al., 2006)	Various *Streptomyces* (Citron et al., 2015)
Dodecanol	Inhibits *Aspergillus fumigatus* biofilm formation (Mowat et al., 2010); Coddling moth sex pheromone (Lofstedt & Bengtsson, 1988)	Various *Streptomyces* (Dias et al., 2017)
Esters		
Methyl benzoate	Repel various insects (Feng & Zhang, 2017; Morrison III et al., 2019; Mostafiz et al., 2018)	Various *Streptomyces* (Citron et al., 2015; Cordovez et al., 2015; Groenhagen et al., 2014; Pollak & Berger, 1996)
Hydrocarbons		
Hexadecane	Nematicidal activity (Gu et al., 2007)	Various *Streptomyces* (Citron, Rabe, et al., 2012; X. Li et al., 2020; Schöller et al., 2002)
Heptadecane	Hydrocabon fuel (Schirmer et al., 2010)	Various *Streptomyces* (Citron, Rabe, et al., 2012; X. Li et al., 2020)
Terpenes		

epi-*isozizaene*	Precursor to antibiotic albaflavenone (B. Zhao et al., 2008)	*S. coelicolor* (B. Zhao et al., 2008)
Geosmin	Springtail attractant (Becher et al., 2020); *Drosophila* repellent (Becher et al., 2010; Stensmyr et al., 2012)	Various *Streptomyces* (Becher et al., 2020)
Methylisoborneol	Springtail attractant (Becher et al., 2020); Antifungal (Logrieco et al., 2002; Z. Wang et al., 2013)	Various *Streptomyces* (Becher et al., 2020)
α-himachalene	Antifungal and antibacterial (Chaudhary et al., 2014)	*S. clavuligerus* (Citron, Gleitzmann, et al., 2012)
1-8 cineole	Mosquito repellent (Klocke et al., 1987)	*S. clavuligerus* (Rinkel et al., 2016)
β-Cedrene	Inhibit sudden oak disease pathogen (*Phytophthora ramorum*) (Manter et al., 2007)	*S. citreus, S. arenae* (Rabe et al., 2013), other soil *Streptomyces* (Cordovez et al., 2015)
Geranyl acetone	Cassava mite's natural predator attractant (Pinto-zevallos et al., 2018); Asian larch bark beetles repellent (Q. Zhang et al., 2007); Tsetse flies repellent (Vale et al., 2012) Mosquito (*Anopheles gambiae*) repellent (Meijerink et al., 2001)	*Streptomyces sp.* GWS-BW-H5 (Martens et al., 2005)
Dihydroagarofuran	Reverse multidrug-resistance (Torres-romero et al., 2009)	*S. citreus* (Pollak & Berger, 1996),*Streptomyces yanglinensis* (Lyu et al., 2020)
Germacrene D	Mosquito (*Aedes aegypti* and *Aedes albopictus*) larvicide (Satyal et al., 2019); Chagas disease vector (*Triatoma rubrofasciata*) repellent (Satyal et al., 2019); Mosquito larvicide (Govindarajan et al., 2016)	*S. citreus* (Pollak & Berger, 1996), *S. fimicarius* BWL-H1 (Xing et al., 2018), and other *Streptomyces* (Citron et al., 2015; Rabe et al., 2013)
Other		
Trimethylamine	Promotes *Streptomyces* exploratory growth (Jones et al., 2019)	*S. venezuelae* (Jones et al., 2019)
Ammonia	Anti-bacterial (Avalos et al., 2019)	Soil *Streptomyces* (Avalos et al., 2019)

from *Streptomyces* are typically identified and documented, creating a foundation for future studies; as such, many VOCs have not yet made it to the application phase. This study aims to add knowledge of *Streptomyces* isolates VOCs to a relatively new field that may see VOCs as potential application in agriculture as natural pest deterrents.

Streptomyces VOCs have the potential to alleviate common problems, including *Drosophila* pests (Kirkpatrick et al., 2018; Wallingford et al., 2017); soft rot disease in tomato plants (Dias et al., 2017); and spoilage to litchi fruit (Xing et al., 2018) and other storage crops (Wang et al., 2013; Yang et al., 2019) caused by fungi. Synthesized by multiple *Streptomyces* species (Citron et al., 2015; Pollak & Berger, 1996), phenylethyl alcohol shows antifungal activity against plant pathogens of peanut kernels (Lyu et al., 2020) and post-harvest sweet potato (Li et al., 2020). The hydrocarbon hexadecane has nematicidal activity (Gu et al., 2007) and is also naturally synthesized in multiple *Streptomyces* (Citron, Rabe, et al., 2012; Li et al., 2020; Schöller et al., 2002).

2-methylisoborneol is an antifungal from *Streptomyces alboflavus* TD-1 that inhibits *Fusarium moniliforme* (Z. Wang et al., 2013), a facultative and non-host necrotrophic pathogen implicated in several plant diseases. *Fusarium moniliforme* broadly contaminates corn, sorghum, wheat, tomatoes, cotton, peanuts, bananas, soybeans, and other feeds where it causes diseases such as corn ear rot, wheat scab, and rice bakanae disease (Logrieco et al., 2002). Interestingly, 2-methylisoborneol is also implicated in attracting the fruit fly *Drosophila melanogaster*; however, the producing *Streptomyces* also synthesizes other specialized metabolites that are larvicidal to the *Drosophila* species (Ho et al., 2020).

The antibacterial and antifungal sesquiterpene α-himachalene has a reported presence in the volatile collection headspace of fermented *S. clavuligerus* (Citron, Gleitzmann, et al., 2012). It is active against the mycotoxigenic fungus, *Aspergillus parasiticus* (Chaudhary et al., 2014), implicated in the disease of widely used spices black and white pepper (Madhyasth & Bhat, 1984). Another terpene, geranyl acetone, plays a defensive role against cassava plant mites and is made by the plant in response to herbivory by the insect leading to the attraction of the cassava mite's natural predator (Pinto-zevallos et al., 2018), thereby protecting the plant. Geranyl acetone was isolated from *Streptomyces* strain GWS-BW-H5 (Martens et al., 2005) and the Actinobacteria, *Amycolatopsis alba* and *Amycolatopsis nigrescens* (Citron et al., 2015). In addition, the metabolite inhibits the growth of green algae (Ikawa et al., 2001), indicating its possible use in the prevention and containment of eutrophication in farm irrigation systems. Geranyl acetone may also act as an insect repellent for Asian larch bark beetles (Zhang et al., 2007), making it potentially valuable in the timber industry. Likewise, ketone acetophenone repels the western pine beetle, *Dendroctonus brevicomis*, but not its predator, *Temnochila chlorodia* (Erbilgin et al., 2007).

The ester, methyl benzoate, is synthesized by various *Streptomyces* species (Citron et al., 2015; Cordovez et al., 2015; Groenhagen et al., 2014; Pollak & Berger, 1996), and its insect repellency has been widely examined using brown marmorated stink bug (*Halyomorpha halys*), diamondback moth (*Plutella xylostella*), tobacco hornworm (*Manduca sexta*), the spotted wing *Drosophila* (*Drosophila suzukii*), American wheat weevil (*Rhyzopertha dominica*), and the red flour beetle (*Tribolium castaneum*) (Feng &

Zhang, 2017; Morrison III et al., 2019). Methyl benzoate also inhibits oviposition, nymphal development and causes mortality of adults in the Silverleaf whitefly (*Bemisia tabaci*) that infects tomato plants (Mostafiz et al., 2018), alluding to its potential application as a biopesticide to control various insects.

Geosmin, known for its earthy smell, can be used as a *Drosophila* repellent to prevent losses in small fruit crops. Geosmin acts as a stimulus to activate specific olfactory pathways that warn the flies of the presence of pathogenic fungi (Stensmyr et al., 2012). The researchers suggested that the geosmin detection system might be conserved across the *Drosophila* genus, indicating its potential use as a broad-spectrum eco-friendly *Drosophila* deterrent. Due to the widespread capacity of *Streptomyces* to produce the specialized metabolite, they are ideal choices to pursue for future applications.

The *Streptomyces* have also been shown to protect the strawberry plants and pollinating bees against pathogenic fungi and pathogenic bacteria (Kim et al., 2019). Although it is not yet known the type of specialized metabolite that provides protection, it is accepted that pollinators are attracted to nectars of plants in flowers. Therefore, VOCs may play a role in the interaction of the strawberry plants, *Streptomyces*, and honeybees. In other symbiotic relationships, *Streptomyces* in disease suppressive soils may increase plant growth and productivity. For example, *Streptomyces* VOCs inhibited hyphal growth of fungal root pathogen, *Rhizoctonia solani*, and significantly enhanced plant shoot and root biomass (Cordovez et al., 2015).

1.4: The main objectives of this thesis

Sustainable farming practices rely on natural community-level interactions to help control pest organisms. The heavy use of harsh chemicals and the lack of community-level species diversity in industrial agriculture can disrupt the ability of natural enemies to control pest populations, so it is vital to reduce harmful chemicals so that natural ecological interactions can be promoted in agroecosystems (Gliessman, 1990). As a proposed solution, naturally occurring VOCs may be used to restore and maintain natural functions to agroecosystems. Among the sources of VOCs, the *Streptomyces* is an excellent candidate genus to study. To date, many studies on *Streptomyces* specialized metabolites focus on antibiotic discovery and their biosynthesis. However, *Streptomyces* VOCs are just as bountiful and display potential for agricultural purposes. The objective of this thesis was to analyze a collection of *Streptomyces* isolates capacity for VOC production and add to the database of growing research on *Streptomyces* VOCs. One hundred and thirty-three *Streptomyces* isolates from the northeastern part of the island of Newfoundland were phylogenetically typed using the *rpoB* gene. Phylogenetically distinct isolates were then subjected to VOC analysis, and compounds with reported bioactivities such as insect repellency were noted. Such compounds may make good targets for agricultural applications, and the isolates should be considered for future insect repellency assays. The isolates were tested for bioactivities against six biological indicators (five bacteria and a fungus). A shortlist of six isolates selected on criteria of the number and types of VOCs they produced were subjected to replicate VOC analysis and phytotoxicity testing.

Chapter 2: Materials and Methods

2.1: General equipment and materials

All media and reagents were prepared with distilled water (dH_2O) and were sterilized by autoclaving. All antibiotic solutions were prepared in dH_2O or DMSO and were filter sterilized using 0.2 µm filters in the case of aqueous solutions (161013311, Basix, Fisher Scientific, Canada). Unless otherwise stated, antibiotics, reagents, and dehydrated media were purchased from Sigma Aldrich, Fisher Scientific, or VWR International (Canada). All procedures involving bacterial cultures were performed using aseptic techniques with sterile equipment and reagents. All centrifugations were performed using an Eppendorf 5424 (0012836), or 5418 microfuge-centrifuge (5418ZR715696), or a ThermoFisher Scientific Sorvall ST 16 R centrifuge (75004381) with a TX-400 Swinging Bucket Rotor (75003629). An Implen P300 NanoPhotometer (1777, Montreal Biotech Inc., Canada) or NanoDrop 1000 (I951, Thermo Scientific, USA) were used to measure nucleic acid concentrations. Isolated DNA was stored at -20°C, and bacterial glycerol stocks and culture supernatants were stored at -80°C.

2.2: *Streptomyces* procedures

2.2.1: Maintenance and stock preparation

All *Streptomyces* stocks were prepared by growing strains/ isolates on ISP-4 agar medium (M359-500G, Himedia) for 7 to 14 days and then scraping the growth with an ethanol flame-sterilized steel spatula into 1.5 ml microfuge tubes containing 500- 1000 µl of 20% glycerol. When required, nystatin (50 µg/ml) and nalidixic acid (50 µg/ml) were

included in the growth media. Mycelial stocks were prepared by centrifuging 1 ml cultures in Eppendorf tubes at 4,000 rpm for 5 min, discarding the growth media, and resuspending the cells in 20% glycerol. All *Streptomyces* isolates liquid cultures were grown at 28°C on a rotary shaker at 200 rpm (08988, New Brunswick Scientific, USA), and all cultures on solid media were incubated at 28°C (C2574, Hotpack, Canada).

2.2.2: Reviving of *Streptomyces* isolates from the strain collection

The *Streptomyces* isolates analyzed in this project were collected between 2011 and 2018 from soil samples collected from St John's, Newfoundland, Canada, and processed into spore stocks (K. Tahlan, personal communication). They were kept frozen at -80°C, and each isolate was revived by spotting 10 µl onto nutrient agar [3% w/v nutrient broth (234000, BD, Difco), 2% w/v agar in dH_2O] containing nalidixic acid and nystatin and incubated for 2-3 days.

2.2.2.1: Liquid cultures for mycelial stocks and chromosomal DNA extraction

Following incubation (section 2.2.2), some mycelia for each *Streptomyces* isolate mycelia were scraped off the respective agar plate using an ethanol flame-sterilized stainless-steel spatula and used to inoculate 5 ml of nutrient broth in 50 ml Falcon tubes containing 3 mm Borosilicate glass balls (Salem Specialty Ball, USA) to help break up mycelia during growth. Liquid cultures were then incubated for 3- 5 days, after which 1 ml of each mycelia culture was used to prepare and store mycelial stocks. Another 1 ml of each culture was used for chromosomal DNA extraction, as described later.

2.2.2.2: *Streptomyces* cultures for spore stock preparations

Some mycelia from nutrient agar plates described above (section 2.2.2) were scraped and spotted with ethanol flamed-sterilized stainless-steel spatula onto ISP-4 agar containing nalidixic acid and nystatin. Four isolates were spotted on one agar plate divided into quarters. Inoculated ISP-4 agar plates were incubated to allow for sporulation of the *Streptomyces* isolates. Once the *Streptomyces* isolates had sporulated, the spores from each quarter were scraped and transferred to fresh ISP-4 plates (one isolate per plate) for lawn cultures. The ISP-4 media plates were incubated and used to make spore stocks after sporulation had occurred.

2.2.3: Chromosomal DNA extraction

Genomic DNA extraction from environmental isolates was performed using the Presto mini gDNA Bacteria kit (GBB100, Geneaid Biotech Ltd., Taiwan) with an additional step involving bead beating. The protocol for extracting chromosomal DNA from Gram-negative bacteria was used because a higher DNA yield was obtained. Clean bead beating tubes were prepared with fine silica beads (BAWZ 400-250-35, OPS Diagnostics, USA), and 1 ml of 3-5 days old mycelial cultures grown in nutrient broth (Section 2.2.2.1) were added to appropriately labeled tubes. The tubes were centrifuged for 3 mins at 14,000 rpm, and the growth medium (supernatant) was discarded. According to the manufacturer's instructions, GT buffer containing proteinase K was added, and the tubes were capped, sealed with parafilm, and subjected to two rounds of bead beating using the pre-installed bacteria protocol (10-3280-001-26, SpeedMill plus, Montreal Biotech Inc lab equipment, Canada). Following bead beating, the tubes were centrifuged for 3 mins at 14,000 rpm to separate the lysate (DNA) from cell debris and the silica beads. The

supernatant was carefully transferred (without collecting the silica beads) to clean and appropriately labeled Eppendorf tubes. The remaining steps were followed according to the manufacturer's instructions up to the elution step, where 43 µl of 60°C nuclease-free water was used instead of the elution buffer provided in the kit. The DNA concentrations were measured and the DNA stored for future use.

2.2.4: Liquid cultures for bioassays and VOC collection

Duplicate *Streptomyces* seed cultures for inoculating fermentation media were grown in 5 ml of trypticase soy broth (211825, BD Bacto) in 25 - 50 ml capacity Erlenmeyer flasks, with stainless steel springs to break up the mycelia growth. After 48 hours of growth, 250 µl and 500 µl seeds were used to inoculate 25 ml of fermentation media in 250 ml capacity fermentation flasks with stainless steel springs. The fermentations were set up as follows:

Streptomyces scabies 87-22 (Li et al., 2019) seed (250 µl) was inoculated into oat bran broth (OBB) medium (Johnson et al., 2007), and *Streptomyces* isolates seeds (500 µl) were inoculated into one of three media (Appendix A.2.1): soy flour mannitol (SFM) medium (Citron, Gleitzmann, et al., 2012; Citron, Rabe, et al., 2012; Komatsu et al., 2008; Rabe et al., 2013; Riclea et al., 2012); a synthetic medium (Cane et al., 2006; Jáchymová et al., 2002); or yeast extract-malt extract-starch (YMS) medium (Ikeda et al., 1987).

Following 5 days incubation, 2- 3 ml of each *Streptomyces* culture was transferred to labeled 1.5 ml Eppendorf tubes (1 ml per tube) and stored at -80°C for future antimicrobial and growth assays (aliquots of *S. scabies* 87-22 from OBB were stored and

only used for phytopathogenicity testing). The remainder of each *Streptomyces* isolates culture were used for VOC collection as follows.

2.2.4.1: Liquid cultures for pooled media VOC collection

At any one time, a maximum of four *Streptomyces* isolates seed cultures were used to start fermentation using the described media (SFM, YMS, and synthetic). Un-inoculated media controls were set up using the same conditions for each medium. After incubation for 5 days, 2 ml aliquots were removed and stored for future use and, the approximately remaining 23 ml of each of the three media were pooled (69 ml) for each isolate and mixed well. For VOC collection, an aliquot of 25 ml per Streptomyces isolate was transferred to labeled 8 oz mason jars (Bernardin Ltd., Canada).

2.2.4.2: Liquid cultures for replicate VOC collection

Single *Streptomyces* isolates were used to set up seed cultures for replicate fermentations and VOC collection. The two seed cultures for each isolate were pooled and mixed well, then used to start three flasks of production cultures in SFM medium. After incubation, 3 ml from each replicate culture were drawn and stored for future use. The approximately remaining 22 ml of each culture was poured into 8 oz mason jars for VOC collection. Small, sterilized magnetic stir bars were added to each mason jar to agitate the cultures during VOC collection. The mason jars were placed atop a hot plate (07287218, VWR, USA) to use the stirring feature only (without heating). The same *Streptomyces* isolate was subjected to replicate culture fermentation in YMS and the synthetic media

following fermentation in SFM medium and VOC collection. Un-inoculated media was used under the same conditions as the *Streptomyces* isolates for control purposes.

2.2.5: Media growth profile of *Streptomyces* isolates

Each *Streptomyces* isolate was tested for growth on the three fermentation media used in this study, SFM YMS and synthetic media; each liquid media formulation (section 2.2.4) was supplemented with 2% w/v agar. A single agar plate was divided to accommodate 20 *Streptomyces* isolates. A glycerol stock of each isolate (2 µl) was spotted onto each of the three different agar media. The plates were then incubated for a maximum of 3 days. Following incubation, the plates were observed for the growth of the *Streptomyces* isolates.

2.2.6: Antibiotic susceptibility profiling of *Streptomyces* isolates

Each *Streptomyces* isolate was tested against three antibiotics: apramycin (50 µg/ml), kanamycin (50 µg/ml), and thiostrepton (25 µg/ml). A single nutrient agar plate was divided to accommodate 20 *Streptomyces* isolates. A glycerol stock of each isolate (2 µl) was spotted onto three different agar media containing one of the three antibiotics. These agar plates were incubated for 1-3 days. Following incubation, the plates were observed for the presence of growth of each *Streptomyces* isolates.

2.2.7: DNA assay for *Streptomyces* growth in fermentation media

An indirect assay was used to measure the growth of *Streptomyces* in the fermentation media as described previously (Burton, 1968; Zhao et al., 2013). To prepare the diphenylamine reagent: 1.5% w/v diphenylamine in 150 ml glacial acetic acid, 1.5%

v/v (final) of concentrated sulfuric acid was added to this solution (the solution was stored in the dark). Before use, aqueous acetaldehyde was prepared for addition as follows: 2% v/v of cooled acetaldehyde (> 99.95% purity, kept at 4°C, and in the dark) was pipetted into 49 ml cooled dH_2O. To the 152.25 ml of diphenylamine- glacial acetic acid- sulfuric acid reagent, 0.5 % v/v (final) of the aqueous acetaldehyde was added. This diphenylamine reagent was set aside in the dark for later use.

One milliliter aliquots of *Streptomyces* cultures (Section 2.2.4.2) were centrifuged at 10,000 rpm for 10 min, and the supernatants were removed and transferred to freshly labeled Eppendorf tubes and stored for future use. The cell pellets were washed with 1 ml of PBS buffer (0.8% w/v NaCl, 0.02% w/v KCl, 0.144% w/v Na_2HPO_4, 0.024% w/v KH_2PO_4, adjusted to pH 7.4 with HCl), the suspension was centrifuged at the same speed and time as before, and the supernatant was discarded. The cells pellets were resuspended a second time with 1 ml of PBS. Aliquots (500 µl) of the cell suspensions were transferred to two freshly labeled Eppendorf tubes (duplicate for each culture). The 500 µl cell suspensions were centrifuged, and the supernatants were discarded. The cell pellets were resuspended in 1 ml of diphenylamine reagent, vortexed to break up the cell pellets, and incubated at 60°C for 1 hour (Oven: 07-32353, Binder, U.S.A.). The Eppendorf tubes were removed from 60°C, vortexed once more, and kept in the dark at room temperature overnight. Following overnight incubation, the tubes were centrifuged at 7,000 rpm for 5 min, and 150 µl of the supernatant from each tube was transferred to a well in 96 – well microtiter plates. The absorbance was measured at 595 nm with a multifunctional microtiter plate reader (270789, Synergy Hybrid Reader, BioTek industries, USA).

2.3: DNA procedures

2.3.1: Polymerase chain reaction and agarose gel electrophoresis conditions

Twenty microliter PCRs to amplify 880 bp DNA and 1465 bp DNA fragments from *rpoB* and 16s rDNA gene, respectively, were set up using 20- 250 ng of DNA, 5% v/v DMSO, 0.5 µM forward primer, 0.5 µM reverse primer, 0.2 mM dNTPs, 1× buffer (5× GC or HF) and 1% v/v polymerase from the Phusion® High-Fidelity PCR Kit (E0553L, NEB, Canada), and nuclease-free water to the final volume. PCRs were performed using an Eppendorf VapoProtect Mastercycler (6321AO714802, Fisher Scientific, Canada) or C1000 Touch Thermal cycler (CC003324, Bio-Rad, Canada). The *rpoB* PCR cycles were set up as follows: initial denaturation 98°C for 30 sec; followed by 30 cycles of 98°C for 10 sec, 58°C for 30 sec, and 72°C for 20 sec; and a final extension of 72°C for 5 min. The 16s rDNA PCR cycles were set up as follows: initial denaturation 98°C for 30 sec; followed by 30 cycles of 98°C for 10 sec, 50°C for 30 sec, and 72°C for 40 sec; and a final extension of 72°C for 5 min. DNA template for PCRs consisted of chromosomal DNA (section 2.2.3) or nuclease-free water in the case of negative controls. All primers used in the current study are listed in Table 2.3.1.1.

Two microliters from each PCR were loaded onto a 1% w/v agarose gel using 6× gel loading dye (B7024S, NEB, Canada) alongside 5 µl of 1 kb ladder (DM010-R500F, FroggaBio, Canada) for a band size estimation. Electrophoresis was conducted at 100 V in constant voltage mode using 1× Tris-borate EDTA (TBE) buffer (89 mM Tris base, 89 mM Boric acid, 2 mM EDTA, pH 8) as follows: 45 minutes for *rpoB* gene amplicons and 60 minutes for 16S rDNA gene amplicons. All gels were imaged following staining with

Table 2.3.1.1: Oligonucleotide primers used for PCR amplification and DNA sequencing in the current study.

Primer name	Sequence (5'- 3')	Annealing temperature [+]	Purpose	Reference
27F 1492R	AGAGTTTGATCCTGGCTCAG ACGGCTACCTTGTTACGACT	59	Amplifying 1465 bp of the 16s rDNA gene and sequencing for molecular identification	(Shi et al., 2019)
516F	TGCCAGCAGCCGCGGTA	-	Sequencing 16s rDNA gene amplicons for molecular identification	(Nagashima et al., 2003)
rpoBF1 rpoBR1	TTCATGGACCAGAACAACC CGTAGTTGTGACCCTCCC	58	Amplifying 880 bp of the *rpoB* gene for use in phylogeny	(Guo et al., 2008)
rpoBF2 rpoBR2	CCAGAACAACCCGCT CTCCCACGGCATGAA	-	Sequencing *rpoB* gene amplicons for phylogeny	This Study

[+] Annealing temperatures were calculated using the NEB online Tm calculator (https://tmcalculator.neb.com/#!/main) for the Phusion system.
- means not applicable.

ethidium bromide using a ChemiImager™ Ready (Alpha Innotech Corporation, USA) or a benchtop UV transilluminator (95-0458-01, UVP, USA). The reaction products were purified using the EZ-10 Spin Column PCR Products Purification Kit (Bio Basic Inc., Canada) according to the manufacturer's instructions, except that DNA was eluted using 30 µl nuclease-free water heated to 60°C instead of the elution buffer. The PCR product concentrations were measured, and the DNA was sent for sequencing at The Centre for Applied Genomics (TCAG) sequencing facility (The Hospital for Sick Children, Canada).

2.3.2: Primer design for sequencing the *rpoB* gene amplicon

Primers for the sequencing of the *rpoB* gene amplicon (Table 2.3.1.1) were designed in this study by creating a multiple sequence alignment in the Molecular Evolutionary Genetics Analysis version ten software (MEGA-X) (Kumar et al., 2018) of six *Streptomyces* species *rpoB* genes, *Streptomyces lividans, Streptomyces avermitillis, Streptomyces clavuligerus, Streptomyces coelicolor, Streptomyces scabies, and Streptomyces venezuelae*. These sequences were obtained from the National Center for Biotechnology Information (NCBI) or the StrepDB database. The consensus sequence (Appendix A.2.3.2.1) was compared against the NCBI non-redundant nucleotide collection database, specific to *Streptomyces* (taxid:1883), using the Basic Local Alignment Search Tool (BLAST). The top hits were selected and downloaded via the multiple sequence alignment viewer in aln file format (102 *rpoB* gene sequences including the query). The aln file was converted to fasta file format using the HIV sequence database format converter (https://www.hiv.lanl.gov/content/sequence/FORMAT_CONVERSION/form.html) then aligned using ClustalW in MEGA-X. One sequence (Accession CP040941.1 for

Streptomyces variabilis) was removed as it was not of sufficient length, leaving 101 sequences inclusive of the query (Appendix A.2.3.2.2). Two primer sequences were then selected from conserved regions of the alignment, flanking a variable 880 bp region. Primers were designed and then purchased from Integrated DNA Technologies, Inc. (USA).

2.3.3: Phylogeny for determining relatedness of *Streptomyces* isolates

rpoB gene phylogenetic trees of all *Streptomyces* isolates from the current study and those from *Streptomyces* genome assemblies available in the NCBI database as of April 8, 2021 (n = 282) were used to determine relatedness. Only one strain per *Streptomyces* species downloaded from NCBI genome assemblies was included in the phylogenetic analyses (n = 211). Each *Streptomyces* isolate *rpoB* gene sequence was compiled by aligning the DNA sequences for the forward and reverse sequencing primers in the online Benchling software (https://www.benchling.com/). The sequences were manually curated using the respective Sanger sequencing chromatograms (peak height, width, and periodicity) to correct any ambiguities, i.e., only one nucleotide (A, G, C, or T) was included at a given position in the alignment. The target *rpoB* gene segments of the NCBI genome assemblies were extracted with the *rpoB* primers created in this study. All *Streptomyces* and isolates *rpoB* sequences were made into two multiple sequence alignments, one with isolates from this study and reference sequences from the NCBI genome database (Appendix A.2.3.3.1) and one with only isolates from this study (Appendix A.2.3.3.2), using the Benchling software MAFFT program, which was downloaded in FASTA format for evolutionary analyses in MEGA-X.

The multiple sequence alignments obtained above were trimmed, and the evolutionary history was inferred using the Maximum Likelihood method and General Time Reversible model (Nei & Kumar, 2000). The trees with the highest log-likelihood were selected. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach and then selecting the topology with a superior log-likelihood value. The trees were drawn to scale, with branch lengths measured in the number of substitutions per site. The analyses involved 345 and 134 nucleotide sequences for *Streptomyces* species and isolates phylogeny (211 *Streptomyces* species and 133 isolates sequences) and isolates only phylogeny (133 isolates sequences) with *M. tuberculosis* as an outgroup for both trees. There were 623 and 618 positions in the final datasets, respectively. Both trees were run using 100 bootstrap replicates, and a minimum of 50% consensus cut-off value was used to create branches. The isolates phylogenetic tree was used to define OTUs for this thesis, where isolates with identical *rpoB* sequences were clustered together. The multiple sequence alignment was checked to ensure that all isolates in an OTU had 100% identity of *rpoB* sequences.

2.3.4: 16S rDNA identification of *Streptomyces* isolates

Select *Streptomyces* isolates were identified using 16S rDNA and *rpoB* sequences. Following the curation of 16S rDNA sequences as described above, each *Streptomyces* isolate sequence was compared against the NCBI 16S rRNA sequences (Bacteria and Archaea) database using BLASTN. The top hit was selected as the species ID, and if there were more than one species with the same percent identity to the query (the isolate's 16S

rDNA sequence) in the top hits, the *rpoB* sequence was then used for further resolution. In this case, the species in the 16S rDNA search top hits respective *rpoB* sequences were pulled from NCBI available genome sequences (accession numbers in Table 3.5.1) using *rpoB* primers in this study (Table 2.3.1.1) and then compared to the *Streptomyces* isolate *rpoB* sequence (Table 3.5.1) in a multiple sequence alignment using Benchling. A minimum recommended 96–97% identity for a 600– 825 bp fragment of *rpoB* between the isolate and the reference sequences (Adékambi et al., 2009) was used to determine the isolate's ID.

2.3.5: AntiSMASH analysis of JAC60 possible species IDs

An antiSMASH (Blin et al., 2021) analysis was done to predict specialized metabolites BGCs in the genomes of two shortlisted species ID for JAC60, *Streptomyces griseofuscus* and *Streptomyces murinus* (NCBI accession numbers NZ_CP051006.1 and NZ_CP0066774.1, respectively). The NCBI accession numbers for the two *Streptomyces* genomes sequences were inputted into the antiSMASH online analysis software (https://antismash.secondarymetabolites.org), and the results were downloaded following completion of the analyses.

2.4: Bioactivity screening of culture supernatants

Fermented culture supernatants (section 2.2.4) were tested against several indicator organisms (Table 2.4.1). Growth conditions and a description of each indicator organism are provided. Bioassays were performed within two months of supernatant collection. One milliliter of each supernatant (thawed from the frozen culture aliquots) was centrifuged at

Table 2.4.1: Indicator organisms used for bioactivity assessment of culture supernatants.

Indicator organism	Testing media and incubation conditions	Description/ Relevant features	Source
Saccharomyces cerevisiae ATCC 208352	YPD[e]; 28ºC, 24-48 hours	Yeast or single-celled fungus	ATCC[g]
Bacillus subtilis ATCC 23857[a]	MHA[f]; 28ºC, 24-48 hours	Gram-positive, rod-shape	ATCC[g]
Escherichia coli ESS	MHA[f]; 37ºC, 24 hours	Gram-negative, rod-shape; β-lactam sensitive	(Jensen et al., 1983)
Pseudomonas fluorescens ATCC 13525[b]	MHA[f]; 28ºC, 24-48 hours	Gram-negative, rod-shape,	ATCC[g]
Micrococcus luteus ATCC 4698[c]	MHA[f]; 28ºC, 24-48 hours	Gram-positive, coccus	ATCC[g]
Staphylococcus epidermidis ATCC 14990[d]	MHA[f]; 37ºC, 24-48 hours	Gram-positive, coccus	ATCC[g]

ATCC conditions for indicators are like those used except as follows:

[a] Grows on potato extract agar, 26°C for 18 to 24 hours.

[b] Grows on nutrient agar, 26ºC for 24-48 hours.

[c] Grows on tryptic soy agar, 28ºC for 24-48 hours.

[d] Grows on nutrient agar, 37ºC for 24-48 hours.

[e]YPD- yeast extract-peptone-dextrose agar.

[f]MHA- Mueller Hinton agar.

[g]ATCC- American Type Culture Collection.

14,000 rpm for 3 minutes, and the supernatants were collected and separated from the cell pellet and debris. Two hundred and ten milliliters of Muller Hinton agar (MHA) (V310937, Mikrobiologie, Thermo Fisher, Canada) were inoculated with 2 ml of *B. subtilis, E. coli, M. luteus, P. fluorescence, or S. epidermidis* glycerol stock. Similarly, 210 ml of yeast extract-peptone-dextrose agar (YPD) was inoculated with 2 ml of *Saccharomyces cerevisiae* glycerol stock. Inoculated media were poured into sterile 22 × 22 cm plastic bioassay trays, and 42 sterile 6 mm Whatman filter paper discs (prepared using a hole punch) were then placed on top of the agar plate. *Streptomyces* supernatants (10 µl) were spotted onto individual discs, after which the plates were incubated under conditions listed in Table 2.4.1, at 28°C or 37°C (01050214, VWR, USA). Following incubation, the zones of inhibitions were measured in millimeters.

2.5: VOC collection and GC-MS analysis procedures

2.5.1: VOC collection, including apparatus set-up and cleaning

VOC collection was performed according to the method described by Light et al. (Light et al., 2020) with some modifications. After fermentation of *Streptomyces* (section 2.2.4), a total of four mason jars containing ~25 ml of cultures each were connected to a pump (PVAS22, portable Volatile Assay System [VAS], USA) that pushed unsterile air into the headspace of each jar and simultaneously pulled that air through pre-cleaned divinylbenzene polymer VOC traps (HayeSep-Q 80/100 mesh, 27+ mg; VAS, USA). The traps were pre-cleaned using established protocols (Light et al., 2020), which involved flushing sequentially with methanol, acetone, and HPLC-grade hexane. The jars were connected to the pump with polypropylene tubing made airtight at connecting junctions as

best as possible. Plastic nuts and bolts wrapped with ½ inch Teflon tape (PS6036, PlumbShop, Canada) were used to connect tubing through two holes made in the lids of each jar. One hole linked tubing with air being pushed into the jar, and the other connected tubing with air being pulled through a trap to collect VOCs emitted from the cultures. A continuous airflow through hydrocarbon filters (PN10033.00, VAS, USA) was passed through each mason jar (0.3 L/min push and 0.1L/min pull).

The first VOC collection set-up included two fermented *Streptomyces* and two controls. One jar was left empty, and one contained uninoculated media treated the same as the test samples for control purposes. Following overnight (~17 hours) VOC trapping, the tubing was disassembled for cleaning, first with tap water, then with 95-100% ethanol. The mason jars were washed with unscented soap (043204, Sparkleen, Fisher Scientific, Canada). The traps were sealed at both ends with Teflon tape, labeled accordingly, and stored at -80°C before being sent for gas chromatography-mass spectrometry (GC-MS) analysis by Dr. Hillier's group (Acadia University, NS, Canada).

2.5.2: VOC collection trap processing

At Acadia University, GC-MS (Scion 456 Gas Chromatograph–Single Quad Mass Spectrometer, SCION Instruments, UK) was used to determine the composition of VOCs collected in the VOC traps (section 2.5.1). VOCs in the traps were eluted with 200 µl of hexane (CHROMASOLV, >98.5% purity, Sigma-Aldrich, USA) under a flow of nitrogen into 1.5 ml vials. Aliquots of 40 µl per eluted samples were transferred to 250 µl limited volume glass inserts within septum-capped autosampler vials. Ten microliters of 50 ng/µl racemic menthol (Sigma-Aldrich) was added as an internal standard to each of these

aliquots. Autosampler vials were capped, sealed with Teflon tape, and stored at 4°C before further analysis.

A non-polar capillary Rxi®-5silMS column (30 m x 0.25 mm, film thickness 0.25 mm; Restek Corporation, USA) linked to a Bruker mass spectrometer (Bruker Daltonics Ltd., UK) was used for analysis. Samples underwent splitless injection for 1 min, followed by a 20:1 split state after that (injector held at 250°C). The oven temperature was maintained for 1 min at 40°C then increased to 100°C at 6°C/ min and held for 1 min, then increased to 250°C at 10°C/ min, and the temperature held for 5 min. The injector and transfer line were maintained at 250°C. Helium was used as a carrier gas at a flow of 1 ml/min.

2.5.3: Identification of VOCs

Peaks from GC-MS analysis were identified using the Bruker-supplied MS Data Review software (version 8.0.1, Bruker Daltonics Ltd., UK) using a detection threshold of 2000 CPS, a maximum peak width of 6.0 sec, a slope sensitivity of 20, and a tangent of 10%. VOCs were identified using standards and NIST Mass Spectral Search Program for the NIST/EPA/NIH Mass Spectral Library Version 2.0g built in 2011 (abbreviated as NIST library) (Scion Instruments UK Ltd, UK).

Spectrum lists of VOCs (libraries of compound IDs) were created for each *Streptomyces* isolate and control chromatogram in the MS Data review software. An empty jar and uninoculated media subjected to the same conditions of volatile trapping, sampling, and analysis served as controls. Text files of the integrated peaks were made, and spectrum

lists were saved as MSP files. Compound tables saved as MSP files in the MS Data review software were imported in the Bruker supplied Method Builder software (version 8.0.1, Bruker Daltonics Ltd., UK). The compound tables were then copied and saved as Microsoft excel CSV files. R-studio (version 1.3.1093, RStudio, USA) was used to calculate integrations and combine the Microsoft excel CSV files and the previously created text file of the spectrum list (Appendices A.2.5.3.1a and A.2.5.3.1b). VOC matches suggested by the NIST library were then evaluated based on spectral similarity and comparing their Kovats retention indices (RI) to published values.

Compounds from the bacterial samples detected in control samples were considered media-derived or artifacts of the collection and sampling procedure, thus not of bacterial origin. The compounds with good spectral matches to the NIST library references and subsequent injection of authentic standards or supported by published RI data (absolute difference ≤ 10) (Linstrom & Mallard, 2001) were considered high confidence matches. If no authentic standards were available and the published RI could not support a NIST library suggested compound ID, or there was no NIST library suggested compound ID, the VOC was deemed a low confidence match or unknown, respectively. A further classification assigned the identified compounds with a functional group and putative origin (bacterial or control-derived).

2.6: Phytopathogenicity screening of select *Streptomyces* isolates

Six *Streptomyces* isolates selected for replicate culture analysis were tested for phytotoxic effects along with the known plant pathogen, *S. scabies* 87-22 (Li et al., 2019). Inoculants for the procedure were prepared as follows: *Streptomyces* spores (~200 µl) from

glycerol stocks were used to inoculate ISP-4 plates, one plate per *Streptomyces* isolate. The spores were spread as lawns with sterilized spreaders to cover the agar plates. The *Streptomyces* isolates and *S. scabies* 87-22 were incubated for 10 days, following which a sterile spatula was used to scrape the spores from each plate and transfer them into 2 ml of sterile dH_2O in 50 ml Falcon tubes. JAC45 was suspended in 1.5 ml of sterile dH_2O as it sporulated less than the other *Streptomyces* isolates. The spores were then vortexed to prepare a homogenous suspension to inoculate the germinated radish seeds described below. Supernatants were drawn from 1 ml aliquots of fermented *S. scabies* 87-22 in OBB (section 2.2.4) and two replicate cultures of the six isolates in SFM (section 2.2.4.2) were also used as inoculum for the germinated seeds.

Radish seeds (White Icicle, The Seed Company by EW Gaze, Canada) were placed into a 50 ml Falcon tube containing 70% v/v ethanol. The seeds were incubated for 5 min with gentle agitation on a gyro shaker (13687705, Fisher Scientific, EU). The ethanol was then poured off, 40 ml of 15% v/v bleach solution added (099-5345-6, Old Dutch, Canada), and the seeds were incubated for 10 min with gentle agitation using the gyro shaker. Following this, the seeds were rinsed 8-10 times with 40 ml of sterile dH_2O to remove all bleach. The seeds were then placed into a sterile deep Petri-dish containing sterile Whatman filter paper prewetted with 2 ml sterile dH_2O. The plate was then sealed with parafilm, and the seeds were left to germinate for 36-48 hours at room temperature in the dark.

2.6.1: Phytotoxicity assessment of *Streptomyces* isolates

Approximately 50 ml of sterile 1.5% w/v agar in dH_2O was poured into eight sterile deep Petri-dishes, 20 mm thick/deep and 140 mm wide. Petri-dishes were labeled to

accommodate one *Streptomyces* isolate per plate. Six wells (or five when the supernatant inoculum was used) per plate were generated in the agar using a sterile test tube with a mouth diameter of 13 mm. The agar plugs were removed with sterile forceps, placed into a sterile beaker, remelted in a microwave, and 200 µl was used to seal the bottom of each well. Once the agar in the wells had set, individual germinated seeds were transferred aseptically to each well using sterile forceps. Two hundred microliters of freshly prepared *Streptomyces* isolate spore suspension or supernatant were pipetted into each agar well containing one germinated seed. *Streptomyces scabies* 87-22 was used as a positive control and sterile dH_2O as a negative control. The plates were sealed with parafilm and incubated at room temperature (~23°C) under a 16-hour photoperiod for 7 days. Following the incubation period, the plates were photographed for a qualitative assessment (spores treated seedlings), or the seedlings were measured for a quantitative evaluation (supernatant treated seedlings). The shortest and longest seedling in each treatment was discarded for the quantitative assessment, and the remaining three were used to calculate an average length. The experiment was repeated using the *Streptomyces* spores.

Chapter 3: Results and Discussion

In this study, a collection of 133 *Streptomyces* isolates from St. John's, Newfoundland, were phylogenetically typed to determine relatedness using the *rpoB* gene. Thirty-seven representative *Streptomyces* isolates from different OTUs were selected for fermentation, VOC collection, and subsequent GC-MS analysis. A shortlist of six *Streptomyces* isolates was further typed using 16S rDNA and was tested for phytotoxic effects on plants. The following sections present and discuss the results obtained.

3.1: The collection of *Streptomyces* isolates is phylogenetically diverse

All isolates used in this study were obtained using soil samples from the vicinity of St. John's, located on the Avalon Peninsula of the island of Newfoundland. These soils have developed from materials derived from the underlying slate, siltstone, sandstone, shale, limestone, conglomerate, and granitic and volcanic rocks (Heringa, 1981). Most of the mineral soils are coarse to moderately coarse-textured, stony, acidic to highly acidic in pH, and low in natural fertility. The ocean heavily influences the climate, with July temperatures ranging from 13-16 °C and winter temperatures ranging from -4–2 °C in December to February.

A specific region of *rpoB* (position 1347 to 2207) from the *Streptomyces* isolates was amplified by PCR using custom primers (Table 2.3.1.1). For comparison, all *Streptomyces* genome assemblies available in the NCBI genome database (n = 282) as of April 2021 were downloaded, and *rpoB* sequences were extracted. However, only one strain per species was included in a phylogenetic analysis (211 reference *Streptomyces*

species) to remove over-represented strains. Phylogenetic analysis with the *Streptomyces* species showed that the *Streptomyces* isolates were distributed throughout the phylogenetic tree, indicating diversity in the isolates collection (Appendix A.3.1.1).

A closer examination of *rpoB* sequences of the 133 isolates from the current study showed that they could be clustered into 31 OTUs, where isolates in an OTU have 100% identity of *rpoB* sequence (Figure 3.1.1). Each isolate *rpoB* sequence was run against the NCBI non-redundant database, and it was assumed that the *rpoB* provided resolution below the genus level. One OTU contained two *Dermacoccus* species, a genus belonging to the class Actinobacteria (Parte et al., 2020; Stackebrandt et al., 1995) and was not included in further analyses. The remaining 30 OTUs contained *Streptomyces* isolates. One clade contained 25 isolates closely related to *Streptomyces pratensis* (Shirling & Gottlieb, 1969). *Streptomyces pratensis* contains a cryptic clavulanic acid-like BGC, which might produce a β-lactamase inhibitor (Álvarez-Álvarez et al., 2013). Other OTUs had one to ten *Streptomyces* isolates, and seven OTUs contained one isolate each.

It is believed that phylogeny, a reliable tool to examine evolutionary relationships, may be used to predict natural product biosynthesis (Adamek et al., 2019). Phylogenetic application may differentiate between orthologous (homologues created through speciation) and paralogous (homologues created through duplication) genes. Orthologous genes are assumed to retain equivalent functions in different species (Gabaldón & Koonin, 2013), which is also applicable for specialized metabolite BGCs.

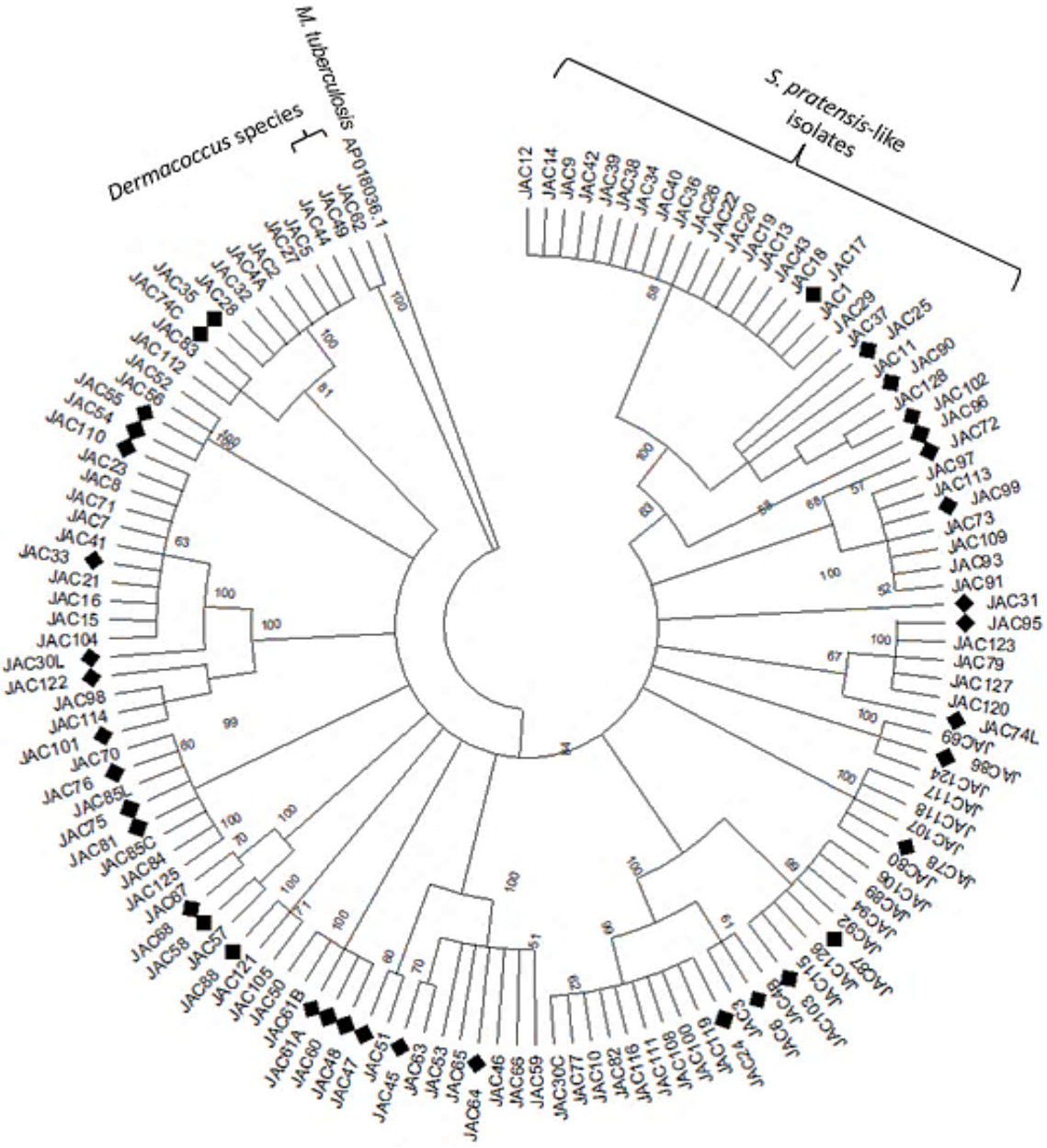


Figure 3.1.1: Phylogenetic comparison of *Streptomyces* isolates in this study. The tree consists of 133 *Streptomyces* isolates *rpoB* gene sequences condensed into 31 OTUs. The tree was constructed with 618 nucleotide sequence lengths using 100 bootstraps replicates. A minimum of 50% consensus cut-off value was used to create the tree branches. Isolates defined as one OTU *rpoB* sequences were verified to be 100% identity. ♦ denotes isolates selected for VOC analysis. *Mycobacterium tuberculosis* served as an outgroup.

3.2: The *Streptomyces* isolates had various phenotypes

The lawn morphology of each *Streptomyces* isolates grown on ISP-4 medium was determined, and they were tested on nutrient agar for apramycin, kanamycin, and thiostrepton susceptibility (Table 3.2.1). The isolates that displayed resistance to the respective antibiotics probably harbour antibiotic resistance genes (Van Hoek et al., 2011), which may allude to their ability to produce similar or related antibiotics (Peterson & Kaur, 2018).

The growth of the *Streptomyces* isolates on solid fermentation media used in this study was also tested. All *Streptomyces* grew successfully on the three media except for a few (Table 3.2.1). The media included SFM, used to encourage VOC production by *Streptomyces* species in other studies (Citron, Gleitzmann, et al., 2012; Citron, Rabe, et al., 2012; Komatsu et al., 2008; Rabe et al., 2013; Riclea et al., 2012); a synthetic and minimal medium used previously to promote geosmin production (Cane et al., 2006; Jáchymová et al., 2002); and YMS (Ikeda et al., 1987), a general *Streptomyces* specialized metabolite test medium.

Thirty-seven *Streptomyces* isolates representing each OTU were selected for VOC analysis (Figure 3.1.1). A maximum of three Streptomyces isolates was chosen from one OTU in a few cases as bacteria with identical *rpoB* gene sequences might have differences in specialized metabolism (Antony-Babu et al., 2017). Bacteria from OTUs containing single isolates were selected for VOC analysis regardless of their growth on all three-fermentation media used in the study. Growth on at least one of the fermentation media was sufficient for the selection of such isolates. Therefore, all phylogenetically different

Table 3.2.1: Characteristics of *Streptomyces* isolates from this study.

Isolate	Growth on ISP-4 medium				Antibiotic resistance	Growth on fermentation media		
	Growth	Sporul-ation	lawn color	diffusible pigment		SFM	Synthetic	YMS
JAC1	+++	+++	white	-	-	**	**	**
JAC2	++	+	cream	-	-	**	**	**
JAC3	+++	+++	cream	red-brown	-	**	**	**
JAC4A	+++	+++	white	-	-	**	**	**
JAC4B	+++	+++	dark yellow and mucoid	red-brown	-	**	**	**
JAC5	+	+	translucent	-	-	-	-	-
JAC6	+++	+++	cream	red-brown	-	**	**	**
JAC7	++	+	cream and translucent	brown	-	**	**	**
JAC8	+++	+	cream	-	-	**	*	**
JAC9	+++	+++	white	-	-	**	**	**
JAC10	+++	+++	cream	red-brown	-	**	**	**
JAC11	+++	+++	white	-	-	**	**	**
JAC12	+++	+++	white	-	-	**	**	**
JAC13	+++	+++	white	-	-	**	**	**
JAC14	+++	++	white	-	-	**	**	**
JAC15	++	+	cream	brown	-	**	**	**
JAC16	++	+	cream	brown	-	**	**	**
JAC17	+++	+++	white	-	-	**	**	**
JAC18	+++	++	white	-	-	**	**	**
JAC19	+++	++	white	-	-	**	**	**
JAC20	+++	+++	white	-	-	**	**	**
JAC21	+++	+	cream	brown	-	-	-	*
JAC22	+++	+++	white	-	-	**	**	**
JAC23	+++	+++	cream	brown	-	**	**	**
JAC24	+++	+++	cream	brown	-	**	**	**
JAC25	+++		white	-	-	**	**	**
JAC26	+++	+++	white	-	-	**	**	**
JAC27	+++	+++	white	-	-	**	**	**
JAC28	+++	+++	white	-	-	**	**	**
JAC29	+++	+++	white	-	-	**	**	**
JAC30C	+++	+++	purple	dark purple	-	**	**	**
JAC30L	+++	+	white	-	-	**	**	**

JAC31	+++		cream	-	-	**	**	**
JAC32	+++	+++	grey	-	-	**	**	**
JAC33	+++	+	white	-	-	**	**	**
JAC34	+++	+++	grey	-	-	**	**	**
JAC35	+++	+++	white	slight orange	-	**	**	**
JAC36	+++	+++	grey	-	-	**	**	**
JAC37	+++	+++	grey	-	-	**	**	**
JAC38	+++	+++	grey	-	-	**	**	**
JAC39	+++	+++	grey	-	-	**	**	**
JAC40	+++	+++	grey	-	-	**	**	**
JAC41	+++	+++	white	dark purple	-	**	**	**
JAC42	+++	+++	white with grey spores	-	-	**	**	**
JAC43	+++	+++	white with grey spores	-	-	**	**	**
JAC44	+++	+++	white	-	-	**	**	**
JAC45	+++	+++	white with grey spores	pink	-	**	**	**
JAC46	+++	+++	white with grey spores	pink	-	**	**	**
JAC47	+++	+++	white with grey spores	pink	-	**	**	**
JAC48	+++	+++	black	golden yellow	-	**	**	**
JAC49	+++	+++	yellow and mucoid	-	thio	**	**	**
JAC50	+++	+++	white with grey spores	golden yellow	-	**	**	**
JAC51	+++	+	grey	pink	-	**	**	**
JAC52	+++	++	grey	-	-	**	**	**
JAC53	+++	++	white	slight pink	-	**	**	**
JAC54	+++	++	grey	-	-	**	**	**
JAC55	+++	++	grey	-	-	**	**	**
JAC56	+++	++	grey	-	-	**	**	**
JAC57	+++	++	grey	-	-	**	**	**
JAC58	+++	+++	grey	-	-	**	**	**
JAC59	+++	+++	grey	slight pink	-	**	**	**
JAC60	+++	+++	slight pink with grey spores	golden yellow	-	**	**	**
JAC61A	+++	+++	brown	yellow	-	**	**	**
JAC61B	+++	-	orange, sporulation not obvious	orange	-	**	**	**

JAC62	+++	+	orange and mucoid	-	apra, kan, thio	**	**	**
JAC63	+++	++	grey	slight pink	-	**	**	**
JAC64	+++	+++	grey	slight	-	**	**	**
JAC65	+++	+++	white with grey spores	slight orange/pink	-	**	**	**
JAC66	+++	+++	white with grey spores	slight orange/pink	-	**	**	**
JAC67	+++	+++	grey	-	-	**	**	**
JAC68	+++	+++	grey	-	-	**	**	**
JAC69	+++	++	white	-	-	**	**	**
JAC70	+++	+++	grey	dark green/brown	-	**	**	**
JAC71	+++	+	grey	brown/slight orange	-	**	**	**
JAC72	+++	+++	white	slight yellow	-	**	**	**
JAC73	+++	+++	white with grey spores	yellow	-	**	**	**
JAC74C	+++	+++	dark grey	dark green	-	**	**	**
JAC74L	+++	+++	light grey	-	-	**	**	**
JAC75	+++	+++	green with grey spores	brown	-	**	**	**
JAC76	+++	+++	white with grey spores	-	-	**	**	**
JAC77	+++	+++	grey	dark purple	-	**	**	**
JAC78	+++	+++	white	-	-	**	**	**
JAC79	++	+	white and translucent	brown	-	**	**	**
JAC80	+++	+++	white and translucent	-	-	-	-	**
JAC81	+++	++	grey	brown	-	**	**	**
JAC82	+++	+++	grey	dark purple	-	**	**	**
JAC83	+++	+++	grey	black/ dark green	-	**	**	**
JAC84	+++	++	white with grey spores	brown	-	**	**	**
JAC85C	++	++	white	brown	-	**	**	**
JAC85L	+++	+++	dark grey with light grey spores	dark green	-	**	**	**
JAC86	+++	+++	black with grey spores	slight pink	-	**	**	**
JAC87	+++	+++	dark grey	black/ grey	-	**	**	**
JAC88	++	+	white and translucent	-	-	**	**	**
JAC89	+++	+++	black and mucoid	black/ grey	-	**	**	**

JAC90	++	+++	dark grey	-	-	**	**	**
JAC91	+++	+++	white with grey spores	yellow	-	**	**	**
JAC92	+++	+++	dark grey	grey/ black	-	**	**	**
JAC93	+++	+++	white with grey spores	yellow	-	**	**	**
JAC94	+++	+++	grey and translucent	grey	-	**	**	**
JAC95	++	+	white and translucent	-	-	**	**	**
JAC96	++	+	white	-	-	-	*	**
JAC97	+++	++	white	yellow	-	**	**	**
JAC98	+	+	translucent	-	-	**	*	*
JAC99	+++	++	white	yellow	-	**	**	**
JAC100	+++	+++	dark grey	dark purple	-	**	**	**
JAC101	+	+	grey	-	-	**	**	**
JAC102	+++	+++	grey	-	-	**	**	**
JAC103	+++	+++	translucent	brown	-	**	**	**
JAC104	+	+	white and translucent	-	-	**	**	**
JAC105	+++	+++	white	slight pink	-	**	**	**
JAC106	+++	+++	dark grey	dark green/ black	-	**	**	**
JAC107	+++	+++	grey	slight orange	-	**	**	**
JAC108	+++	+	red and mucoid	dark purple	thio	**	**	**
JAC109	+++	++	white	yellow	-	**	**	**
JAC110	+++	+++	grey	-	thio	**	**	**
JAC111	+++	+++	red and mucoid	dark purple	thio	**	**	**
JAC112	+++	+++	grey	black/ grey	-	**	-	**
JAC113	+++	++	white	yellow	-	**	**	**
JAC114	++	+	grey	-	-	**	*	**
JAC115	+++	+++	dark grey	black grey	-	**	**	**
JAC116	+++	+++	red	dark purple	kan, thio	**	**	**
JAC117	+++	+++	white	slight orange	-	-	-	**
JAC118	+++	+++	white	slight orange	-	-	-	**
JAC119	+++	+	brown and translucent	brown	-	**	**	**
JAC120	+++	+++	dark grey	yellow	apra	**	**	**
JAC121	+++	+	translucent	-	-	**	**	**
JAC122	+++	+	yellow and mucoid	yellow	kan	**	**	**

JAC123	+++	+++	grey	red-brown	-	**	**	**
JAC124	+++	+++	white	slight yellow	-	*	-	**
JAC125	+++	+++	light grey	slight brown	-	**	**	**
JAC126	+++	+	white	slight yellow	-	**	**	**
JAC127	+++	+++	white	-	-	**	**	**
JAC128	+++	++	translucent	slight green	-	**	**	**

\+ poor growth or sporulation on ISP-4 medium.
++ some (not good but not poor) growth or sporulation on ISP-4 medium.
+++ good growth or sporulation on ISP-4 medium.
\- antibiotic susceptibility, no diffusible pigment, or no growth on selected fermentation media.
** good growth observed on select fermentation media.
* poor growth on select fermentation media.
apra means resistance to apramycin antibiotic at 50 µg/ml.
kan means resistance to kanamycin antibiotic at 50 µg/ml.
thio means resistance to thiostrepton antibiotic at 25 µg/ml.

Streptomyces isolates were selected to capture the full range of VOCs that the collection of bacteria could produce.

3.3: VOC analysis

3.3.1: Screening of VOC production by select *Streptomyces* isolates

The 37 *Streptomyces* isolates representing each OTU were grown in the three fermentation media, SFM, YMS, and the synthetic medium. The cultures were pooled together for VOC collection and GC-MS analysis. The VOCs identifications were primarily based on spectra availability in the NIST library, and the stereochemistry assigned should be interpreted with caution. It was assumed that not all compounds present in the samples were identified due to limitations in the identification procedure. One hundred and sixty-three bacterially derived VOCs were detected from the 37 *Streptomyces* isolates and were further classified based on the level of confidence in their identities (Figure 3.3.1.1 and Table 3.3.1.1). Those compounds with a high confidence level identity were further divided into major VOC classes, of which the alcohols were the most common (Table 3.3.1.2). There were more unknown compounds than those identified with high or low confidence, where 116 VOCs did not have published RI values for comparison. Therefore, there may be novel VOCs in the 116 unknown compounds with potentially useful bioactivities. Forty-nine of the bacterially derived VOCs were produced by isolates from more than one OTU, and 22 of those were from two of the 30 OTUs in the current study. Geosmin, [4a(2H)-Naphthalenol, octahydro-4,8a-dimethyl-,(4.alpha.,4a.alpha.,8a.beta.)- or trans-1,10-Dimethyl-trans-9-decalol], was the most common VOCs detected in 23 isolates (VOC 27 in Table 3.3.1.1). Compounds detected in single OTUs totalled 114 (Table 3.3.1.2),

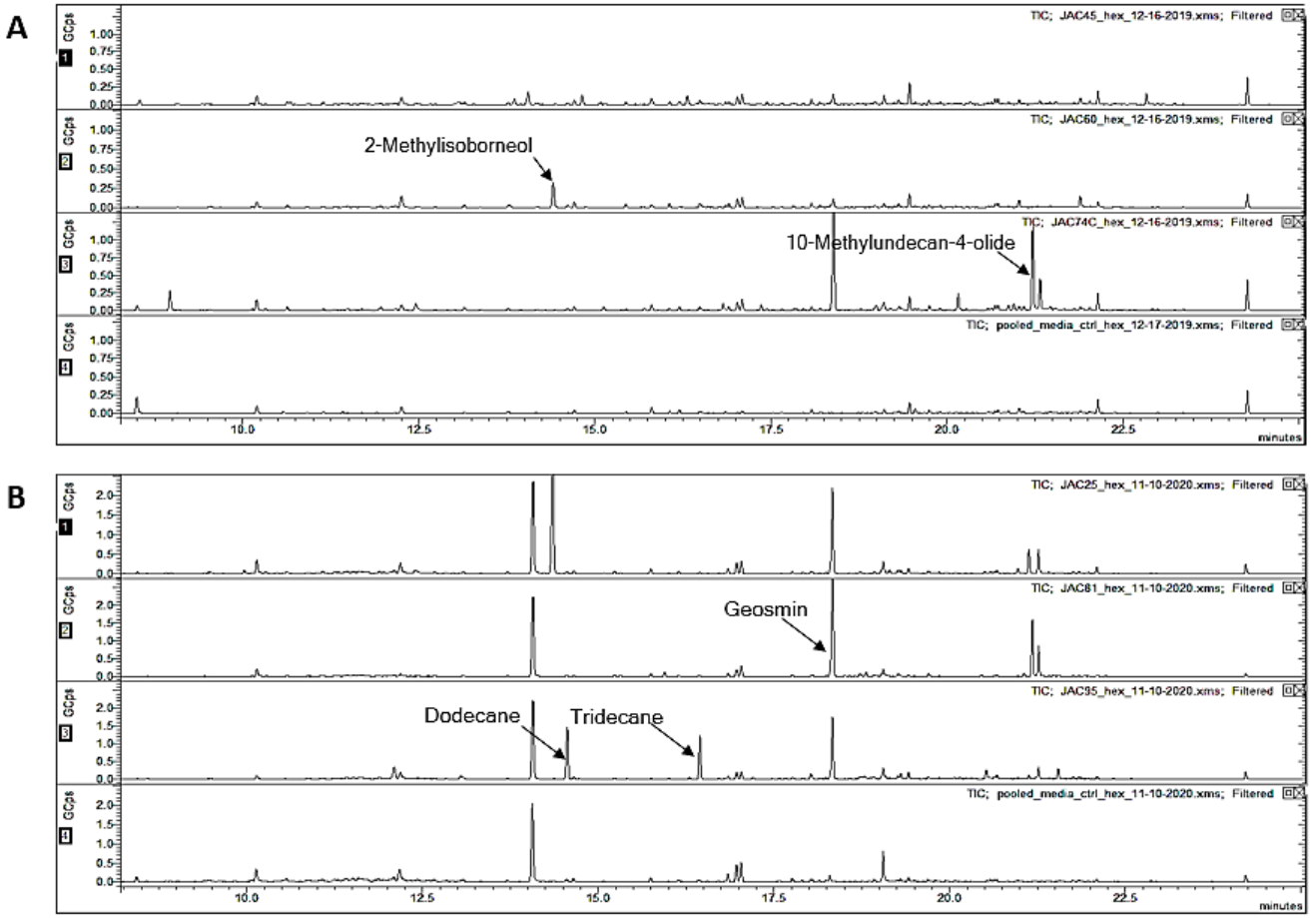


Figure 3.3.1.1: Example view (~9- 23 min) of pooled culture GC-MS chromatograms for six *Streptomyces* isolates. Some prominent peaks with >38% identity to NIST reference spectra are labelled. GC-MS analysis was done for cultures fermented in two batches of media for isolates (top to bottom) **A.** JAC45, JAC60, JAC74C, and a pooled media control **B.** JAC25, JAC81, JAC95, and a pooled media control.

Table 3.3.1.1: Bacterially derived compounds detected in pooled culture VOC analysis.

No.	VOC [a]	IUPAC name	RT	RT Stdev [b]	Detected in	Confidence in the VOC ID [c]	Detected in >1 OTU [d]
1	(R)-(+)-3-Methylcyclopentanone	(3*R*)-3-methylcyclopentan-1-one	5.742	0.000	JAC6	high	no
2	1-Hexanol	hexan-1-ol	6.223	0.034	JAC60, JAC61A	high	no
3	1-Octen-3-ol	oct-1-en-3-ol	8.963	0.000	JAC74C	high	no
4	3-Octanone	octan-3-one	9.038	0.000	JAC122	high	no
5	Ethanol, 2-(2-ethoxyethoxy)-	2-(2-ethoxyethoxy)ethanol	9.425	0.000	JAC17	high	no
6	2-Methyl-2-bornene	1,2,7,7-tetramethylbicyclo[2.2.1]hept-2-ene	9.964	0.000	JAC25	high	no
7	2,6-Octadienoic acid, 3,7-dimethyl-, methyl ester	methyl (2E)-3,7-dimethylocta-2,6-dienoate	16.820	0.000	JAC74C	high	no
8	2,4-Furandicarboxylic acid, dimethyl ester	dimethyl furan-2,4-dicarboxylate	17.364	0.000	JAC74C	high	no
9	1H-Benzocycloheptene, 2,4a,5,6,7,8,9,9a-octahydro-3,5,5-trimethyl-9-methylene-, (4aS-cis)- Alpha-Himachalene [e]	3,5,5-trimethyl-9-methylidene-2,4a,6,7,8,9a-hexahydro-1H-benzo[7]annulene	18.341	0.000	JAC102	high	no
10	1-Dodecanol	dodecan-1-ol	19.054	0.000	JAC68	high	no
11	Dodecanoic acid, methyl ester	methyl dodecanoate	19.695	0.000	JAC35	high	no
12	Cyclohexanemethanol, 4-ethenyl-.alpha.,.alpha.,4-trimethyl-3-(1-methylethenyl)-, [1R-(1.alpha.,3.alpha.,4.beta.)]- Elemol group [e]	2-(4-ethenyl-4-methyl-3-prop-1-en-2-ylcyclohexyl)propan-2-ol	20.162	0.000	JAC74C	high	no

13	Pentanoic acid, 2,2,4-trimethyl-3-carboxyisopropyl, isobutyl ester	1-O-(2-methylpropyl) 4-O-propan-2-yl 2,2-dimethyl-3-propan-2-ylbutanedioate	20.535	0.000	JAC122	high	no
14	Cetene	hexadec-1-ene	20.633	0.000	JAC47	high	no
15	DL-2,3-Butanediol	butane-2,3-diol	4.894	0.003	JAC17, JAC47	high	yes
16	Hexanenitrile	hexanenitrile	6.403	0.025	JAC60, JAC78, JAC61A	high	yes
17	Pyrazine, 2,5-dimethyl-	2,5-dimethylpyrazine	7.230	0.012	JAC17, JAC24, JAC31, JAC33, JAC47, JAC64, JAC78, JAC74C, JAC25, JAC35, JAC96, JAC110, JAC74L, JAC61A, JAC30L, JAC87, JAC81	high	yes
18	Methyl 2-methylhexanoate	methyl 2-methylhexanoate	8.537	0.001	JAC45, JAC64	high	yes
19	Benzyl alcohol	phenylmethanol	10.290	0.027	JAC17, JAC58, JAC25, JAC122, JAC30L	high	yes
20	1-Octanol	octan-1-ol	11.282	0.002	JAC99, JAC101, JAC64	high	yes
21	Undecane	Undecane	12.099	0.001	JAC95, JAC6, JAC122	high	yes
22	Phenylethyl Alcohol	2-phenylethanol	12.432	0.032	JAC17, JAC31, JAC86, JAC101, JAC33, JAC58, JAC74C, JAC25, JAC110, JAC74L, JAC30L	high	yes
23	Cyclopropane, pentyl-	pentylcyclopropane	12.917	0.001	JAC99, JAC101	high	yes

24	2-Methylisoborneol	1,2,7,7-tetramethylbicyclo[2.2.1]heptan-2-ol	14.380	0.032	JAC102, JAC99, JAC31, JAC101, JAC60, JAC78, JAC25, JAC72, JAC55, JAC54, JAC61A, JAC90	high	yes
25	Dodecane	dodecane	14.560	0.001	JAC95, JAC6, JAC122, JAC68	high	yes
26	heptylcyclohexane	heptylcyclohexane	17.204	0.000	JAC95, JAC6, JAC122, JAC68	high	yes
27	4a(2H)-Naphthalenol, octahydro-4,8a-dimethyl-,(4.alpha.,4a.alpha.,8a.beta.)- trans-1,10-Dimethyl-trans-9-decalol Geosmin [e]	(4S,4aS,8aR)-4,8a-dimethyl-1,2,3,4,5,6,7,8-octahydronaphthalen-4a-ol	18.356	0.014	JAC95, JAC35, JAC88, JAC55, JAC110, JAC74L, JAC54, JAC61A, JAC75, JAC90, JAC87, JAC78, JAC25, JAC81, JAC102, JAC31, JAC86, JAC101, JAC33, JAC76, JAC47, JAC64, JAC74C	high	yes
28	1,6-Cyclodecadiene, 1-methyl-5-methylene-8-(1-methylethyl)-, [S-(E,E)]- Germacrene D [e]	(1Z,6Z)-1-methyl-5-methylidene-8-propan-2-ylcyclodeca-1,6-diene	19.270	0.002	JAC25, JAC81	high	yes
29	epi-cubenol	(1S,4R,4aS,8aR)-4,7-dimethyl-1-propan-2-yl-2,3,4,5,6,8a-hexahydro-1H-naphthalen-4a-ol	21.128	0.001	JAC25, JAC95, JAC90, JAC87	high	yes

30	1-Naphthalenol, 1,2,3,4,4a,7,8,8a-octahydro-1,6-dimethyl-4-(1-methylethyl)-, [1R-(1.alpha.,4.beta.,4a.beta.,8a.beta.)]- Cadinol [e]	(1R,4S)-1,6-dimethyl-4-propan-2-yl-3,4,4a,7,8,8a-hexahydro-2H-naphthalen-1-ol	21.182	0.003	JAC17, JAC102, JAC24, JAC86	high	yes
31	cubenol	(1S,4R,4aR,8aR)-4,7-dimethyl-1-propan-2-yl-2,3,4,5,6,8a-hexahydro-1H-naphthalen-4a-ol	21.277	0.024	JAC47, JAC25, JAC110, JAC74L, JAC81	high	yes
32	2-Naphthalenemethanol, decahydro-α,α,4a-trimethyl-8-methylene-, [2R-(2α,4aα,8aβ)]- β-Selinenol [e]	2-(4*a*-methyl-8-methylidene-1,2,3,4,5,6,7,8*a*-octahydronaphthalen-2-yl)propan-2-ol	21.320	0.001	JAC86, JAC101, JAC33, JAC76, JAC74C	high	yes
33	Methyl tetradecanoate	methyl tetradecanoate	22.094	0.001	JAC35, JAC110, JAC74L	high	yes
34	Toluene-Dup1 [f]	[g]	4.322	0.000	JAC103	low	no
35	3-Hexanone, 5-methyl-	5-methylhexan-3-one	5.486	0.000	JAC64	low	no
36	Cyclohexane, 1,1,2-trimethyl-	1,1,2-trimethylcyclohexane	6.691	0.000	JAC99	low	no
37	7-Tetradecene	(E)-tetradec-7-ene	19.307	0.000	JAC35	low	no
38	10-Methylundecan-4-olide	5-(6-methylheptyl)oxolan-2-one	21.220	0.000	JAC74C	low	no
39	7-epi-.alpha.-eudesmol	2-[(2S,4aR)-4a,8-dimethyl-2,3,4,5,6,8a-hexahydro-1H-naphthalen-2-yl]propan-2-ol	21.599	0.000	JAC86	low	no
40	10-epi-.gamma.-Eudesmol	2-(4a,8-dimethyl-2,3,4,5,6,7-hexahydro-1H-naphthalen-2-yl)propan-2-ol	21.655	0.000	JAC31	low	no
41	10-Methyldodecan-5-olide	6-(5-methylheptyl)oxan-2-one	22.831	0.000	JAC45	low	no
42	3-Hexanone	hexan-3-one	4.493	0.030	JAC99, JAC76, JAC48, JAC6	low	yes

43	Cyclopentanol, 1-methyl-	1-methylcyclopentan-1-ol	4.676	0.012	JAC99, JAC103, JAC76, JAC48	low	yes
44	Dodecane, 2,6,10-trimethyl-	2,6,10-trimethyldodecane	18.875	0.001	JAC6, JAC68, JA95	low	yes
	Farnesane [e]						
45	Alloaromadendrene	1,1,7-trimethyl-4-methylidene-2,3,4a,5,6,7,7a,7b-octahydro-1aH-cyclopropa[e]azulene	21.265	0.002	JAC88, JAC75, JAC87	low	yes
46	7-epi-.alpha.-selinene	5,8a-dimethyl-3-prop-1-en-2-yl-2,3,4,4a,7,8-hexahydro-1H-naphthalene	21.544	0.000	JAC95, JAC87	low	yes
47	1-Nonadecene	nonadec-1-ene	21.757	0.000	JAC110, JAC74L	low	yes
48	Propanedioic acid, (bromomethyl)methyl-, bis(1,1-dimethylethyl) ester	ditert-butyl 2-(bromomethyl)-2-methylpropanedioate	4.019	0.000	JAC61A	unknown	no
49	t-Butyl cyclopentaneperoxycarboxylate	tert-butyl cyclopentanecarboperoxoate	4.026	0.000	JAC96	unknown	no
50	t-Butyl cyclopentaneperoxycarboxylate-Dup1 [f]	[g]	4.057	0.000	JAC96	unknown	no
51	1-n-Butoxy-2,3-dimethyldiaziridine	1-butoxy-2,3-dimethylaziridine	4.065	0.000	JAC75	unknown	no
52	(S)-3,4-Dimethylpentanol	(3S)-3,4-dimethylpentan-1-ol	4.129	0.000	JAC122	unknown	no
53	1-Heptene, 3-methoxy-	3-methoxyhept-1-ene	4.268	0.000	JAC74C	unknown	no
54	Propanoic acid, 2-methyl-, 2-ethyl-1-propyl-1,3-propanediyl ester	[2-ethyl-3-(2-methylpropanoyloxy)hexyl] 2-methylpropanoate	4.273	0.000	JAC64	unknown	no
55	2-Hexanone, 6-methoxy-	6-methoxyhexan-2-one	4.290	0.000	JAC6	unknown	no
56	2-Buten-1-ol, 2-methyl-	(E)-2-methylbut-2-en-1-ol	4.316	0.000	JAC74C	unknown	no
57	Sulfurous acid, isobutyl pentyl ester	2-methylpropyl pentyl sulfite	4.485	0.000	JAC103	unknown	no

58	2-Hexanone	hexan-2-one	4.523	0.000	JAC6	unknown	no
59	(2,3,3-Trimethyloxiranyl)methanol	(2,3,3-trimethyloxiran-2-yl)methanol	4.714	0.000	JAC88	unknown	no
60	meso-3,4-Hexanediol	hexane-3,4-diol	4.724	0.000	JAC30L	unknown	no
61	(S)-3,4-Dimethylpentanol	(3S)-3,4-dimethylpentan-1-ol	4.744	0.000	JAC87	unknown	no
62	2-Methyl-2,3-pentanediol	2-methylpentane-2,3-diol	4.757	0.000	JAC17	unknown	no
63	2-Hexanol	hexan-2-ol	4.766	0.000	JAC88	unknown	no
64	DL-2,3-Butanediol	butane-2,3-diol	4.798	0.000	JAC103	unknown	no
65	Methyl 2-methoxypropenoate	methyl 2-methoxyprop-2-enoate	4.824	0.000	JAC30L	unknown	no
66	Propanoic acid, 2,2-dimethyl-, 2,4-dinitrophenyl ester	(2,4-dinitrophenyl) 2,2-dimethylpropanoate	5.488	0.000	JAC45	unknown	no
67	2-Furanmethanol	furan-2-ylmethanol	5.826	0.000	JAC30L	unknown	no
68	2-Butenoic acid, 2-methyl-, methyl ester	methyl (E)-2-methylbut-2-enoate	6.158	0.000	JAC74C	unknown	no
69	Butanedioic acid, phenyl-	2-phenylbutanedioic acid	6.733	0.000	JAC88	unknown	no
70	Vinyl butyrate	ethenyl butanoate	6.751	0.000	JAC61A	unknown	no
71	Carbonic acid, bis(1-methylethyl) ester	dipropan-2-yl carbonate	6.791	0.000	JAC33	unknown	no
72	4-Octanol, 2-methyl-	2-methyloctan-4-ol	6.918	0.000	JAC99	unknown	no
73	1,6:3,4-Dianhydro-2-deoxy-.beta.-d-ribo-hexopyranose	3,7,9-trioxatricyclo[4.2.1.02,4]nonane	8.253	0.000	JAC88	unknown	no
74	Dimethyl trisulfide	(methyltrisulfanyl)methane	8.636	0.000	JAC90	unknown	no
75	Formic acid, heptyl ester	heptyl formate	8.714	0.000	JAC99	unknown	no
76	3(2H)-Thiophenone, dihydro-2-methyl-	2-methylthiolan-3-one	9.139	0.000	JAC96	unknown	no
77	3-Octanol	octan-3-ol	9.342	0.000	JAC122	unknown	no
78	4-Octene, 2,3,7-trimethyl-, [S-(E)]-	(E,3S)-2,3,7-trimethyloct-4-ene	9.601	0.000	JAC25	unknown	no

79	Tricyclo[2.2.1.0(1,4)]heptan-2-one, 6-nitro-	6-nitrotricyclo[2.2.1.01,4]heptan-2-one	9.615	0.000	JAC58	unknown	no
80	4-Nonene, 3-methyl-, (Z)-	3-methylnon-4-ene	10.026	0.000	JAC31	unknown	no
81	Tricyclo[3.2.1.0(2,4)]octane, 8-methylene-, (1.alpha.,2.alpha.,4.alpha.,5.alpha.)-	8-methylidenetricyclo[3.2.1.02,4]octane	10.075	0.000	JAC35	unknown	no
82	6-Azabicyclo[3,2,0]heptan-7-one	6-azabicyclo[3.2.0]heptan-7-one	10.079	0.000	JAC61A	unknown	no
83	Cyclopropane, 1-butyl-2-(2-methylpropyl)-	1-butyl-2-(2-methylpropyl)cyclopropane	10.337	0.000	JAC87	unknown	no
84	Propanoic acid, 3-chloro-, 4-formylphenyl ester	(4-formylphenyl) 3-chloropropanoate	10.547	0.000	JAC54	unknown	no
85	DL-4,5-Octanediol	octane-4,5-diol	11.475	0.000	JAC17	unknown	no
86	6-Dodecene, (E)-	(E)-dodec-6-ene	12.024	0.000	JAC87	unknown	no
87	1-Octanol, 2-butyl-	2-butyloctan-1-ol	12.098	0.000	JAC35	unknown	no
88	Hydrazine, (phenylmethyl)-	benzylhydrazine	12.396	0.000	JAC55	unknown	no
89	Bicyclo[3.2.0]hepta-2,6-diene	bicyclo[3.2.0]hepta-2,6-diene	12.400	0.000	JAC90	unknown	no
90	Cyclopropane, 1-butyl-2-(2-methylpropyl)-	1-butyl-2-(2-methylpropyl)cyclopropane	13.103	0.000	JAC86	unknown	no
91	Iron, tricarbonyl[(2,3,4,5-.eta.)-2,4-cycloheptadien-1-ol]-	[g]	13.856	0.000	JAC45	unknown	no
92	Benzene, (iodomethyl)-	iodomethylbenzene	13.969	0.000	JAC30L	unknown	no
93	Benzene, [(methylsulfonyl)methyl]-	methylsulfonylmethylbenzene	14.023	0.000	JAC58	unknown	no
94	Iron, tricarbonyl[(2,3,4,5-.eta.)-2,4-cycloheptadien-1-ol]--Dup1 [f]	[g]	14.818	0.000	JAC45	unknown	no
95	Sydnone, 3-(phenylmethyl)-	3-benzyl-2H-oxadiazol-3-ium-5-one	14.820	0.000	JAC64	unknown	no
96	Benzothiazole	1,3-benzothiazole	15.082	0.000	JAC61A	unknown	no

97	6-Octen-1-ol, 3,7-dimethyl-, formate	3,7-dimethyloct-6-enyl formate	15.123	0.000	JAC74C	unknown	no
98	(2S,4R)-p-Mentha-[1(7),8]-diene 2-hydroperoxide	2-hydroperoxy-1-methylidene-4-prop-1-en-2-ylcyclohexane	15.292	0.000	JAC86	unknown	no
99	(1,2,2-trimethyl-3-cyclopenten-1-yl)acetaldehyde	2-(1,2,2-trimethylcyclopent-3-en-1-yl)acetaldehyde	15.315	0.000	JAC81	unknown	no
100	Phenol, 4-[2-(methylamino)ethyl]-	4-[2-(methylamino)ethyl]phenol	15.317	0.000	JAC25	unknown	no
101	Bicyclo[3.1.1]heptane, 6,6-dimethyl-2-methylene-, (1S)- β-pinene [e]	6,6-dimethyl-2-methylidenebicyclo[3.1.1]heptane	15.528	0.000	JAC122	unknown	no
102	2-Hydroxy-5-ethyl-5-methylcyclopent-2-en-1-one	5-ethyl-2-hydroxy-5-methylcyclopent-2-en-1-one	15.944	0.000	JAC81	unknown	no
103	3-Hexadecene, (Z)-	(Z)-hexadec-3-ene	16.299	0.000	JAC122	unknown	no
104	1-Tridecene	tridec-1-ene	16.300	0.000	JAC6	unknown	no
105	Tricyclo[4.1.0.0(2,7)]heptane	tricyclo[4.1.0.02,7]heptane	16.311	0.000	JAC47	unknown	no
106	Phenylacetaldehyde N-methyl-N-formylhydrazone	N-methyl-N-[(E)-2-phenylethylideneamino]formamide	16.707	0.000	JAC25	unknown	no
107	1-Methylene-2b-hydroxymethyl-3,3-dimethyl-4b-(3-methylbut-2-enyl)-cyclohexane	[2,2-dimethyl-4-(3-methylbut-2-enyl)-6-methylidenecyclohexyl]methanol	18.507	0.000	JAC81	unknown	no
108	1-Methylene-2b-hydroxymethyl-3,3-dimethyl-4b-(3-methylbut-2-enyl)-cyclohexane-Dup1 [f]	[g]	18.575	0.000	JAC81	unknown	no
109	(3R,2E)-2-(Hexadec-15-ynylidene)-3-hydroxy-4-methylenebutanolide	(3E,4R)-3-hexadec-15-ynylidene-4-hydroxy-5-methylideneoxolan-2-one	18.654	0.000	JAC81	unknown	no
110	1,4-Dimethyladamantane #	1,4-dimethyladamantane	18.726	0.000	JAC95	unknown	no

111	Cyclobutaneacetonitrile, 1-methyl-2-(1-methylethylidene)-	2-(1-methyl-2-propan-2-ylidenecyclobutyl)acetonitrile	18.780	0.000	JAC86	unknown	no
112	Pyrrol-2(5H)-one, 4-acetyl-3-hydroxy-5-(3-nitrophenyl)-1-[2-(1-piperazinyl)ethyl]-	3-acetyl-4-hydroxy-2-(3-nitrophenyl)-1-(2-piperazin-1-ylethyl)-2H-pyrrol-5-one	18.795	0.000	JAC87	unknown	no
113	4,7-Methanoazulene, 1,2,3,4,5,6,7,8-octahydro-1,4,9,9-tetramethyl-, [1S-(1.alpha.,4.alpha.,7.alpha.)]- β-Patchoulene [e]	1,5,11,11-tetramethyltricyclo[6.2.1.02,6]undec-2(6)-ene	18.812	0.000	JAC81	unknown	no
114	n-Hexyl methylphosphonofluoridate	1-[fluoro(methyl)phosphoryl]oxyhexane	18.841	0.000	JAC86	unknown	no
115	trans,trans- and trans,cis-1,8-Dimethylspiro[5.5]undecane	4,11-dimethylspiro[5.5]undecane	18.950	0.000	JAC86	unknown	no
116	3-Isopropoxy-1,1,1,7,7,7-hexamethyl-3,5,5-tris(trimethylsiloxy)tetrasiloxane	propan-2-yl bis(trimethylsilyl) tris(trimethylsilyloxy)silyl silicate	18.960	0.000	JAC87	unknown	no
117	5,10-Pentadecadiyn-1-ol	pentadeca-5,10-diyn-1-ol	18.969	0.000	JAC81	unknown	no
118	tert-Butyl 3-methylbutyl disulfide	1-(tert-butyldisulfanyl)-3-methylbutane	18.977	0.000	JAC33	unknown	no
119	2H-Benzocyclohepten-2-one, decahydro-4a-methyl-, trans-	4a-methyl-3,4,5,6,7,8,9,9a-octahydro-1H-benzo[7]annulen-2-one	19.147	0.000	JAC25	unknown	no
120	6-Tridecene	(E)-tridec-6-ene	19.307	0.000	JAC88	unknown	no
121	Dichlorphen, O,O'-bis(4-fluoro-2-trifluoromethylbenzoyl)-	[4-chloro-2-[[5-chloro-2-[4-fluoro-2-(trifluoromethyl)benzoyl]oxyphenyl]methyl]phenyl] 4-fluoro-2-(trifluoromethyl)benzoate	19.496	0.000	JAC61A	unknown	no

122	Benzoic acid, 4-ethoxy-, ethyl ester	ethyl 4-ethoxybenzoate	19.728	0.000	JAC88	unknown	no
123	Cyclohexanol, 4-ethyl-4-methyl-3-(1-methylethyl)-, (1.alpha.,3.alpha.,4.beta.)-	4-ethyl-4-methyl-3-propan-2-ylcyclohexan-1-ol	20.279	0.000	JAC87	unknown	no
124	2,7-Octadiene-1,6-diol, 2,6-dimethyl-, (E)-	(2E)-2,6-dimethylocta-2,7-diene-1,6-diol	20.454	0.000	JAC81	unknown	no
125	Naphthalene, 1,2,4a,5,6,8a-hexahydro-4,7-dimethyl-1-(1-methylethyl)- α-Amorphene [e]	4,7-dimethyl-1-propan-2-yl-1,2,4a,5,6,8a-hexahydronaphthalene	20.502	0.000	JAC25	unknown	no
126	1-Hexyl-2-nitrocyclohexane	1-hexyl-2-nitrocyclohexane	20.947	0.000	JAC74C	unknown	no
127	Hexanethioic acid, S-methyl ester	S-methyl hexanethioate	21.058	0.000	JAC81	unknown	no
128	Dihydro-cis-.alpha.-copaene-8-ol-Dup1 [f]	[g]	21.267	0.000	JAC95	unknown	no
129	3,4-Nonadien-6-yne, 5-ethyl-3-methyl-	5-Ethyl-3-methyl-3,4-nonadien-6-yne	21.482	0.000	JAC81	unknown	no
130	1-Allyl-cyclohexane-1,2-diol	1-prop-2-enylcyclohexane-1,2-diol	21.544	0.000	JAC81	unknown	no
131	10-Heneicosene (c,t)	(E)-henicos-10-ene	21.758	0.000	JAC25	unknown	no
132	Cyclohexene, 1,5,5-trimethyl-6-(2-propenylidene)-	(6Z)-1,5,5-trimethyl-6-prop-2-enylidenecyclohexene	21.921	0.000	JAC81	unknown	no
133	Benzoic acid, 3-fluoro-, 2-oxo-2-phenylethyl ester	phenacyl 3-fluorobenzoate	21.976	0.000	JAC25	unknown	no
134	Cyclohexanol, 1-ethyl-2,2-dimethyl-6-methylene-	1-ethyl-2,2-dimethyl-6-methylidenecyclohexan-1-ol	22.219	0.000	JAC81	unknown	no
135	5,6-Decadien-3-yne, 5,7-diethyl-	[g]	24.274	0.000	JAC35	unknown	no
136	1-Nitrosoadamantane	1-nitrosoadamantane	24.376	0.000	JAC35	unknown	no
137	5,10-Pentadecadien-1-ol, (Z,Z)-	(5Z,10Z)-pentadeca-5,10-dien-1-ol	25.467	0.000	JAC102	unknown	no
138	Hexanedioic acid, mono(2-ethylhexyl)ester	6-(2-ethylhexoxy)-6-oxohexanoic acid	28.637	0.000	JAC81	unknown	no

139	Phthalic acid, di(hept-3-yl) ester	diheptan-3-yl benzene-1,2-dicarboxylate	30.506	0.000	JAC81	unknown	no
140	Butane, 2-azido-2,3,3-trimethyl-	2-azido-2,3,3-trimethylbutane	4.022	0.000	JAC72, JAC75, JAC81	unknown	yes
141	3-Pentanol, 3-methyl-	3-methylpentan-3-ol	4.036	0.005	JAC99, JAC48	unknown	yes
142	Cyclobutanone, 2,3-dimethyl-, trans-	(2S,3R)-2,3-dimethylcyclobutan-1-one	4.061	0.006	JAC88, JAC30L	unknown	yes
143	Butanoic acid, 2-methyl-, methyl ester	methyl 2-methylbutanoate	4.359	0.005	JAC45, JAC47	unknown	yes
144	Propanoic acid, 2,2-dimethyl-, 2,4-dinitrophenyl ester	(2,4-dinitrophenyl) 2,2-dimethylpropanoate	4.361	0.001	JAC60, JAC64	unknown	yes
145	2-Hydroxy-3-pentanone	2-hydroxypentan-3-one	4.856	0.007	JAC88, JAC30L	unknown	yes
146	Acetamidoacetaldehyde	N-(2-oxoethyl)acetamide	7.523	0.000	JAC110, JAC74L	unknown	yes
147	Ether, 3-butenyl propyl	4-propoxybut-1-ene	7.524	0.002	JAC6, JAC88, JAC55	unknown	yes
148	(S)-3-Ethyl-4-methylpentanol	3-ethyl-4-methylpentan-1-ol	9.462	0.053	JAC25, JAC88, JAC61A	unknown	yes
149	Formic acid, 2-ethylhexyl ester	2-ethylhexyl formate	11.490	0.003	JAC99, JAC101, JAC64	unknown	yes
150	Benzoyl bromide	benzoyl bromide	11.954	0.004	JAC86, JAC76	unknown	yes
151	1,3-Methanopentalene, octahydro-	tricyclo[5.1.1.02,6]nonane	14.053	0.001	JAC45, JAC64	unknown	yes
152	Tricyclo[3.3.1.1(3,7)]decanone, 4-iodo-, (1.alpha.,3.beta.,4.beta.,5.alpha.,7.beta.)-	[g]	15.236	0.000	JAC25, JAC81	unknown	yes
153	6-Tridecene, (Z)-	(Z)-tridec-6-ene	16.301	0.001	JAC95, JAC68	unknown	yes
154	Bicyclo[3.2.0]hept-2-en-6-one, 7-chloro-	7-chlorobicyclo[3.2.0]hept-2-en-6-one	16.313	0.000	JAC45, JAC64	unknown	yes
155	Dodecanal	dodecanal	18.199	0.001	JAC24, JAC33	unknown	yes
156	trans-1,10-Dimethyl-trans-9-decalinol	8,8a-dimethyl-2,3,4,4a,5,6,7,8-octahydro-1H-naphthalen-1-ol	18.385	0.001	JAC24, JAC60, JAC58, JAC45	unknown	yes

157	3-Heptyne, 5-methyl-	5-methylhept-3-yne	18.726	0.001	JAC25, JAC81	unknown	yes
158	Cyclopentane, nonyl-	nonylcyclopentane	18.755	0.002	JAC95, JAC6, JAC122	unknown	yes
159	n-Heptyl methylphosphonofluoridate	1-[fluoro(methyl)phosphoryl]oxyheptane	18.818	0.032	JAC78, JAC95	unknown	yes
160	n-Heptadecylcyclohexane	heptadecylcyclohexane	20.143	0.000	JAC95, JAC6, JAC122, JAC68	unknown	yes
161	Dihydro-cis-.alpha.-copaene-8-ol	2,8-dimethyl-5-propan-2-yltricyclo[4.4.0.02,7]decan-4-ol	20.549	0.031	JAC24, JAC86, JAC95, JAC87	unknown	yes
162	β-Vatirenene	2-Isopropylidene-8,8a-dimethyl-1,2,3,7,8,8a-hexahydronaphthalene	21.177	0.002	JAC72, JAC75, JAC81	unknown	yes
163	5,10-Pentadecadiyn-1-ol	pentadeca-5,10-diyn-1-ol	21.247	0.025	JAC76, JAC90	unknown	Yes

[a] Compound names were assigned by the NIST library. This was used in conjunction with the retention time for more accurate identification of compounds, i.e., VOCs with the same name but different RT should be interpreted cautiously.

[b] Standard deviation from the average retention time (RT) for a given VOC.

[c] The level of confidence assigned to the identity of the VOC. High confidence means there was a good spectral match to the NIST library reference and a Kovats retention index (RI) within 10 points of a literature value available on the NIST WebBook (https://webbook.nist.gov/chemistry/name-ser/), low confidence means the difference in RI of a NIST library suggested VOC to that of the literature value is greater 10 points, and unknown means that a literature RI value could not be found for a NIST library suggested compound.

[d] Indicates if the VOC was from one or more isolates/OTU.

[e] common name for the given VOC.

[f] Dup1 suffix means a NIST reference spectra with close identity to the given VOC was not available in the library used in this study.

[g] not applicable, or the IUPAC name is not available.

Table 3.3.1.2: The numbers of bacterially derived VOCs detected in the pooled cultures.

Confidence in VOC identity	**No. of VOCs present in 1 OTU**	**No. of VOCs Present in >1 OTU**	**Total**
High	**14**	**19**	**33**
Alcohol	*4*	*4*	*8*
Aldehyde	*0*	*0*	*0*
Ketone	*2*	*0*	*2*
Ester	*4*	*2*	*6*
Hydrocarbon	*1*	*4*	*5*
Terpene	*2*	*1*	*3*
Terpenoid	*1*	*6*	*7*
Other	*0*	*2*	*2*
Low	**8**	**6**	**14**
Unknown	**92**	**24**	**116**
Total	**114**	**49**	**163**

indicating differences in the VOC production profiles of phylogenetically different isolates.

There was variability in VOCs produced among the isolates. For example, JAC81 produced the most (14) VOCs that were not produced by the other isolates, while nine isolates did not produce any unique VOCs (Table 3.3.1.3). Variability was also evident within the OTUs, where two or more isolates were selected for analysis. For example, OTU-22 comprising JAC75, JAC76, and JAC81 shows notable differences in the number of unique VOCs produced by each mentioned *Streptomyces* isolates. JAC81 produced 14 unique VOCs, while JAC75 and JAC76 produced one and no unique VOCs, respectively (Table 3.3.1.3). This event might mean that genes or BGCs for VOCs are present in JAC81 but absent in the others, though all three *Streptomyces* are from the same OTU. However, there could be variability in the growth of the isolates during fermentation. Similar results were seen for OTUs 13, 21, and 28, where the VOC production profiles by the isolates were inconsistent with each other.

Table 3.3.1.3: The number of VOCs detected from each *Streptomyces* isolate in pooled culture VOC analysis.

***Streptomyces* isolate**	**OTU [a]**	**Empty jar VOCs [b]**	**Medium VOCs [c]**	**Bacteria VOCs [d]**	**Total VOC**	**Unique bacteria VOCs [e]**
Empty jar (Dec 2019)	[g]	26	[g]	[g]	26	[g]
empty jar (2020)	[g]	83	[g]	[g]	83	[g]
Medium (Dec 2019)	[g]	15	2	[g]	17	[g]
Medium (2020)	[g]	38	13	[g]	51	[g]
Medium_2 (2020)	[g]	40	12	[g]	52	[g]
JAC17	1	22	0	8	30	3
JAC25 [f]	2	46	11	20	77	8
JAC90	3	50	6	6	62	2
JAC102	4	20	1	5	26	2
JAC96	5	51	8	6	65	3
JAC72	6	42	7	4	53	0
JAC99	7	21	2	11	34	3
JAC31	8	22	2	12	36	2
JAC95 [f]	9	22	6	15	43	2
JAC74L	10	38	6	7	51	0
JAC86	11	20	1	12	33	5
JAC78	12	27	1	5	33	0
JAC87	13	60	3	12	75	4
JAC103	13	25	2	3	30	3
JAC6	14	21	5	13	39	4
JAC24	15	32	2	5	39	0
JAC64	16	20	0	12	32	3
JAC45 [f]	17	22	2	9	33	3
JAC47	18	29	2	7	38	2
JAC61A	19	50	6	12	68	5
JAC48	19	19	2	3	24	0
JAC60 [f]	19	20	1	5	26	0
JAC88	20	59	4	12	75	6
JAC58	21	22	2	5	29	2
JAC68	21	20	6	8	34	1
JAC81 [f]	22	33	4	22	59	14
JAC75	22	36	4	4	44	1
JAC76	22	14	1	5	20	0
JAC101	23	24	2	8	34	0
JAC122	24	25	4	13	42	5
JAC30L	25	43	11	10	64	4

JAC33	26	23	2	8	33	2
JAC110	27	38	6	7	51	0
JAC54	28	37	4	4	45	1
JAC55	28	52	6	4	62	1
JAC74C [f]	29	22	2	18	42	10
JAC35	30	32	5	8	45	6

[a] The OTU that the isolate was assigned.

[b] The total VOCs determined as coming from the empty jar control (the VOC collection system).

[c] The total VOCs determined as coming from the pooled media control.

[d] The total VOCs that were not present in the controls but detected from the *Streptomyces* isolate.

[e] VOCs detected from only the *Streptomyces* isolate in question and not the others.

[f] *Streptomyces* isolates selected for replicate culture VOC analysis.

[g] means not applicable.

3.3.2: Six isolates selected for replicate VOC analysis

Six of the 37 isolates (JAC25, JAC45, JAC60, JAC74C, JAC81, and JAC95) were selected for replicate VOC analysis (fermentation of isolates in triplicate followed by VOC collection and subsequent GC-MS analysis) based on the numbers and types of VOCs they produced during the pooled culture screens (Figure 3.3.1.1 and Table 3.3.1.3). Those VOCs identified with high confidence in the pooled culture screen were noted, and published bioactivities were reported (Table 3.3.2.1), which aided in selecting the six isolates for further analysis. Some of these VOCs were described in the literature as insect repellents, making them target candidates for future studies. Phenylethyl alcohol detected in JAC25 and JAC74C has mosquito repellency (Zeng et al., 2019) and antifungal activity (Li et al., 2020; Lyu et al., 2020). 1-hexanol produced by JAC60 suppresses mosquito CO_2 sensing receptors (Turner et al., 2011) and plays a role in oriental fruit moth (*Grapholita molesta*) signaling (Chen et al., 2019). 1-octen-3-ol detected in JAC74C has a variety of bioactivities, including nematicidal activity (Hummadi et al., 2020; Khoja et al., 2020), and it is produced by the foreign grain beetle (*Ahasverus advena*) (Pierce et al., 1990) and bont ticks (*Amblyomma variegatum* and *Amblyomma hebraeum*) (Mcmahon et al., 2001) and has a role in attracting the insects. Dodecane, heptylcyclohexane, and undecane produced by JAC95 are used in various insect communication systems (El-Sayed, 2007; Minaeimoghadam et al., 2017). Cubenol detected in JAC25 and JAC81 was found in plant essential oils that repel berry pests (Espinoza et al., 2016) and has larvicidal and adulticidal activity against various mosquito species (Dua et al., 2008). Geosmin detected in JAC25, JAC74C, JAC81, and JAC95 has various bioactivities, including springtail insect

Table 3.3.2.1: VOCs identified with high confidence in the pooled culture screen and their published bioactivities.

VOC	IUPAC name	Functional Group	Present in	Reported in plants	Bioactivities	Reported in *Streptomyces*
Alcohol						
Phenylethyl Alcohol	2-phenylethanol	aromatic alcohol	JAC17, JAC31, JAC86, JAC101, JAC33, JAC58, JAC74C, JAC25, JAC110, JAC74L, JAC30L	[b]	Mosquito repellent (Zeng et al., 2019); Antifungal (Li et al., 2020; Lyu et al., 2020)	(X. Li et al., 2020; Schöller et al., 2002)
DL-2,3-Butanediol	butane-2,3-diol	diol	JAC17, JAC47	[b]	[b]	[b]
1-Dodecanol	dodecan-1-ol	alcohol	JAC68	[b]	[b]	[b]
1-Hexanol	hexan-1-ol	alcohol	JAC60, JAC61A	Prickly pear (Andreu-Coll et al., 2020), mangos (Haocheng et al., 2020), korla fragrant pears (J. Chen et al., 2020)	Suppresses mosquito CO_2 receptors (Turner et al., 2011); beetle attractant (Zhu et al., 2020); role in the oriental fruit moth signaling & attractant (L. H. Chen et al., 2019)	(Cheng et al., 2020)
1-Octen-3-ol	oct-1-en-3-ol	fatty alcohol	JAC74C	orange essence (Högnadóttir & Rouseff, 2003) ; Kiwi fruit essence and Kiwi fruit puree, (El-Sayed, 2007); edible Korean chamchwi (Chung et al., 1993)	Nematicidal (Hummadi et al., 2020; Khoja et al., 2020); produced by and attract foreign grain beetle (Pierce et al., 1990); produced by and attract bont ticks (Mcmahon et al., 2001)	[b]
1-Octanol	octan-1-ol	fatty alcohol	JAC99, JAC101, JAC64	[b]	multiple insects use in their communication system – (El-Sayed, 2007)	[b]

Benzyl alcohol	phenylmethanol	aromatic alcohol	JAC17, JAC58, JAC25, JAC122, JAC30L	In climbing shrub essential oil (Thangadurai et al., 2002)	Head lice repellent (Toloza et al., 2006)	(Cheng et al., 2020; Narayana et al., 2008)
Ester						
Methyl 2-methylhexanoate	methyl 2-methylhexanoate	fatty ester	JAC45, JAC64	[b]	[b]	[b]
Dodecanoic acid, methyl ester	methyl dodecanoate	fatty ester	JAC35		Drosophila sex pheromone (Dweck et al., 2015)	(Cheng et al., 2020)
Methyl tetradecanoate	methyl tetradecanoate	fatty ester	JAC35, JAC110, JAC74L	emitted by multiple plant species – (El-Sayed, 2007)	Multiple insects use in their communication system (El-Sayed, 2007); ant repellent allomone (Posy et al., 1984)	[b]
Hydrocarbon						
Dodecane	dodecane	alkane	JAC95, JAC6, JAC122, JAC68	Released by tomato plants when exposed to pests (Errard et al., 2015)	clearwing moth sex pheromone (Minaeimoghadam et al., 2017)	(El-Din Mohamed & Zaki, 2019)
heptylcyclohexane	heptylcyclohexane	cyclic alkane	JAC95, JAC6, JAC122, JAC68	[b]	used in insect communication systems (El-Sayed, 2007)	[b]
Cetene	hexadec-1-ene	alkene	JAC47	[b]	[b]	[b]
Cyclopropane, pentyl-	pentylcyclopropane	cyclic alkane	JAC99, JAC101	in Persian Hogweed essential oil (Ehsani et al., 2019)	[b]	[b]
Undecane	undecane	alkane	JAC95, JAC6, JAC122	From herbaceous fern-root exudate (K. mei Zhang et al., 2019)	Ant sex and alarm pheromone (Monnin et al., 2018)	[b]
Ketone						

(R)-(+)-3-Methylcyclopentanone	(3R)-3-methylcyclopentan-1-one	cyclic ketone	JAC6	[b]	In the territorial marking fluid of Siberian tigers (Soso & Koziel, 2016)	[b]
3-Octanone	octan-3-one	ketone	JAC122	[b]	[b]	[b]
Terpene					[b]	
1,6-Cyclodecadiene, 1-methyl-5-methylene-8-(1-methylethyl)-, [S-(E,E)]- Germacrene D [a]	(1Z,6Z)-1-methyl-5-methylidene-8-propan-2-ylcyclodeca-1,6-diene	sesquiterpene	JAC25, JAC81	Synthase gene in grapevine (Lücker et al., 2004)	[b]	(Cheng et al., 2020; Pollak & Berger, 1996; Rabe et al., 2013);
2-Methyl-2-bornene	1,2,7,7-tetramethylbicyclo[2.2.1]hept-2-ene	monoterpene	JAC25	[b]	[b]	(Cheng et al., 2020; Rabe et al., 2013)
1H-Benzocycloheptene, 2,4a,5,6,7,8,9,9a-octahydro-3,5,5-trimethyl-9-methylene-, (4aS-cis)- α-Himachalene [a]	3,5,5-trimethyl-9-methylidene-2,4a,6,7,8,9a-hexahydro-1H-benzo[7]annulene	sesquiterpene	JAC102	In various plant extracts (Alves et al., 2019; Chaudhary et al., 2009; Joycharat et al., 2014; Ogunbinua et al., 2009; Paoli et al., 2011)	In essential oil that repels Mediterranean Fruit Fly (Alves et al., 2019)	(Cheng et al., 2020; Citron, Gleitzmann, et al., 2012)
Terpenoid						
1-Naphthalenol, 1,2,3,4,4a,7,8,8a-octahydro-1,6-dimethyl-4-(1-methylethyl)-, [1R-(1.alpha.,4.beta.,4a.beta.,8a.beta.)]- Cadinol [a]	(1R,4S)-1,6-dimethyl-4-propan-2-yl-3,4,4a,7,8,8a-hexahydro-2H-naphthalen-1-ol	sesquiterpene, alcohol	JAC17, JAC102, JAC24, JAC86	In giant Tangier fennel essential oil (Elghwaji et al., 2017)	[b]	[b]

cubenol	(1S,4R,4aR,8aR)-4,7-dimethyl-1-propan-2-yl-2,3,4,5,6,8a-hexahydro-1H-naphthalen-4a-ol	sesquiterpene, alcohol	JAC47, JAC25, JAC110, JAC74L, JAC81	floral compound emitted by various plants (El-Sayed, 2007); in various plant essential oil (Ciccarelli et al., 2016; Jirovetz et al., 2002; Mulyaningsih et al., 2010)	In essential oil that repels berry pests (Espinoza et al., 2016); in an essential oil that has larvicidal and adulticidal activity against various mosquito species (Dua et al., 2008)	(Cheng et al., 2020)
epi-cubenol	(1S,4R,4aS,8aR)-4,7-dimethyl-1-propan-2-yl-2,3,4,5,6,8a-hexahydro-1H-naphthalen-4a-ol	sesquiterpene, alcohol	JAC25, JAC95, JAC90, JAC87	[b]	[b]	(Rabe et al., 2013)
Geosmin	(4S,4aS,8aR)-4,8a-dimethyl-1,2,3,4,5,6,7,8-octahydronaphthalen-4a-ol	monoterpene, alcohol	JAC102, JAC31, JAC86, JAC101, JAC33, JAC76, JAC47, JAC64, JAC78, JAC74C, JAC25, JAC95, JAC35, JAC88, JAC55, JAC110, JAC74L, JAC54, JAC61A, JAC75, JAC90, JAC87, JAC81	[b]	Springtail insect attractant (Becher et al., 2020); Drosophila repellent (Stensmyr et al., 2012)	(Pollak & Berger, 1996; Rabe et al., 2013; Schöller et al., 2002)
2-Methylisoborneol	1,2,7,7-tetramethylbicyclo[2.2.1]heptan-2-ol	monoterpene, alcohol	JAC102, JAC99, JAC31, JAC101, JAC60, JAC78, JAC25, JAC72, JAC55, JAC54, JAC61A, JAC90	[b]	Springtail insect attractant (Becher et al., 2020); *Drosophila* attractant (Ho et al., 2020) ; Antifungal (Z. Wang et al., 2013)	(Becher et al., 2020; Rabe et al., 2013; Schrader et al., 2015)
2,6-Octadienoic acid, 3,7-dimethyl-, methyl ester	methyl (2E)-3,7-dimethylocta-2,6-dienoate	monoterpene, ester	JAC74C	[b]	[b]	[b]

Cyclohexanemethanol, 4-ethenyl-.alpha.,.alpha.,4-trimethyl-3-(1-methylethenyl)-, [1R-(1.alpha.,3.alpha.,4.beta.)]- elemol group [a]	2-(4-ethenyl-4-methyl-3-prop-1-en-2-ylcyclohexyl)propan-2-ol	sesquiterpene, alcohol	JAC74C	[b]	used in insect communication systems (El-Sayed, 2007)	(Cheng et al., 2020)
2-Naphthalenemethanol, decahydro-α,α,4a-trimethyl-8-methylene-, [2R-(2α,4aα,8aβ)]- β-Eudesmol ; β-Selinenol [a]	2-[(2R,4aR,8aS)-4a-methyl-8-methylidene-1,2,3,4,5,6,7,8a-octahydronaphthalen-2-yl]propan-2-ol	sesquiterpene, alcohol	JAC86, JAC101, JAC33, JAC76, JAC74C	In perennial medicinal plant extract (Tsusaka et al., 2020)	Insecticidal to *Drosophila* species (Chu et al., 2011)	(Cheng et al., 2020)
Other						
Pentanoic acid, 2,2,4-trimethyl-3-carboxyisopropyl, isobutyl ester	1-O-(2-methylpropyl) 4-O-propan-2-yl 2,2-dimethyl-3-propan-2-ylbutanedioate	ester	JAC122	[b]	[b]	[b]
Ethanol, 2-(2-ethoxyethoxy)- Carbitol [a]	2-(2-ethoxyethoxy)ethanol	alcohol, ether	JAC17	[b]	[b]	[b]
Pyrazine, 2,5-dimethyl-	2,5-dimethylpyrazine	pyrazine	JAC17, JAC24, JAC31, JAC33, JAC47, JAC64, JAC78, JAC74C, JAC25, JAC35, JAC96, JAC110, JAC74L, JAC61A, JAC30L, JAC87, JAC81	[b]	[b]	[b]
2,4-Furandicarboxylic acid, dimethyl ester	dimethyl furan-2,4-dicarboxylate	[b]	JAC74C	[b]	[b]	[b]
Hexanenitrile	hexanenitrile	cyanide	JAC60, JAC78, JAC61A	[b]	Used in aphid insect communication system (Pope et al., 2004)	[b]

[a] common name for the given VOC.

[b] means no relevant publications were found.

attractancy (Becher et al., 2020) and *Drosophila* repellency (Stensmyr et al., 2012). 2-Methylisoborneol detected in JAC25 and JAC60 attracts the springtail insect (Becher et al., 2020) and *Drosophila* (Ho et al., 2020) and has antifungal activity (Z. Wang et al., 2013). The terpene, β-Eudesmol detected in JAC74C, is insecticidal to *Drosophila* species (Chu et al., 2011). JAC45 did not produce any notable VOCs (Table 3.3.2.1). JAC45 and JAC60 produced the least number of VOCs, 9 and 5, respectively, while JAC81 produced the most VOCs (Table 3.3.1.3). The six isolates were confirmed for further analysis by Dr. Sean McCann (Hillier Lab, Acadia University). A total of seventy-one VOCs were detected in the six isolates during the pooled culture screens (Table 3.3.2.2), where there were more unknown compounds (low and unknown confidence level VOCs) than known compounds (high confidence VOCs) except for JAC74C (Figure 3.3.2.1). Further analysis is needed to identify the unknown VOCs, where potentially useful ones could be present.

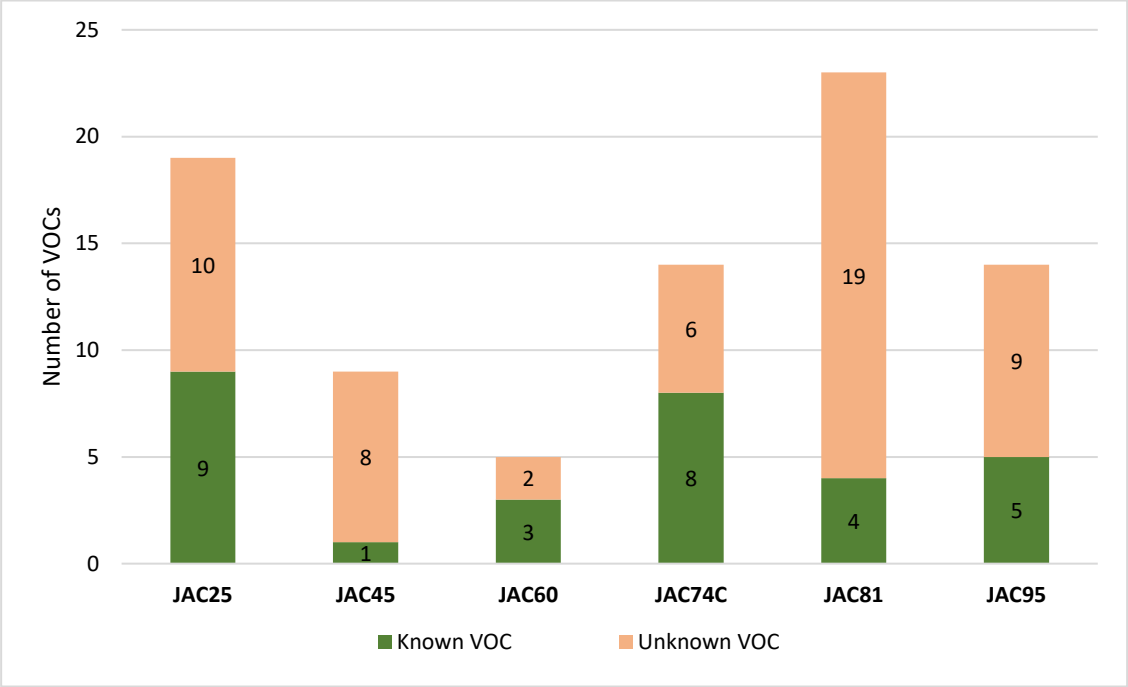


Figure 3.3.2.1: Number of VOCs detected in the pooled culture screens for the six *Streptomyces* isolates selected for replicate VOC analysis. Known VOCs were identified with a high confidence level, and unknown VOCs were identified with a low and unknown confidence level.

Table. 3.3.2.2: VOCs detected in the pooled culture screen for the six *Streptomyces* isolates selected for replicate VOC analysis.

No.	VOC [a]	IUPAC name	RT	RT Stdev [b]	Detected in [c]	Confidence in the VOC ID [d]	Detected in >1 OTU [e]
1	Hexanenitrile	hexanenitrile	6.403	0.025	JAC60,	high	yes
2	Pyrazine, 2,5-dimethyl-	2,5-dimethylpyrazine	7.230	0.012	JAC25, JAC74C, JAC81	high	yes
3	Methyl 2-methylhexanoate	methyl 2-methylhexanoate	8.537	0.001	JAC45	high	yes
4	Benzyl alcohol	phenylmethanol	10.290	0.027	JAC25	high	yes
5	Undecane	Undecane	12.099	0.001	JAC95	high	yes
6	Phenylethyl Alcohol	2-phenylethanol	12.432	0.032	JAC25, JAC74C	high	yes
7	2-Methylisoborneol	1,2,7,7-tetramethylbicyclo[2.2.1]heptan-2-ol	14.380	0.032	JAC25, JAC60	high	yes
8	Dodecane	dodecane	14.560	0.001	JAC95	high	yes
9	heptylcyclohexane	heptylcyclohexane	17.204	0.000	JAC95	high	yes
10	geosmin	(4S,4aS,8aR)-4,8a-dimethyl-1,2,3,4,5,6,7,8-octahydronaphthalen-4a-ol	18.356	0.014	JAC25, JAC74C, JAC81, JAC95	high	yes
11	1,6-Cyclodecadiene, 1-methyl-5-methylene-8-(1-methylethyl)-, [S-(E,E)]-	(1Z,6Z)-1-methyl-5-methylidene-8-propan-2-ylcyclodeca-1,6-diene	19.270	0.002	JAC25, JAC81	high	yes
12	epi-cubenol	(1S,4R,4aS,8aR)-4,7-dimethyl-1-propan-2-yl-2,3,4,5,6,8a-hexahydro-1H-naphthalen-4a-ol	21.128	0.001	JAC25, JAC95	high	yes
13	cubenol	(1S,4R,4aR,8aR)-4,7-dimethyl-1-propan-2-yl-2,3,4,5,6,8a-hexahydro-1H-naphthalen-4a-ol	21.277	0.024	JAC25, JAC81	high	yes

14	2-Naphthalenemethanol, decahydro-α,α,4a-trimethyl-8-methylene-, [2R-(2α,4aα,8aβ)]-	2-(4*a*-methyl-8-methylidene-1,2,3,4,5,6,7,8*a*-octahydronaphthalen-2-yl)propan-2-ol	21.320	0.001	JAC74C	high	yes
15	1-Hexanol	hexan-1-ol	6.223	0.034	JAC60	high	no
16	1-Octen-3-ol	oct-1-en-3-ol	8.963	0.000	JAC74C	high	no
17	2-Methyl-2-bornene	1,2,7,7-tetramethylbicyclo[2.2.1]hept-2-ene	9.964	0.000	JAC25	high	no
18	2,6-Octadienoic acid, 3,7-dimethyl-, methyl ester	methyl (2E)-3,7-dimethylocta-2,6-dienoate	16.820	0.000	JAC74C	high	no
19	2,4-Furandicarboxylic acid, dimethyl ester	dimethyl furan-2,4-dicarboxylate	17.364	0.000	JAC74C	high	no
20	Cyclohexanemethanol, 4-ethenyl-.alpha.,.alpha.,4-trimethyl-3-(1-methylethenyl)-, [1R-(1.alpha.,3.alpha.,4.beta.)]-	2-(4-ethenyl-4-methyl-3-prop-1-en-2-ylcyclohexyl)propan-2-ol	20.162	0.000	JAC74C	high	no
21	7-epi-.alpha.-selinene	5,8a-dimethyl-3-prop-1-en-2-yl-2,3,4,4a,7,8-hexahydro-1H-naphthalene	21.544	0.000	JAC95	low	yes
22	10-Methylundecan-4-olide	5-(6-methylheptyl)oxolan-2-one	21.220	0.000	JAC74C	low	no
23	10-Methyldodecan-5-olide	6-(5-methylheptyl)oxan-2-one	22.831	0.000	JAC45	low	no
24	Butane, 2-azido-2,3,3-trimethyl-	2-azido-2,3,3-trimethylbutane	4.022	0.000	JAC81	unknown	yes
25	Butanoic acid, 2-methyl-, methyl ester	methyl 2-methylbutanoate	4.359	0.005	JAC45	unknown	yes
26	Propanoic acid, 2,2-dimethyl-, 2,4-dinitrophenyl ester	(2,4-dinitrophenyl) 2,2-dimethylpropanoate	4.361	0.001	JAC60	unknown	yes
27	(S)-3-Ethyl-4-methylpentanol	3-ethyl-4-methylpentan-1-ol	9.462	0.053	JAC25	unknown	yes
28	1,3-Methanopentalene, octahydro-	tricyclo[5.1.1.02,6]nonane	14.053	0.001	JAC45	unknown	yes
29	Tricyclo[3.3.1.1(3,7)]decanone, 4-iodo-,	[g]	15.236	0.000	JAC25, JAC81	unknown	yes

	(1.alpha.,3.beta.,4.beta.,5.alpha.,7.beta.)-						
30	6-Tridecene, (Z)-	(Z)-tridec-6-ene	16.301	0.001	JAC95	unknown	yes
31	Bicyclo[3.2.0]hept-2-en-6-one, 7-chloro-	7-chlorobicyclo[3.2.0]hept-2-en-6-one	16.313	0.000	JAC45	unknown	yes
32	trans-1,10-Dimethyl-trans-9-decalinol	8,8a-dimethyl-2,3,4,4a,5,6,7,8-octahydro-1H-naphthalen-1-ol	18.385	0.001	JAC45, JAC60	unknown	yes
33	3-Heptyne, 5-methyl-	5-methylhept-3-yne	18.726	0.001	JAC25, JAC81	unknown	yes
34	Cyclopentane, nonyl-	nonylcyclopentane	18.755	0.002	JAC95	unknown	yes
35	n-Heptyl methylphosphonofluoridate	1-[fluoro(methyl)phosphoryl]oxyheptane	18.818	0.032	JAC95	unknown	yes
36	n-Heptadecylcyclohexane	heptadecylcyclohexane	20.143	0.000	JAC95	unknown	yes
37	Dihydro-cis-.alpha.-copaene-8-ol	2,8-dimethyl-5-propan-2-yltricyclo[4.4.0.02,7]decan-4-ol	20.549	0.031	JAC95	unknown	yes
38	β-Vatirenene	2-Isopropylidene-8,8a-dimethyl-1,2,3,7,8,8a-hexahydronaphthalene	21.177	0.002	JAC81	unknown	yes
39	1-Heptene, 3-methoxy-	3-methoxyhept-1-ene	4.268	0.000	JAC74C	unknown	no
40	2-Buten-1-ol, 2-methyl-	(E)-2-methylbut-2-en-1-ol	4.316	0.000	JAC74C	unknown	no
41	Propanoic acid, 2,2-dimethyl-, 2,4-dinitrophenyl ester	(2,4-dinitrophenyl) 2,2-dimethylpropanoate	5.488	0.000	JAC45	unknown	no
42	2-Butenoic acid, 2-methyl-, methyl ester	methyl (E)-2-methylbut-2-enoate	6.158	0.000	JAC74C	unknown	no
43	4-Octene, 2,3,7-trimethyl-, [S-(E)]-	(E,3S)-2,3,7-trimethyloct-4-ene	9.601	0.000	JAC25	unknown	no
44	Iron, tricarbonyl[(2,3,4,5-.eta.)-2,4-cycloheptadien-1-ol]-	[g]	13.856	0.000	JAC45	unknown	no
45	Iron, tricarbonyl[(2,3,4,5-.eta.)-2,4-cycloheptadien-1-ol]--Dup1 [f]	[g]	14.818	0.000	JAC45	unknown	no

46	6-Octen-1-ol, 3,7-dimethyl-, formate	3,7-dimethyloct-6-enyl formate	15.123	0.000	JAC74C	unknown	no
47	(1,2,2-trimethyl-3-cyclopenten-1-yl)acetaldehyde	2-(1,2,2-trimethylcyclopent-3-en-1-yl)acetaldehyde	15.315	0.000	JAC81	unknown	no
48	Phenol, 4-[2-(methylamino)ethyl]-	4-[2-(methylamino)ethyl]phenol	15.317	0.000	JAC25	unknown	no
49	2-Hydroxy-5-ethyl-5-methylcyclopent-2-en-1-one	5-ethyl-2-hydroxy-5-methylcyclopent-2-en-1-one	15.944	0.000	JAC81	unknown	no
50	Phenylacetaldehyde N-methyl-N-formylhydrazone	N-methyl-N-[(E)-2-phenylethylideneamino]formamide	16.707	0.000	JAC25	unknown	no
51	1-Methylene-2b-hydroxymethyl-3,3-dimethyl-4b-(3-methylbut-2-enyl)-cyclohexane	[2,2-dimethyl-4-(3-methylbut-2-enyl)-6-methylidenecyclohexyl]methanol	18.507	0.000	JAC81	unknown	no
52	1-Methylene-2b-hydroxymethyl-3,3-dimethyl-4b-(3-methylbut-2-enyl)-cyclohexane-Dup1 [f]	[g]	18.575	0.000	JAC81	unknown	no
53	(3R,2E)-2-(Hexadec-15-ynylidene)-3-hydroxy-4-methylenebutanolide	(3E,4R)-3-hexadec-15-ynylidene-4-hydroxy-5-methylideneoxolan-2-one	18.654	0.000	JAC81	unknown	no
54	1,4-Dimethyladamantane #	1,4-dimethyladamantane	18.726	0.000	JAC95	unknown	no
55	4,7-Methanoazulene, 1,2,3,4,5,6,7,8-octahydro-1,4,9,9-tetramethyl-, [1S-(1.alpha.,4.alpha.,7.alpha.)]-	1,5,11,11-tetramethyltricyclo[6.2.1.02,6]undec-2(6)-ene	18.812	0.000	JAC81	unknown	no
56	Dodecane, 2,6,10-trimethyl-	2,6,10-trimethyldodecane	18.876	0.000	JAC95	unknown	no
57	5,10-Pentadecadiyn-1-ol	pentadeca-5,10-diyn-1-ol	18.969	0.000	JAC81	unknown	no
58	2H-Benzocyclohepten-2-one, decahydro-4a-methyl-, trans-	4a-methyl-3,4,5,6,7,8,9,9a-octahydro-1H-benzo[7]annulen-2-one	19.147	0.000	JAC25	unknown	no

59	2,7-Octadiene-1,6-diol, 2,6-dimethyl-, (E)-	(2E)-2,6-dimethylocta-2,7-diene-1,6-diol	20.454	0.000	JAC81	unknown	no
60	Naphthalene, 1,2,4a,5,6,8a-hexahydro-4,7-dimethyl-1-(1-methylethyl)-	4,7-dimethyl-1-propan-2-yl-1,2,4a,5,6,8a-hexahydronaphthalene	20.502	0.000	JAC25	unknown	no
61	1-Hexyl-2-nitrocyclohexane	1-hexyl-2-nitrocyclohexane	20.947	0.000	JAC74C	unknown	no
62	Hexanethioic acid, S-methyl ester	S-methyl hexanethioate	21.058	0.000	JAC81	unknown	no
63	Dihydro-cis-.alpha.-copaene-8-ol-Dup1 [f]	[g]	21.267	0.000	JAC95	unknown	no
64	3,4-Nonadien-6-yne, 5-ethyl-3-methyl-	5-Ethyl-3-methyl-3,4-nonadien-6-yne	21.482	0.000	JAC81	unknown	no
65	1-Allyl-cyclohexane-1,2-diol	1-prop-2-enylcyclohexane-1,2-diol	21.544	0.000	JAC81	unknown	no
66	10-Heneicosene (c,t)	(E)-henicos-10-ene	21.758	0.000	JAC25	unknown	no
67	Cyclohexene, 1,5,5-trimethyl-6-(2-propenylidene)-	(6Z)-1,5,5-trimethyl-6-prop-2-enylidenecyclohexene	21.921	0.000	JAC81	unknown	no
68	Benzoic acid, 3-fluoro-, 2-oxo-2-phenylethyl ester	phenacyl 3-fluorobenzoate	21.976	0.000	JAC25	unknown	no
69	Cyclohexanol, 1-ethyl-2,2-dimethyl-6-methylene-	1-ethyl-2,2-dimethyl-6-methylidenecyclohexan-1-ol	22.219	0.000	JAC81	unknown	no
70	Hexanedioic acid, mono(2-ethylhexyl)ester	6-(2-ethylhexoxy)-6-oxohexanoic acid	28.637	0.000	JAC81	unknown	no
71	Phthalic acid, di(hept-3-yl) ester	diheptan-3-yl benzene-1,2-dicarboxylate	30.506	0.000	JAC81	unknown	No

[a] VOC names were assigned by the NIST library. This was used in conjunction with the retention time for more accurate identification of compounds, i.e., VOCs with the same name but different RT should be interpreted cautiously.

[b] Standard deviation from the average retention time (RT) of a VOC.

[c] Only the six isolates selected for replicate VOC analysis are included here. A complete list of isolates a VOC was detected in can be found in Table 3.3.1.1 if the VOC is listed as detected in more than one OTU.

[d] The level of confidence assigned to the identity of the VOC. High confidence means there was a good spectral match to the NIST library reference and a Kovats retention index (RI) within 10 points of a literature value available on the NIST WebBook (https://webbook.nist.gov/chemistry/name-ser/), low confidence means the difference in RI of a NIST library suggested VOC to that of the literature value is greater 10 points, and unknown means that a literature RI value could not be found for a NIST library suggested compound.

[e] Indicates if the VOC was from one or more isolates/OTU.

[f] Dup1 suffix means a NIST reference spectra with close identity to the given VOC was not available in the library used in this study.

[g] not applicable, or the IUPAC name is not available.

3.3.3: VOC production profiles for replicate cultures of the six selected *Streptomyces* isolates

Six *Streptomyces* isolates described above were chosen from the 37 phylogenetically different ones used in pooled culture VOC analysis. These six were fermented in triplicate in the three media for replicate VOC analysis. Ninety-five VOCs were identified (Table 3.3.3.1), where VOCs present in at least two of three replicate cultures were considered produced by a *Streptomyces* isolate (Figure 3.3.3.1). JAC81 fermented in SFM produced VOCs that elute at RTs 21.056, 21.177, and 21.266 with relatively prominent peaks (Figure 3.3.3.1). However, the confidence in their identity is not high. That is to say that the published literature or standards were not available to confirm that these VOCs elute at the RTs mentioned above. VOCs were considered unknown because the reference spectra were not available to identify them unambiguously. In these instances, more than one VOCs eluted at the same RTs (Table 3.3.3.1) and could not be distinguished with the spectra alone. Therefore, they should be interpreted cautiously.

Replicate VOC analysis revealed that some compounds detected in the pooled VOC screen were only produced in certain fermentation media (Table 3.3.3.1 and Figure 3.3.3.2). Overall, the isolates produced the highest number of VOCs in SFM medium and the least number in the synthetic medium, except for JAC25 and JAC74C, where the numbers were the same or more in the YMS medium, respectively (Figure 3.3.3.2). JAC81 and JAC95 were relatively the most productive in the three media regarding the numbers of VOCs (Figure 3.3.3.2).

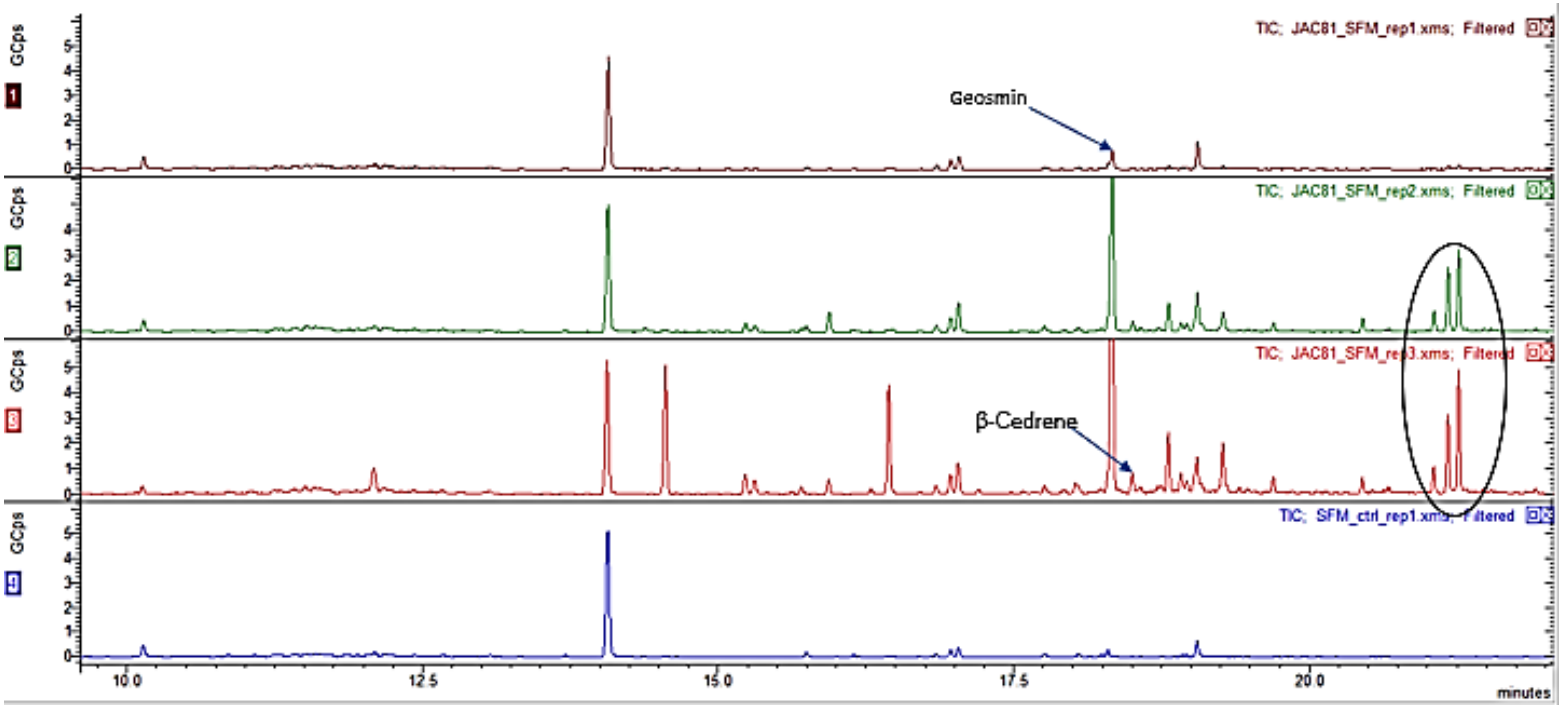


Figure 3.3.3.1: Example (~10- 22 min) of GC-MS chromatogram for JAC81 replicate cultures (x3) fermented in SFM medium. Two peaks with >40% probability match to the reference spectra are labeled. JAC81 chromatograms are compared to the SFM medium chromatogram used as a control. Only peaks common to at least two of JAC81 replicate cultures are considered VOC products of said bacteria. Peaks at RT 21.056 to 21.266 (circled) were identified with an "unknown" confidence level.

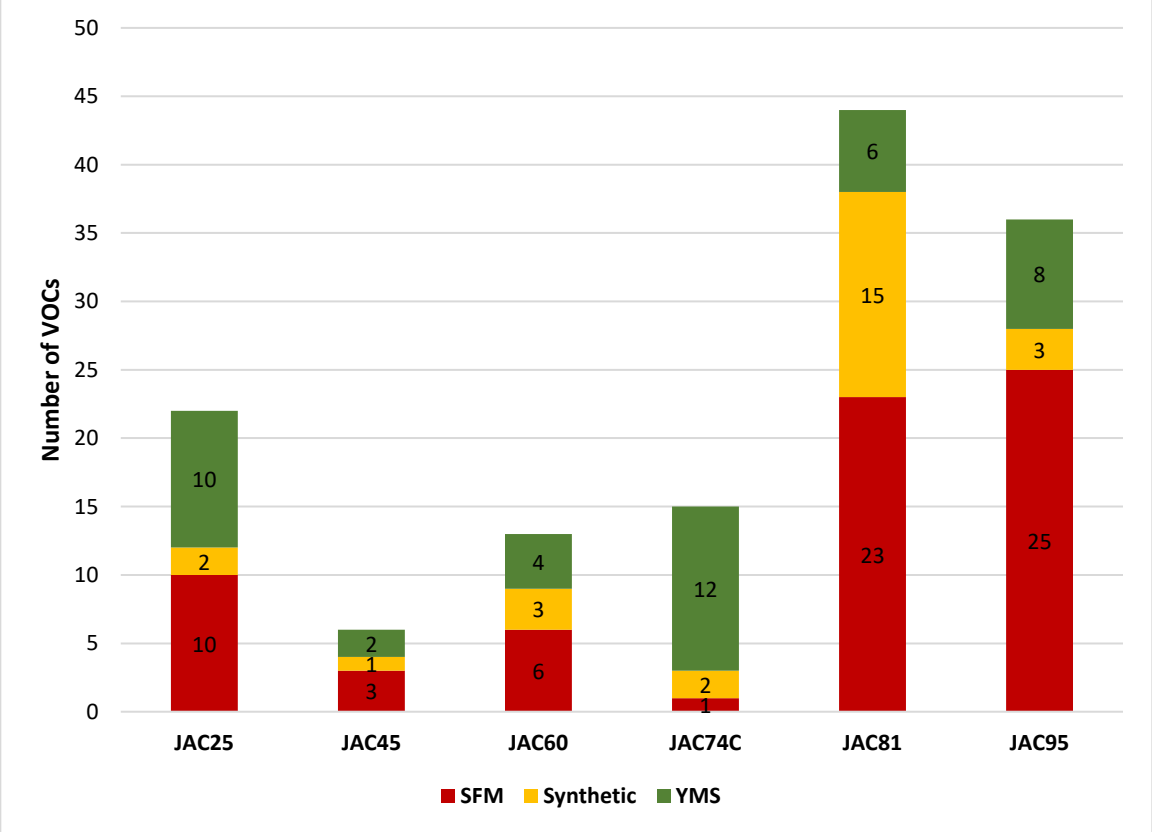


Figure 3.3.3.2: The numbers of VOCs produced by the *Streptomyces* isolates in replicate culture analysis. VOCs were counted if they were present in at least two replicate cultures of a given isolate.

Table 3.3.3.1: Bacterially derived compounds identified in replicate VOC analysis.

No.	VOC [a]	IUPAC name	Confidence in VOC ID [b]	RT	Stdev from the RT [c]	Medium	Isolate	Present in total replicates [d]
1	(1,2,2-trimethyl-3-cyclopenten-1-yl)acetaldehyde	2-(1,2,2-trimethylcyclopent-3-en-1-yl)acetaldehyde	low	15.711	0.001	SFM	JAC25	2
				15.714	0.002	SFM	JAC95	3
				15.707	0.004	SFM	JAC81	2
2	(S)-3,4-Dimethylpentanol	(3S)-3,4-dimethylpentan-1-ol	unknown	6.257	0.045	YMS	JAC95	2
3	.alfa.-Copaene	1,3-dimethyl-8-propan-2-yltricyclo[4.4.0.02,7]dec-3-ene	unknown	21.130	0.001	SFM	JAC95	3
4	1,5-Cyclodecadiene, 1,5-dimethyl-8-(1-methylethenyl)-, [S-(Z,E)]-	(1E,5Z,8S)-1,5-dimethyl-8-prop-1-en-2-ylcyclodeca-1,5-diene	unknown	17.933	0.001	SFM	JAC95	3
				17.930	0.002	YMS	JAC74C	2
5 [a]	1,5-Cyclodecadiene, 1,5-dimethyl-8-(1-methylethenyl)-, [S-(Z,E)]--Dup1 [f]	[g]	unknown	20.109	0.001	YMS	JAC74C	3
	1-Methylene-2b-hydroxymethyl-3,3-dimethyl-4b-(3-methylbut-2-enyl)-cyclohexane	[2,2-dimethyl-4-(3-methylbut-2-enyl)-6-methylidenecyclohexyl]methanol						
6 [a]	1,6:3,4-Dianhydro-2-deoxy-.beta.-d-ribo-hexopyranose	3,7,9-trioxatricyclo[4.2.1.02,4]nonane	unknown	5.499	0.001	YMS	JAC25	2
	Oxalic acid, cyclohexyl tetradecyl ester	2-O-cyclohexyl 1-O-tetradecyl oxalate						

7	1,6-Cyclodecadiene, 1-methyl-5-methylene-8-(1-methylethyl)-, [S-(E,E)]-	(1Z,6Z)-1-methyl-5-methylidene-8-propan-2-ylcyclodeca-1,6-diene	high	19.273	0.000	SFM	JAC25	2
	Germacrene D [e]							
				19.274	0.001	SFM	JAC95	3
				19.271	0.001	SFM	JAC81	3
				19.271	0.002	synthetic	JAC81	3
				19.274	0.002	YMS	JAC74C	2
8 [a]	1,8-Nonanediol, 8-methyl-	8-methylnonane-1,8-diol	unknown	4.722	0.014	YMS	JAC25	2
	2-Methyl-2,3-pentanediol	2-methylpentane-2,3-diol		4.720	0.004	YMS	JAC60	2
9	10-Methylundecan-4-olide	5-(6-methylheptyl)oxolan-2-one	low	21.155	0.004	synthetic	JAC74C	2
				21.157	0.002	YMS	JAC74C	3
10	1H-3a,7-Methanoazulene, octahydro-3,8,8-trimethyl-6-methylene-, [3R-(3.alpha.,3a.beta.,7.beta.,8a.alpha.)]-	2,6,6-trimethyl-8-methylidenetricyclo[5.3.1.01,5]undecane	high	18.506	0.003	SFM	JAC81	3
	β-Cedrene [e]							
11	1-Hexanol	hexan-1-ol	high	6.197	0.002	synthetic	JAC60	3
12 [a]	1-Hexyl-2-nitrocyclohexane	1-hexyl-2-nitrocyclohexane	unknown	19.803	0.001	SFM	JAC95	2
	Hexadecen-1-ol, trans-9-	(E)-hexadec-9-en-1-ol						
13 [a]	1-Hexyl-2-nitrocyclohexane-Dup1 [f]	[g]	unknown	27.638	0.001	YMS	JAC25	3
	Oxiraneoctanoic acid, 3-octyl-, methyl ester	methyl 8-[(2S,3R)-3-octyloxiran-2-yl]octanoate						
	3-Methyl-2-(2-oxopropyl)furan	1-(3-methylfuran-2-yl)propan-2-one						

14	1H-Indene, 1-ethylideneoctahydro-7a-methyl-, (1E,3aα,7aβ)-	(1E)-1-Ethylidene-7a-methyloctahydro-1H-indene	unknown	15.626	0.005	SFM	JAC81	2
15	1H-Indene, 1-ethylideneoctahydro-7a-methyl-, (1Z,3aα,7aβ)-	(1E)-1-Ethylidene-7a-methyloctahydro-1H-indene	unknown	15.316	0.003	SFM	JAC25	2
				15.317	0.000	SFM	JAC95	3
				15.311	0.002	SFM	JAC81	2
				15.314	0.003	synthetic	JAC81	2
16	1H-Indene, 1-ethylideneoctahydro-7a-methyl-, cis-	[g]	unknown	15.234	0.003	SFM	JAC60	3
				15.235	0.001	SFM	JAC25	2
				15.237	0.001	SFM	JAC95	3
				15.234	0.005	SFM	JAC81	3
				15.233	0.003	synthetic	JAC81	3
				15.237	0.001	YMS	JAC74C	2
17 [a]	1-Methyl-4-isopropyl-cyclohexyl 2-hydroperfluorobutanoate-Dup1 [f]	[g]	unknown	21.836	0.001	YMS	JAC45	2
	4-[(Bicyclo[2.2.1]hepta-2,5-dien-7-yl)oxy]benzonitrile-Dup1 [f]	[g]						
18	1-Methylene-2b-hydroxymethyl-3,3-dimethyl-4b-(3-methylbut-2-enyl)-cyclohexane	[2,2-dimethyl-4-(3-methylbut-2-enyl)-6-methylidenecyclohexyl]methanol	unknown	18.421	0.001	SFM	JAC95	2
				18.505	0.001	synthetic	JAC81	3
19	1-Methylene-2b-hydroxymethyl-3,3-dimethyl-	[g]	unknown	18.507	0.002	YMS	JAC81	2

	4b-(3-methylbut-2-enyl)-cyclohexane-Dup1 [f]							
20 [a]	1-Methylene-2b-hydroxymethyl-3,3-dimethyl-4b-(3-methylbut-2-enyl)-cyclohexane-Dup1 [f]	[g]	unknown	21.479	0.001	SFM	JAC81	2
	10-Pentadecen-5-yn-1-ol, (E)-	(E)-pentadec-10-en-5-yn-1-ol						
21 [a]	1-Methylene-2b-hydroxymethyl-3,3-dimethyl-4b-(3-methylbut-2-enyl)-cyclohexane-Dup1 [f]	[g]	unknown	21.479	0.002	synthetic	JAC81	3
	5,10-Pentadecadiyn-1-ol ; 1,5-Diphenylhex-3-ene-Dup1 [f]	[g]						
22 [a]	1-Methylene-2b-hydroxymethyl-3,3-dimethyl-4b-(3-methylbut-2-enyl)-cyclohexane-Dup1 [f]	[g]	unknown	21.266	0.001	YMS	JAC74C	3
	Dihydro-cis-.alpha.-copaene-8-ol	2,8-dimethyl-5-propan-2-yltricyclo[4.4.0.02,7]decan-4-ol		21.263	0.002	synthetic	JAC81	3
				21.266	0.001	SFM	JAC25	2
				21.266	0.002	SFM	JAC95	3
	5,10-Pentadecadiyne, 1-chloro-	1-chloropentadeca-5,10-diyne						
23	1-Phenyl-2-methyl-oct-1-ene	[(E)-2-methyloct-1-enyl]benzene	unknown	21.177	0.002	SFM	JAC81	3
				21.177	0.002	synthetic	JAC81	3
				21.177	0.001	YMS	JAC81	2
24 [a]	2(5H)-Furanone, 5-(bromomethyl)-5-phenyl-	5-(bromomethyl)-5-phenylfuran-2-one	unknown	21.060	0.003	synthetic	JAC81	2

	3-[4-Hydroxyisopent-2(Z)-enyl]-4-hydroxyacetophenone	1-[4-hydroxy-3-[(Z)-4-hydroxy-3-methylbut-2-enyl]phenyl]ethanone						
25 [a]	2,3,4-Trifluorobenzoic acid, 2,6-dimethylnon-1-en-3-yn-5-yl ester	2,6-dimethylnon-1-en-3-yn-5-yl 2,3,4-trifluorobenzoate	unknown	18.967	0.002	synthetic	JAC81	3
	1-Phenyl-2-methyl-oct-1-ene	[(E)-2-methyloct-1-enyl]benzene						
26 [a]	2,5-Octadiyne, 4,4-diethyl-	4,4-diethylocta-2,5-diyne	unknown	21.914	0.001	SFM	JAC81	2
	Cyclohexene, 1,5,5-trimethyl-6-(2-propenylidene)-	(6Z)-1,5,5-trimethyl-6-prop-2-enylidenecyclohexene		21.916	0.001	synthetic	JAC81	3
27	2,6-Octadienoic acid, 3,7-dimethyl-, methyl ester	methyl (2E)-3,7-dimethylocta-2,6-dienoate	high	16.772	0.001	YMS	JAC74C	2
	Methyl geranate [e]							
28 [a]	2-Furanol, tetrahydro-2-methyl-	2-methyloxolan-2-ol	unknown	4.020	0.003	synthetic	JAC81	2
	propanoic acid, 2-methyl- ethyl ester	ethyl 2-methylpropanoate						
	Propanedioic acid, oxo-, bis(1-methylethyl) ester	dipropan-2-yl 2-oxopropanedioate		4.027	0.005	SFM	JAC81	2
	Butane, 2,2'-[methylenebis(oxy)]bis[2-methyl-	2-methyl-2-(2-methylbutan-2-yloxymethoxy)butane		4.022	0.002	SFM	JAC95	3
29	2H-3,9a-Methano-1-benzoxepin, octahydro-2,2,5a,9-tetramethyl-, [3R-(3.alpha.,5a.alpha.,9.alpha.,9a.alpha.)]-	(2S,6R,9S)-2,6,10,10-tetramethyl-11-oxatricyclo[7.2.1.01,6]dodecane	low	19.699	0.001	SFM	JAC81	2
	Dihydroagarofuran [e]							

30	2H-Pyran-2-one, 6-hexyltetrahydro-	6-hexyloxan-2-one	unknown	18.799	0.005	SFM	JAC95	3
31 [a]	2-Hydroxy-3-pentanone	2-hydroxypentan-3-one	low	4.854	0.001	YMS	JAC95	2
	DL-2,3-Butanediol	butane-2,3-diol		4.856	0.006	YMS	JAC60	2
32 [a]	2-Hydroxy-5-ethyl-5-methylcyclopent-2-en-1-one	5-ethyl-2-hydroxy-5-methylcyclopent-2-en-1-one	unknown	15.940	0.006	SFM	JAC81	3
	Bicyclo[2.2.1]heptan-2-ol, 2-(2-cyclopenten-1-yl)-	2-cyclopent-2-en-1-ylbicyclo[2.2.1]heptan-2-ol						
33	2-Methyl-2,3-pentanediol	2-methylpentane-2,3-diol	unknown	4.720	0.004	YMS	JAC60	2
34	2-Methyl-2-bornene	1,2,7,7-tetramethylbicyclo[2.2.1]hept-2-ene	high	9.966	0.002	SFM	JAC60	3
				9.965	0.001	SFM	JAC25	2
35	2-Methyl-3-(3-methyl-but-2-enyl)-2-(4-methyl-pent-3-enyl)-oxetane	2-methyl-3-(3-methylbut-2-enyl)-2-(4-methylpent-3-enyl)oxetane	unknown	18.724	0.001	SFM	JAC81	2
36	2-Methylisoborneol	1,2,7,7-tetramethylbicyclo[2.2.1]heptan-2-ol	high	14.348	0.001	SFM	JAC60	3
				14.349	0.001	SFM	JAC25	2
				14.347	0.004	synthetic	JAC25	2
				14.351	0.002	YMS	JAC25	2
37 [a]	2R-Acetoxymethyl-1,3,3-trimethyl-4t-(3-methyl-2-buten-1-yl)-1t-cyclohexanol	[6-hydroxy-2,2,6-trimethyl-3-(3-methylbut-2-enyl)cyclohexyl]methyl acetate	unknown	18.918	0.002	SFM	JAC81	2
	Dihydro-cis-.alpha.-copaene-8-ol-Dup1 [f]	[g]						
38	3-Heptyne, 5-methyl-	5-methylhept-3-yne	low	18.729	0.001	SFM	JAC95	3
39	3-Isopropoxy-1,1,1,7,7,7-hexamethyl-3,5,5-tris(trimethylsiloxy)tetrasiloxane	propan-2-yl bis(trimethylsilyl) tris(trimethylsilyloxy)silyl silicate	unknown	18.965	0.002	SFM	JAC81	2

40 [a]	4,4-Dimethyl-3-(3-methylbut-3-enylidene)-2-methylenebicyclo[4.1.0]heptane	(3E)-4,4-dimethyl-3-(3-methylbut-3-enylidene)-2-methylidenebicyclo[4.1.0]heptane	unknown	18.578	0.001	synthetic	JAC81	3
	Benzene, 1-(5,5-dimethyl-1-cyclopenten-1-yl)-2-methoxy-	1-(5,5-dimethylcyclopenten-1-yl)-2-methoxybenzene		18.576	0.002	SFM	JAC81	2
	1-Phenyl-2-methyl-oct-1-ene	[(E)-2-methyloct-1-enyl]benzene						
41	4-Acetylbenzoic acid	4-Acetylbenzoic acid	unknown	14.294	0.002	YMS	JAC95	3
42	4-Methyl-2,4-bis(4'-trimethylsilyloxyphenyl)pentene-1-Dup1 [f]	[g]	unknown	4.862	0.003	synthetic	JAC60	3
43	5,6-Decadien-3-yne, 5,7-diethyl-	[g]	unknown	21.126	0.001	synthetic	JAC25	3
44 [a]	5,6-Decadien-3-yne, 5,7-diethyl-	[g]	unknown	18.811	0.002	synthetic	JAC81	3
	Cadina-1(10),6,8-triene	2,5-dimethyl-8-propan-2-yl-1,2,3,4-tetrahydronaphthalene		18.812	0.002	YMS	JAC81	2
45	5-Hepten-3-one, 5-ethyl-2-methyl-	(E)-5-ethyl-2-methylhept-5-en-3-one	unknown	20.451	0.004	SFM	JAC81	2
46	6,7-Dodecanedione	dodecane-6,7-dione	unknown	21.056	0.003	SFM	JAC81	3
47	7-epi-.alpha.-eudesmol	2-[(2S,4aR)-4a,8-dimethyl-2,3,4,5,6,8a-hexahydro-1H-naphthalen-2-yl]propan-2-ol	low	21.547	0.001	SFM	JAC95	3
48 [a]	Acetamide, N,N'-ethylenebis(N-nitro-	N-[2-[acetyl(nitro)amino]ethyl]-N-nitroacetamide	unknown	4.579	0.021	YMS	JAC81	2
	Pentane, 1-bromo-3,4-dimethyl-	1-bromo-3,4-dimethylpentane						
49	Acetic acid, (dodecahydro-7-hydroxy-1,4b,8,8-tetramethyl-10-oxo-2(1H)-	2-(dimethylamino)ethyl (2E)-2-(7-hydroxy-1,4b,8,8-tetramethyl-10-oxo-1,3,4,4a,5,6,7,8a,9,10a-	unknown	19.482	0.002	SFM	JAC81	2

	phenanthrenylidene)-,2-(dimethylamino)ethyl ester	decahydrophenanthren-2-ylidene)acetate						
50 [a]	Benzene, (propoxymethyl)-	propoxymethylbenzene	unknown	4.231	0.002	SFM	JAC45	3
	1-Heptene, 3-methoxy-	3-methoxyhept-1-ene						
51	Benzoic acid, methyl ester	methyl benzoate	high	11.889	0.001	SFM	JAC95	2
52	Benzyl alcohol	phenylmethanol	high	10.269	0.001	SFM	JAC95	2
53 [a]	Bicyclo[3.2.0]hepta-2,6-diene	bicyclo[3.2.0]hepta-2,6-diene	unknown	12.398	0.010	YMS	JAC95	2
	3-Methoxyamphetamine	1-(3-methoxyphenyl)propan-2-amine						
54	Butane, 2-azido-2,3,3-trimethyl-	2-azido-2,3,3-trimethylbutane	unknown	5.683	0.017	synthetic	JAC95	3
55 [a]	Butanedioic acid, phenyl-	2-phenylbutanedioic acid	unknown	6.730	0.001	synthetic	JAC60	3
	Oxalic acid, pentyl propyl ester	2-O-pentyl 1-O-propyl oxalate						
	(SS)- or (RR)-4-methyl-2,3-pentanediol	4-methylpentane-2,3-diol						
56 [a]	Butanoic acid, 2-methyl-, methyl ester	methyl 2-methylbutanoate	unknown	4.314	0.000	SFM	JAC45	3
				4.314	0.001	synthetic	JAC45	2
	Propanoic acid, 2,2-dimethyl-, 2,4-dinitrophenyl ester	(2,4-dinitrophenyl) 2,2-dimethylpropanoate		4.310	0.004	SFM	JAC60	3
57	Butanoic acid, 3-methyl-	3-methylbutanoic acid	high	5.582	0.007	synthetic	JAC95	3
58	Butanoic acid, 3-methyl-	3-methylbutanoic acid	low	5.634	0.002	synthetic	JAC95	2
59 [a]	Butanoic acid, 3-methyl-	3-methylbutanoic acid	unknown	6.112	0.002	YMS	JAC74C	2
	Methyl tiglate	methyl (E)-2-methylbut-2-enoate						
60	Cadina-1(10),6,8-triene	2,5-dimethyl-8-propan-2-yl-1,2,3,4-tetrahydronaphthalene	low	18.812	0.003	SFM	JAC81	3
61	Caryophyllenyl alcohol	(3Z)-4,8,11,11-tetramethylbicyclo[7.2.0]undec-3-en-5-ol	unknown	20.523	0.001	SFM	JAC95	3

62 [a]	cis-muurola-3,5-diene	(1R,4R,4aS)-4,7-dimethyl-1-propan-2-yl-1,2,3,4,4a,5-hexahydronaphthalene	unknown	21.129	0.001	SFM	JAC25	2
	5,6-Decadien-3-yne, 5,7-diethyl-	[g]						
63 [a]	Cyclobutaneacetonitrile, 1-methyl-2-(1-methylethylidene)-	[g]	unknown	19.031	0.001	YMS	JAC74C	2
	2H-3,9a-Methano-1-benzoxepin, octahydro-2,2,5a,9-tetramethyl-, [3R-(3.alpha.,5a.alpha.,9.alpha.,9a.alpha.)]-	(2S,6R,9S)-2,6,10,10-tetramethyl-11-oxatricyclo[7.2.1.01,6]dodecane						
64 [a]	Cyclobutaneacetonitrile, 1-methyl-2-(1-methylethylidene)--Dup1 [f]	[g]	unknown	19.855	0.002	SFM	JAC95	3
	Nonadecane	nonadecane						
65 [a]	Cyclobutanone, 2-methyl-	2-methylcyclobutan-1-one	unknown	4.562	0.002	SFM	JAC45	3
	Cyclopentanone, 2-(1-methylpropyl)-	2-butan-2-ylcyclopentan-1-one						
66 [a]	Cyclohexanol, 1,3,3-trimethyl-2-(3-methyl-2-methylene-3-butenylidene)-, (Z)-	(2*Z*)-1,3,3-trimethyl-2-(3-methyl-2-methylidenebut-3-enylidene)cyclohexan-1-ol	unknown	21.836	0.004	YMS	JAC60	2
	Cyclohexanone, 3-ethenyl-3-methyl-2-(1-methylethenyl)-6-(1-methylethylidene)-, cis-	[g]						
67	Cyclohexanol, 1-ethyl-2,2-dimethyl-6-methylene-	1-ethyl-2,2-dimethyl-6-methylidenecyclohexan-1-ol	unknown	22.103	0.002	SFM	JAC81	2
68 [a]	Cyclohexanol, 4-ethyl-4-methyl-3-(1-methylethyl)-, (1.alpha.,3.beta.,4.alpha.)-	4-ethyl-4-methyl-3-propan-2-ylcyclohexan-1-ol	unknown	19.704	0.001	SFM	JAC95	3

	2H-3,9a-Methano-1-benzoxepin, octahydro-2,2,5a,9-tetramethyl-, [3R-(3.alpha.,5a.alpha.,9.alpha.,9a.alpha.)]--Dup1 [f]	[g]						
	11,12-Dihydroxyseychellane	2-(hydroxymethyl)-3,6,8-trimethyltricyclo[5.3.1.03,8]undecan-2-ol						
69 [a]	Decane, 3-bromo-	3-bromodecane	unknown	16.709	0.002	SFM	JAC74C	2
	Nonadecane	nonadecane						
70 [a]	Diacetyl sulphide	S-acetyl ethanethioate	low	4.313	0.003	YMS	JAC45	2
	(S)-3,4-Dimethylpentanol	(3S)-3,4-dimethylpentan-1-ol						
71	Diethyl Phthalate	diethyl benzene-1,2-dicarboxylate	high	20.523	0.001	YMS	JAC25	3
				20.522	0.002	YMS	JAC81	3
				20.523	0.001	YMS	JAC74C	2
72	Dihydro-cis-.alpha.-copaene-8-ol-Dup1 [f]	[g]	unknown	18.387	0.002	SFM	JAC81	3
73	Dihydro-cis-.alpha.-copaene-8-ol	2,8-dimethyl-5-propan-2-yltricyclo[4.4.0.02,7]decan-4-ol	unknown	21.263	0.002	synthetic	JAC81	3
				21.266	0.001	SFM	JAC25	2
				21.266	0.002	SFM	JAC95	3
74 [a]	Dihydro-cis-.alpha.-copaene-8-ol-Dup1 [f]	[g]	unknown	19.482	0.001	SFM	JAC95	3
	Selina-3,7(11)-diene	5,8a-dimethyl-3-propan-2-ylidene-1,2,4,4a,7,8-hexahydronaphthalene						
75 [a]	Ethaneperoxoic acid, 1-cyano-1-[2-(2-phenyl-1,3-dioxolan-2-yl)ethyl]pentyl ester	[3-cyano-1-(2-phenyl-1,3-dioxolan-2-yl)heptan-3-yl] ethaneperoxoate	unknown	20.523	0.001	YMS	JAC74C	2

	Diethyl Phthalate	diethyl benzene-1,2-dicarboxylate						
76	Geosmin	(4S,4aS,8aR)-4,8a-dimethyl-1,2,3,4,5,6,7,8-octahydronaphthalen-4a-ol	high	18.331	0.001	SFM	JAC60	3
				18.332	0.002	SFM	JAC25	2
				18.337	0.005	SFM	JAC95	3
				18.334	0.001	SFM	JAC81	3
				18.331	0.002	synthetic	JAC81	3
				18.333	0.002	YMS	JAC74C	3
				18.330	0.003	YMS	JAC81	2
77 [a]	Heptacosane	heptacosane	unknown	25.901	0.000	YMS	JAC25	3
	Dodecane, 2,6,10-trimethyl-	2,6,10-trimethyldodecane						
78 [a]	Hexacosane	hexacosane	unknown	19.854	0.001	YMS	JAC95	2
	Decane, 2,4,6-trimethyl-	2,4,6-trimethyldecane						
79 [a]	Hexadecanoic acid, 15-methyl-, methyl ester	methyl 15-methylhexadecanoate	unknown	26.128	0.001	YMS	JAC25	3
	Methyl stearate	methyl octadecanoate						
80 [a]	Hexane, 2,3-dimethyl-	2,3-dimethylhexane	unknown	20.282	0.001	SFM	JAC95	3
	n-Hexyl methylphosphonofluoridate	1-[fluoro(methyl)phosphoryl]oxyhexane						
81	Hexane, 3,4-bis(1,1-dimethylethyl)-2,2,5,5-tetramethyl-	3,4-ditert-butyl-2,2,5,5-tetramethylhexane	unknown	4.864	0.002	SFM	JAC95	2
82	Hexanedioic acid, bis(2-ethylhexyl) ester	bis(2-ethylhexyl) hexanedioate	unknown	28.636	0.001	YMS	JAC25	3
83	Hexanoic acid, 5-methyl-, methyl ester	methyl 5-methylhexanoate	high	9.322	0.001	SFM	JAC60	2

84 [a]	Naphthalene, 1,2,3,4,4a,5,6,8a-octahydro-4a,8-dimethyl-2-(1-methylethenyl)-, [2R-(2.alpha.,4a.alpha.,8a.beta.)]-	(3S,8aS)-5,8a-dimethyl-3-prop-1-en-2-yl-2,3,4,4a,7,8-hexahydro-1H-naphthalene	unknown	21.268	0.001	SFM	JAC95	3
	Dihydro-cis-.alpha.-copaene-8-ol-Dup1 [f]	[g]						
85 [a]	Naphthalene, 1,2,3,4,4a,7-hexahydro-1,6-dimethyl-4-(1-methylethyl)-	1,6-dimethyl-4-propan-2-yl-1,2,3,4,4a,7-hexahydronaphthalene	unknown	19.933	0.001	SFM	JAC95	2
	5,6-Decadien-3-yne, 5,7-diethyl-	[g]						
86 [a]	Naphthalene, 1,2,3,4,4a,7-hexahydro-1,6-dimethyl-4-(1-methylethyl)-	1,6-dimethyl-4-propan-2-yl-1,2,3,4,4a,7-hexahydronaphthalene	unknown	21.128	0.001	YMS	JAC25	3
	5,6-Decadien-3-yne, 5,7-diethyl-	[g]						
87 [a]	Phenol, 4-[2-(methylamino)ethyl]-	4-[2-(methylamino)ethyl]phenol	unknown	15.712	0.000	YMS	JAC74C	2
	indene-4(1H-indene, 1-ethylideneoctahydro-7a-methyl-,cis-	[g]						
88 [a]	Phenylacetaldehyde N-methyl-N-formylhydrazone	N-methyl-N-[(E)-2-phenylethylideneamino]formamide	unknown	21.391	0.001	YMS	JAC95	2
	Ethanone, 1-(3-ethyloxiranyl)-	1-(3-ethyloxiran-2-yl)ethanone						
89	Phenylethyl Alcohol	2-phenylethanol	high	12.393	0.002	SFM	JAC95	3
				12.388	0.002	YMS	JAC60	2
90	photocitral B	1,6,6-trimethylbicyclo[2.1.1]hexane-5-carbaldehyde	unknown	19.699	0.001	synthetic	JAC81	3

91 [a]	Propanoic acid, 2-methyl-, anhydride	2-methylpropanoyl 2-methylpropanoate	unknown	5.111	0.001	YMS	JAC95	2
	4-Methyl-2,4-bis(4'-trimethylsilyloxyphenyl)pentene-1-Dup1 [f]	[g]						
92 [a]	Pyrazine, 2-methyl-5-(1-methylethyl)-	2-methyl-5-propan-2-ylpyrazine	unknown	10.541	0.007	YMS	JAC95	2
	Acetamide, N,N'-ethylenebis(N-nitro-	N-[2-[acetyl(nitro)amino]ethyl]-N-nitroacetamide						
93 [a]	Pyrimidine-2,4(1H,3H)-dione, 5-amino-6-nitroso-	5-amino-6-nitroso-1H-pyrimidine-2,4-dione	unknown	4.021	0.001	SFM	JAC25	2
	Pentane, 1-propoxy-	1-propoxypentane						
94	Sulfurous acid, cyclohexylmethyl heptadecyl ester	cyclohexylmethyl heptadecyl sulfite	unknown	4.923	0.001	YMS	JAC25	2
95 [a]	Tetrahydrofuran, 2-propyl-	2-propyloxolane	unknown	4.024	0.004	synthetic	JAC74C	2
	2,2,5,5,6-Pentamethyl-4,7,9-trioxabicyclo[4.2.1]nonane	1,2,2,5,5-pentamethyl-3,8,9-trioxabicyclo[4.2.1]nonane						

[a] VOC names should be interpreted with caution where more than one VOCs have similar retention time (RT).

[b] The level of confidence assigned to the identity of the VOC. High confidence means there was a good spectral match to the NIST library reference and a Kovats retention index (RI) within 10 points of a literature value available on the NIST WebBook (https://webbook.nist.gov/chemistry/name-ser/), low confidence means the difference in RI of a NIST library suggested VOC to that of the literature value is greater 10 points, and unknown means that a literature RI value could not be found for a NIST library suggested compound.

[c] Standard deviation from the average retention time (RT).

[d] The number of replicates of an isolate fermented in each medium that had the VOC in question. Only VOCs present in at least 2 of 3 media replicate cultures are listed as produced by the *Streptomyces* isolate in question.

[e] common name for the given VOC.

[f] Dup1 suffix means a NIST reference spectra with close identity to the given VOC was not available in the library used in this study.

[g] not applicable or IUPAC name not available.

1-hexanol produced by JAC60 was only detected in the synthetic medium and has various bioactivities (Table 3.3.2.1), including attracting the beetle, *Scolytus schevyrewi*, a pest to apricot trees (producer of 1-hexanol) that promote the death of the host trees (Zhu et al., 2020). Other insects that pose a threat to small fruits have found a way to evolve and sense 1-hexanol. For example, the oriental fruit fly, *Grapholita molesta*, has honed its receptors to sense plant volatiles, including 1-hexanol (L. H. Chen et al., 2019). The same compound disrupts CO_2-mediated activation and host-finding behaviour in *Ades aegypti* mosquitoes (Turner et al., 2011). Geosmin was detected from four of the six isolates (JAC25, JAC60, JAC81, and JAC95) fermented in SFM, one of the six (JAC81) fermented in the synthetic medium, and two of the six (JAC74C and JAC81) fermented in YMS (VOC No. 76, Table 3.3.3.1). Geosmin attracts springtails insects (Becher et al., 2020) but repels *Drosophila* (Stensmyr et al., 2012). Other VOCs function similarly (Table 1.3.3.1); hence, it is crucial to not categorize VOCs as having only one mechanism of action, e.g., insect repellency. A species that uses a given VOC as a communication signal evolve to either be repelled or attracted to the location of that VOC, e.g., a predator to avoid or a food source for nourishment, respectively. Therefore, when selecting VOCs for agricultural uses, one needs to do extensive experiments, including insect repellency testing, to investigate the effects of candidate compounds on pests and their natural predators. In addition, the fermentation media used will determine the VOCs produced.

3.3.4: There was no direct relationship between *Streptomyces* growth and VOC production

An indirect DNA assay (Burton, 1968; Zhao et al., 2013) was performed to determine the growth in the fermentation media of the six isolates selected for replicate VOC analysis. DNA content in the cultures was measured and used to correlate with growth. The DNA was treated with a diphenylamine reagent creating products with typical absorbance at 595 nm. Absorbance was then compared for the cultures with a higher absorbance indicating more growth. Comparison of the amount of VOCs produced and the growth of the six isolates showed no direct relationship across the three media (Figure 3.3.3.2 and Figure 3.3.4.1). Only JAC81 out of six isolates showed that more growth equalled more VOCs produced.

Furthermore, JAC74C showed the least growth in YMS (Figure 3.3.4.1) but produced the most VOCs in YMS (Figure 3.3.3.2). JAC45 and JAC60 grew the most in YMS compared to the other two media (Figure 3.3.4.1). However, the number of VOCs produced by each of the two isolates in YMS was less than those made in SFM and not notably more than those produced in the synthetic medium (Figure 3.3.3.2). The results showed that it might not have been the growth of the isolates but other conditions like the ingredients in the media that determined VOC production. Therefore, optimal conditions for specialized metabolite production are not the same for different species (VanderMolen et al., 2013). More growth may not be more ideal for VOC production. Moreover, variations existed within replicates of the same isolate, i.e., VOC production was not the same across replicate cultures (Table 3.3.3.1), and only VOCs produced in at least two replicate cultures were considered. It should also be noted that VOC collection was done after five days of

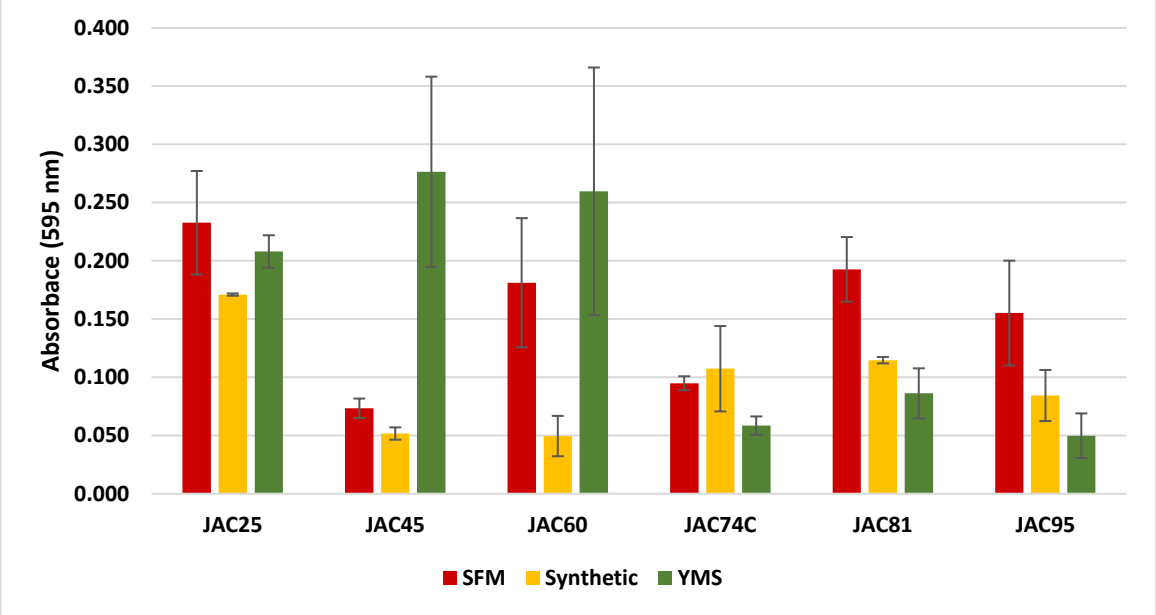


Figure 3.3.4.1: Growth of the *Streptomyces* isolates fermented in SFM, YMS, and the synthetic media. DNA content in the cultures was measured and used to correlate with growth. DNA from the cultures was treated with a diphenylamine reagent creating products with typical absorbance at 595 nm. Error bars represent the standard deviation from the mean calculated for three replicate cultures of the same *Streptomyces* isolate for each medium. Duplicate absorbances for each replicate culture were measured and averaged before calculating the average absorbances for the *Streptomyces* isolates.

fermentation, and VOCs emitted before collection were not trapped for GC-MS analysis. Therefore, cultures that grew the most may have emitted most of their VOCs earlier in the fermentation period and have entered the death phase of its growth closer to the time of collection. At this point, the nutrients in the media would have been primarily used up (Kolter et al., 1993), and VOCs produced are expected to be less.

3.4: Bioassay typing of selected *Streptomyces* isolates

3.4.1: Antibiotic bioactivities varied for phylogenetically different isolates

The 37 *Streptomyces* isolates selected from different OTUs were tested for bioactivities against several indicator organisms, including Gram-positive and Gram-negative bacteria and a fungus (Figure 3.4.1.1). Twelve isolates (JAC24, JAC30L, JAC31, JAC33, JAC45, JAC58, JAC74C, JAC74L, JAC86, JAC101, JAC103, and JAC110) displayed no bioactivity against any of the biological indicators. The remaining 25 isolates produced antibacterials, with ten isolates (JAC25, JAC54, JAC55, JAC68, JAC72, JAC81, JAC88, JAC96, JAC99, JAC122) having activities against Gram-negative bacteria and 19 isolates (JAC17, JAC25, JAC35, JAC47, JAC48, JAC54, JAC55, JAC60, JAC61A, JAC64, JAC68, JAC75, JAC78, JAC81, JAC87, JAC88, JAC90, JAC96, JAC102) having activities against Gram-positive bacteria. Eight isolates (JAC6, JAC35, JAC54, JAC55, JAC75, JAC76, JAC81, and JAC95) had anti-fungal activity.

The isolates produced diverse bioactivities across the three media. For example, supernatant from JAC35 grown in SFM had activity against all indicators except *E. coli* and *P. fluorescence*. JAC35 supernatant from the synthetic medium had similar results to

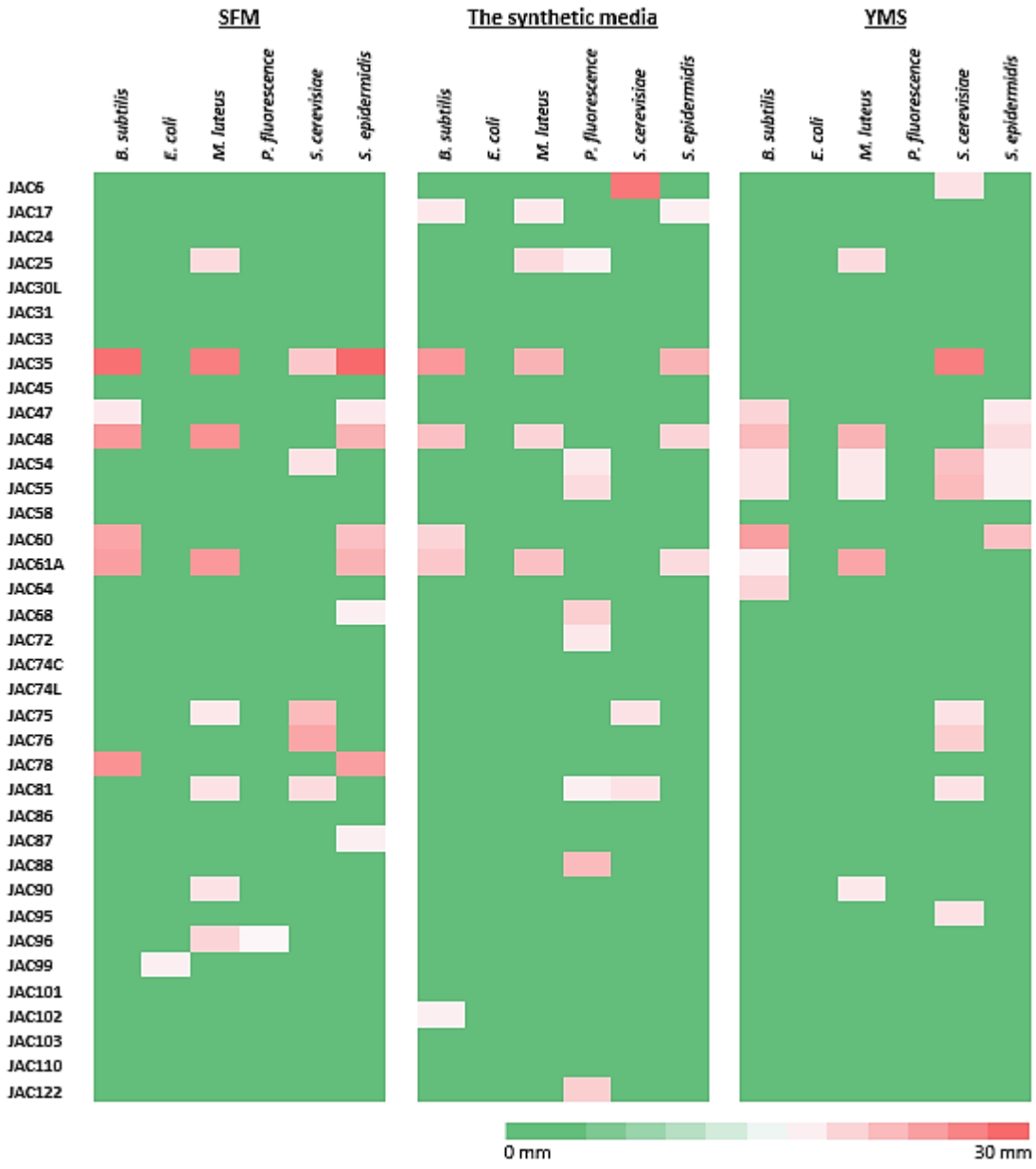


Figure 3.4.1.1: Comparative single-culture supernatant bioassay against six indicator organisms. Supernatants were from the 37 *Streptomyces* isolates representing different OTUs and fermented in SFM, YMS, and the synthetic medium. The data available was for one replicate bioassay analysis. Zones of inhibitions (ZOIs) were measured in mm.

SFM but showed no inhibition against *S. cerevisiae,* while only *S. cerevisiae* was inhibited when exposed to JAC35 supernatant grown on YMS. Only one medium had supernatants showing bioactivities in some cases, e.g., JAC72, JAC102, and JAC122 grown in the synthetic medium, and JAC78, JAC87 JAC96, and JAC99 grown in SFM, and the indicators were not necessarily the same ones inhibited as seen with JAC54.

The six *Streptomyces* isolates used in replicate VOC analysis were fermented in triplicate and tested for bioactivity against the same indicator organisms mentioned above (Figure 3.4.1.2). There were some differences when the results were compared for the bioassays first done with one culture supernatant of each isolate (Figure 3.4.1.1) and subsequently with replicate culture supernatants (Figure 3.4.1.2). JAC25 replicate cultures did not show bioactivity against any biological indicators contrasting with the results obtained before, where *M. luteus* was inhibited with supernatants from the three media and *P. fluorescence* was inhibited by supernatant from the synthetic medium (Figure 3.4.1.1). JAC45 replicate culture supernatants from the synthetic medium displayed bioactivity against *E. coli* and *P. fluorescence* (Figure 3.4.1.2), though the isolate did not show bioactivity when first tested (Figure 3.4.1.1). Likewise, the JAC74C replicate culture supernatants from YMS displayed antibacterial activity against *M. luteus* (Figure 3.4.1.2), but not so when tested at first with one culture supernatant (Figure 3.4.1.1). JAC81 replicate culture supernatants from the synthetic medium did not show bioactivity against *P. fluorescence* and *S. cerevisiae* (Figure 3.4.1.2) but did so when tested at first with supernatant from the same medium (Figure 3.4.1.1). JAC95 replicate culture supernatants from the synthetic medium showed bioactivity against *E. coli* and *P. fluorescence*

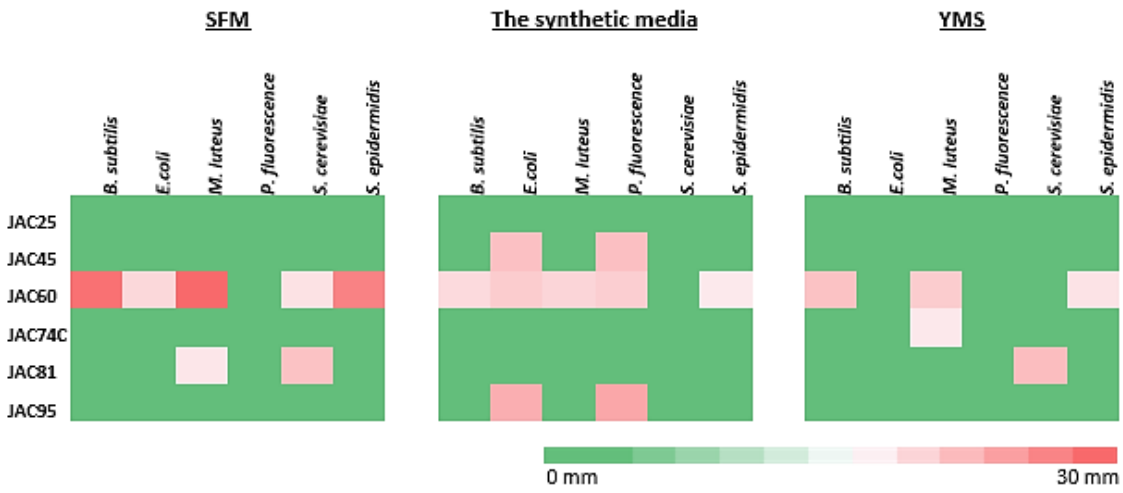


Figure 3.4.1.2: Comparative replicate supernatant bioassay against six indicator organisms. Supernatants were from six *Streptomyces* isolates fermented in triplicate in SFM, YMS, and the synthetic medium. Zones of inhibitions (ZOIs) measured in mm were averaged for the three replicates and used to construct the heat map.

(Figure 3.4.1.2), while only supernatant from YMS inhibited *S. cerevisiae* when JAC95 was first tested (Figure 3.4.1.1). JAC60 replicate culture supernatants inhibited all biological indicators (Figure 3.4.1.2), but only *B. subtilis* and *S. epidermidis* were inhibited by JAC60 single-culture supernatant when tested first (Figure 3.4.1.1).

One should be aware that variability can occur in replicate cultures of the same bacteria as local conditions within a fermentation flask may differ from those of other flasks treated under the same laboratory conditions (Siebenberg et al., 2010). Therefore, these events were not unexpected because slight changes (e.g., the concentration of oxygen in the fermentation flasks) can affect specialized metabolite production, including the volume produced and whether it is enough to cause inhibition of a biological indicator.

Apart from external conditions and uncontrollable circumstances, it is established that media formulation and growth conditions should be considered when targeting a specialized metabolite for production. In the "one strain many compounds" technique, one bacterial strain can produce many different compounds by changing media composition and cultivation conditions (Pan et al., 2019). Hence, one should keep this in mind to capture a bacterial strain's total specialized metabolic capacity.

3.4.2: Phytopathogenicity testing of select *Streptomyces* isolates

The six *Streptomyces* isolates selected for replicate VOC analysis were tested for toxic effects on plants in a simple radish seedlings assay. The *Streptomyces* isolates were compared to a known plant pathogen, *S. scabies* 87-22 (Li et al., 2019), using the respective isolates spores as inoculants to the seedlings. The results (Figure 3.4.2.1) showed that only

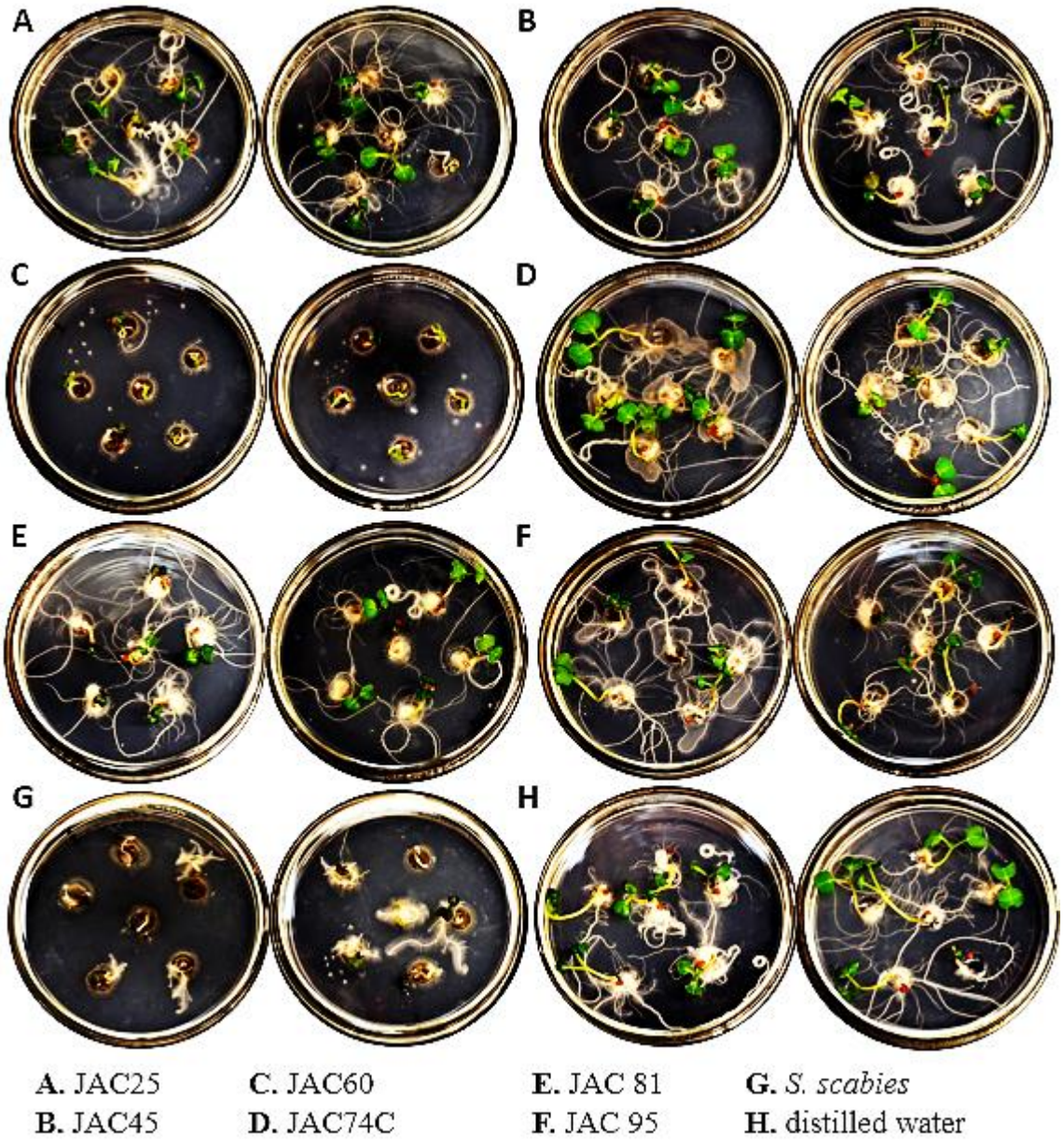


Figure 3.4.2.1: Phytotoxic effects of *Streptomyces* isolates spores on radish seedlings. *S. scabies* 87-22 was used as a positive control and sterile distilled water as a negative control. The test was done in duplicate.

JAC60 significantly hindered the growth of the radish seedlings. The roots and shoots of the seedlings were not well developed compared to the seedlings treated with the other *Streptomyces* spores.

Supernatants collected from the six isolates fermented in the SFM medium were tested for plant toxicity. The average for three of the five seedlings in each treatment (the longest and shortest seedlings were removed) was taken. An ANOVA test was used to compare the radish seedling lengths treated with the SFM supernatants from the isolates and the SFM medium control (Table 3.4.2.1). The ANOVA showed that the p-values for both replicates are less than an alpha value of 0.05. Likewise, the F-values for both the replicates are greater than the critical F-statistics, indicating significant results. When considering both replicate experiments, JAC60 and JAC95 hindered the radish seedlings' growth compared to the SFM medium control, though JAC60 caused the most considerable reduction in plant height (Table 3.4.2.1 and Figure 3.4.2.2). JAC45 and JAC74C appeared to enhance the growth of the radish seedlings (Table 3.4.2.1 and Figure 3.4.2.2) though further analysis is necessary to confirm this. *Streptomyces scabies* 87-22 known to produce phytotoxins in OBB (Johnson et al., 2007) was used to verify that the supernatant test worked, where *S. scabies* 87-22 supernatant from fermented OBB hindered the growth of the radish seedlings when compared to the OBB medium (Table 3.4.2.1 and Figure 3.4.2.2).

It should be noted that there is not a big difference in the variance between the lengths of seedlings treated with *S. scabies* 87-22 OBB supernatant and the OBB control. However, *S. scabies* 87-22 OBB supernatant caused shorter heights in the radish seedlings.

Table 3.4.2.1: Radish seedlings height after treatment with the *Streptomyces* isolates and *S. scabies* 87-22 supernatants from the SFM and OBB media, respectively.

		Replicate1			**Replicate2**		
SUMMARY							
Groups	*Count*	*Sum*	*Average*	*Variance*	*Sum*	*Average*	*Variance*
JAC25	3	124	41.33	56.33	91	30.33	14.33
JAC45	3	146	48.67	158.33	150	50.00	273.00
JAC60	3	65	21.67	12.33	57	19.00	4.00
JAC74C	3	143	47.67	58.33	145	48.33	70.33
JAC81	3	79	26.33	14.33	78	26.00	36.00
JAC95	3	118	39.33	8.33	126	42.00	12.00
SFM-ctrl	3	146	48.67	41.33	115	38.33	20.33
OBB-ctrl	3	81	27.00	1.00			
S. scabies 87-22	3	68	22.67	4.33			

ANOVA	**Replicate1**						**Replicate2**					
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	2185.14	6	364.19	7.29	0.001096	2.85	2430.29	6	405.05	6.59	0.001787	2.85
Within Groups	698.67	14	49.90				860	14	61.43			
Total	2883.81	20					3290.29	20				

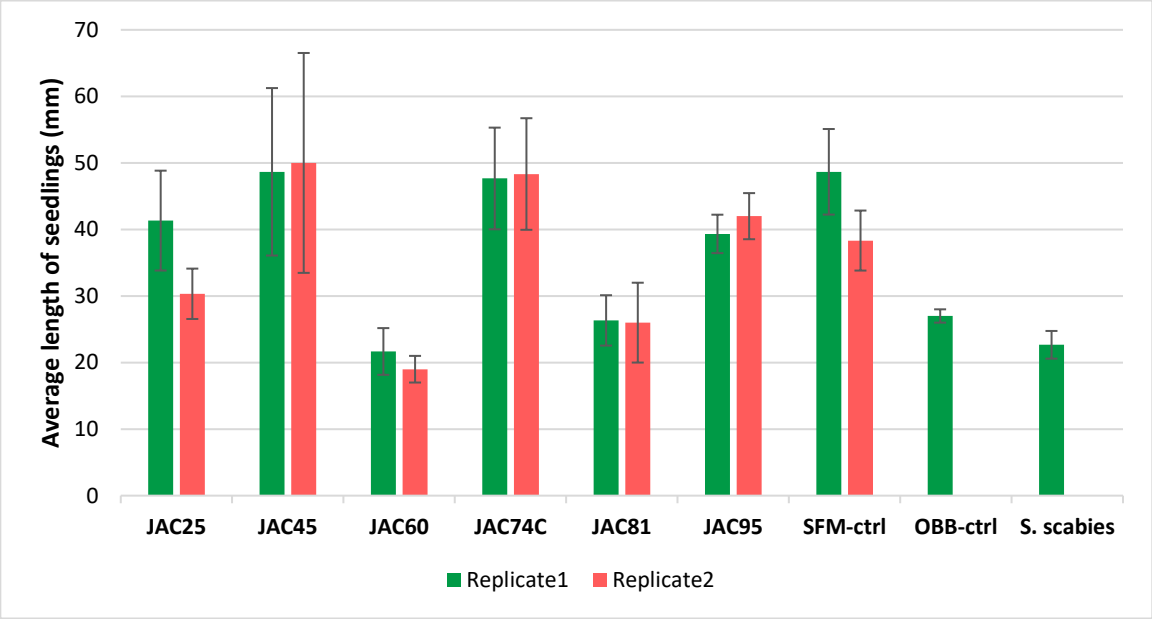


Figure 3.4.2.2: Phytotoxic effects of *Streptomyces* isolates supernatants on radish seedlings. The *Streptomyces* isolates supernatants were drawn from the SFM medium following fermentation. A detailed comparison of the treatments is shown in the ANOVA table (Table 3.4.2.1). The graph depicts averages of technical replicates for the length of seedlings in each supernatant treatment (n=3). The experiment was done in duplicate. Error bars are standard deviations from the means. SFM and Oat-bran broth (OBB) were used as negative controls and *S. scabies* 87-22 supernatant from OBB as a positive control to show that the test worked. There was no positive control for the use of the SFM medium. Only *S. scabies* 87-22 was grown in OBB to collect the supernatant.

In addition, only one replicate treatment was done using the *S. scabies* 87-22 OBB supernatant and the OBB control. The next step would be to repeat the experiment using replicates and with the *Streptomyces* isolates supernatants from OBB, YMS, and the synthetic media.

By and large, the use of SFM supernatants and the *Streptomyces* isolates spores showed comparable results, e.g., JAC60 spores hindered the seedlings' growth the most, as did JAC60 SFM supernatant (Table 3.4.2.1 and Figure 3.4.2.2). These findings suggest that the responsible metabolite(s) produced in SFM medium may also be made when the isolates spores were used as the test substance. Therefore, using supernatants from the fermentation of media or spores for some isolates may not be safe for direct use on plants. Further studies are encouraged to ensure plant safety or VOCs of interest may be isolated for use. One could exploit the plant growth-inhibiting action of JAC60 as a herbicidal strategy, but likewise, further study of the isolate is encouraged.

3.5: *Streptomyces* isolates identification required more than the 16S rDNA in some cases

DNA sequences for the molecular identification of the six isolates used in replicate VOC analysis and phytopathogenicity testing include 16S rDNA and the *rpoB* gene. The six isolates were identified as members of the genus *Streptomyces* from a 16S rDNA search against the NCBI 16S rRNA database. The top hit for a query (the isolate's 16S rDNA sequence) was taken as the species ID for the isolate if it had the highest percent identity to the query sequence. However, if there was more than one top hit with the same percent identity to the query, the *rpoB* gene sequences of those top hits were manually extracted

from the *Streptomyces* species whole-genome sequences obtained from NCBI (accession numbers in Tabe 3.5.1), using *rpoB* primers from this study (Table 2.3.1.1). The top hits *rpoB* sequences were then compared to that of the *Streptomyces* isolate *rpoB* sequence to determine the *Streptomyces* species with the highest percent identity above the recommended minimum of 96–97% percent identity (Adékambi et al., 2009) to the isolate's *rpoB* sequence to determine its ID (Table 3.5.1). The 16S rDNA sequence did not resolve the species identification of four of the six *Streptomyces* isolates (JAC25, JAC60, JAC81, and JAC95). Therefore, their *rpoB* gene sequences were included in determining the species-level IDs (Table 3.5.1).

JAC25 16S rDNA and *rpoB* gene sequences most resemble *S. pratensis*, formerly called *Streptomyces flavogriseus*, and was first named *Actinomyces flavogriseus* by Duché in 1934 (Shirling & Gottlieb, 1969). Following its first isolation, *S. pratensis* has been studied relatively extensively. One study showed that *S. pratensis* was isolated from diseased tomato plants and inhibited pathogenic fungus, *Botrytis cinerea*, infecting the host plant (Lian et al., 2017). Another study showed that *S. pratensis* inhibits the wheat scab causing fungus, *Fusarium graminearum* (Zhang et al., 2020).

JAC45 was identified as *Streptomyces griseoaurantiacus*, first named *Actinomyces griseoaurantiacus* by Krasil'nikov and Yuan in 1965, and then *S. griseoaurantiacus* by Pridham in 1970 (Shirling & Gottlieb, 1972). JAC74C was identified as *Streptomyces xiamenensis*, a *Streptomyces* first isolated from mangrove sediments in the Fujian Province, China (Xu et al., 2009). JAC81 was identified as *Streptomyces sampsonii*, first isolated from garden soil in India (Jain & Jain, 2007). One study reported using *S. sampsonii* to

Table 3.5.1: Molecular identification of select *Streptomyces* isolates.

Isolate	Gene	Query length (nucleotides)	NCBI, BLAST hit [a]	Query cover to reference (%)	Identity to reference (%)	Species ID proposed
JAC25	16S rDNA	1342	*Streptomyces* species: *S. cyaneofuscatus* (NR_115383.1), *S. anulatus* (NR_112527.1), *S. praecox* (NR_115437.1), *S. pratensis* (NR_125616.1)	100	100	*S. pratensis*
	rpoB	709	*Streptomyces* species:			
			S. cyaneofuscatus (NZ_JOEM01000001.1)	100	91.68	
			S. anulatus (NZ_CM003601.1)	100	92.24	
			S. praecox (JF424013.1)	100	92.10	
			S. pratensis (NZ_CP051486.1)	100	98.87	
JAC45	16S rDNA	1372	*Streptomyces griseoaurantiacus* (NR_115231.1)	100	100	*S. griseoaurantiacus*
	rpoB	738	*S. griseoaurantiacus* (NZ_JACERG010000019.1)	100	99.86	
JAC60	16S rDNA	1348	*Streptomyces* species: *S. graminearus* (NR_112461.1), *S. murinus* (NR_112445.1), *S. costaricanus* (NR_041414.1), *S. griseofuscus* (NR_041084.1)	100	100	*S. griseofuscus*
	rpoB	764	*S. murinus* (NZ_MUNF01000071.1)	100	99.21	
			S. griseofuscus (NZ_CP051006.1)	100	99.35	
			S. graminearus (KJ996369.1)	70.68	95.56	
JAC74C	16S rDNA	1354	*Streptomyces xiamenensis* (NR_044035.1)	100	99.41	*S. xiamenensis*
	rpoB	761	*S. xiamenensis* (NZ_CP009922.3)	100	99.60	
JAC81	16S rDNA	1357	*Streptomyces* species: *S. griseochromogenes* (NR_042102.1)	100	100	*S. sampsonii*

			S. resistomycificus (NR_042100.1)			
			S. coelicolor (NR_112305.1)			
			S. sampsonii (NR_112362.1)			
			S. limosus (NR_112279.1)			
			S. felleus (NR_112266.1)			
	rpoB	751	*Streptomyces* sp. GF20			
			S. griseochromogenes (CP016279.1)	100	90.19	
			S. resistomycificus (NZ_KQ948999.1)	100	91.51	
			S. coelicolor (AL645882.2)	100	91.78	
			S. sampsonii (NZ_CP016824.1)	100	99.73	
			S. limosus (KJ996775.1)	72.02	100	
			S. felleus (KT844524.1)	72.02	100	
JAC95	16S rDNA	1345	*Streptomyces* species:	100	100	*S. halstedii*
			S. cyaneus (NR_114820.1)			
			S. griseolus (NR_112493.1)			
			S. halstedii (NR_112275.1)			
	rpoB	743	*S. cyaneus* (NZ_RQJW01000006.1)	100	89.23	
			S. griseolus (NZ_JOFC01000040.1)	100	99.60	
			S. halstedii (NZ_JAAGLQ010000138.1)	100	100	

[a] The accession numbers are provided for *Streptomyces* species obtained using NCBI BLAST against the 16S ribosomal RNA database. The *rpoB* gene sequences of each isolate's 16S ribosomal RNA database top hits were manually extracted from available NCBI whole-genome sequences. Accession numbers of species used for extraction of *rpoB* gene sequences are provided.

degrade poultry feathers to make a fertilizer that improved wheat seed germination and significantly increased plant height (Jain et al., 2016), demonstrating the indirect use of the *Streptomyces* and products.

JAC95 was identified as *Streptomyces halstedii*, a reported producer of compounds that inhibit the growth of fungi, including *Aspergillus fumigatus, Mucor hiemalis, Penicillium roqueforti,* and *Paecilomyces variotii* (Frandberg et al., 2000). Another study showed that red-pepper seeds soaked in a cell-free culture broth (or supernatant) of *S. halstedii* inhibited the fungal phytopathogen, *Phytophthora capsica*, that causes Phytophthora blight in the peppers (Joo, 2005).

JAC60's was identified as *Streptomyces griseofuscus* (Table 3.5.1). Only JAC60 showed severe growth inhibition to the radish seedlings in this study (Table 3.4.2.1, Figure 3.4.2.1, and Figure 3.4.2.2). An antiSMASH analysis to predict specialized metabolite BGCs in the genome of JAC60 species ID, *S. griseofuscus* (Appendix A.3.5.1), and the species with the following closest *rpoB* percent identity, *S. murinus* (Appendix A.3.5.2), was done. A predicted BGC in *S. murinus* had 89% identity to actinomycin D, an antibiotic implicated in inhibiting auxin-induced elongation in oats seedlings (Cleland, 1965). *Streptomyces murinus* was classified as the same species as *S. costaricanus*, by MLSA of five protein-coding genes, including the *rpoB* gene (Komaki, 2021). Actinomycin D was identified in *S. costaricanus* (Liu et al., 2019). Therefore, if actinomycin D was produced by JAC60 and caused the inhibition of the radish seedlings development (Table 3.4.2.1, Figure 3.4.2.1 and Figure 3.4.2.2), *S. murinus* may be the correct ID, or the *S. griseofuscus* strain in this study (JAC60) may have a BGC that produce a plant growth-inhibiting

specialized metabolite. It should be noted that the *rpoB* percent identities between JAC60 and *S. griseoufuscus,* and JAC60 and *S. murinus* were close (99.35% and 99.21%, respectively). Therefore, it is good to use another protein-coding gene like those mentioned in Table 1.2.2.1 to concatenate with the *rpoB* to confirm JAC60's species ID. The same is encouraged for JAC95.

3.6: Concluding remarks

Streptomyces continues to be a resource for relevant natural products. In this study, 133 *Streptomyces* isolates obtained from St. John's, Newfoundland, Canada, showed diversity in the specialized metabolites they produced. It was discussed that VOC biosynthesis is often linked to specialized metabolic pathways, and along these pathways, intermediate products are transformed into new compounds by enzymes to create shunt products. Many other VOCs are yet to be discovered from the *Streptomyces* genus as relatively few are studied to date, suggesting the need for future research of bacteria VOCs that could replace chemicals for sustainable farming practices. For those *Streptomyces* studied, their detected VOCs were not tested for bioactivity in the scope of this thesis. However, some VOCs detected from the *Streptomyces* isolates were described elsewhere (Table 1.3.3.1 and Table 3.3.2.1); nevertheless, these compounds are the same, so reported activities are assumed to be equivalent. Extensive studies must be done when selecting *Streptomyces* or derived products for agriculture because the effects of a given VOC or other specialized metabolite vary across plant and insect species.

This thesis provided new knowledge of *Streptomyces* isolates from Newfoundland and the VOCs they produce. It is vital to determine the effects of the *Streptomyces* spores

or supernatant on the agricultural crop in question. To apply these VOCs in farming practices, experiments including plant and insect assays should be done. Depending on the crop and insect that predates on that plant, bioassays using those specific seedlings and insects are encouraged. In cases where an isolate produces a VOC that repels a predatory insect yet is phytopathogenic, isolating the VOC for use is an alternative. Other solutions include sequencing the genomes of such isolates and identifying BGCs linked to specialized metabolites of interest. These BGCs may then be heterologously expressed in "generally regarded as safe" species to produce the desired compounds. JAC60 could be used as an antiweed strategy. However, plant assays with the weed of interest should be conducted to ensure that only the weed is negatively impacted during application.

References

Adamek, M., Alanjary, M., & Ziemert, N. (2019). Applied evolution: Phylogeny-based approaches in natural products research. *Natural Product Reports*, *36*(9), 1295–1312. https://doi.org/10.1039/c9np00027e

Adékambi, T., Drancourt, M., & Raoult, D. (2009). The rpoB gene as a tool for clinical microbiologists. *Trends in Microbiology*, *17*(1), 37–45. https://doi.org/10.1016/j.tim.2008.09.008

Adekambi, T., Shinnick, T. M., Raoult, D., & Drancourt, M. (2008). Complete rpoB gene sequencing as a suitable supplement to DNA – DNA hybridization for bacterial species and genus delineation. *International Journal of Systematic and Evolutionary Microbiology*, 1807–1814. https://doi.org/10.1099/ijs.0.65440-0

Agarwala, R., Barrett, T., Beck, J., Benson, D. A., Bollin, C., Bolton, E., Bourexis, D., Brister, J. R., Bryant, S. H., Canese, K., Cavanaugh, M., Charowhas, C., Clark, K., Dondoshansky, I., Feolo, M., Fitzpatrick, L., Funk, K., Geer, L. Y., Gorelenkov, V., … Zbicz, K. (2018). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, *46*(D1), D8–D13. https://doi.org/10.1093/nar/gkx1095

Alam, M. T., Merlo, M. E., Takano, E., & Breitling, R. (2010). Genome-based phylogenetic analysis of Streptomyces and its relatives. *Molecular Phylogenetics and Evolution*, *54*(3), 763–772. https://doi.org/10.1016/j.ympev.2009.11.019

Álvarez-Álvarez, R., Martínez-Burgo, Y., Pérez-Redondo, R., Braña, A. F., Martín, J. F., & Liras, P. (2013). Expression of the endogenous and heterologous clavulanic acid cluster in Streptomyces flavogriseus: Why a silent cluster is sleeping. *Applied Microbiology and Biotechnology*, *97*(21), 9451–9463. https://doi.org/10.1007/s00253-013-5148-7

Alves, T. J. S., Murcia, A., Wanumen, A. C., Wanderley-Teixeira, V., Teixeira, Á. A. C., Ortiz, A., & Medina, P. (2019). Composition and Toxicity of a Mixture of Essential Oils Against Mediterranean Fruit Fly, Ceratitis capitata (Wiedemann) (Diptera: Tephritidae). *Journal of Economic Entomology*, *112*(1), 164–172. https://doi.org/10.1093/jee/toy275

Andreu-Coll, L., Noguera-Artiaga, L., Carbonell-Barrachina, Á. A., Legua, P., & Hernández, F. (2020). Volatile composition of prickly pear fruit pulp from six Spanish cultivars. *Journal of Food Science*, *85*(2), 358–363. https://doi.org/10.1111/1750-3841.15001

Antony-Babu, S., Stien, D., Eparvier, V., Parrot, D., Tomasi, S., & Suzuki, M. T. (2017). Multiple Streptomyces species with distinct secondary metabolomes have identical 16S rRNA gene sequences. *Scientific Reports*, *September 2016*, 1–8. https://doi.org/10.1038/s41598-017-11363-1

Apprill, A., Mcnally, S., Parsons, R., & Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology*, *75*(2), 129–137. https://doi.org/10.3354/ame01753

Audrain, B., Farag, M. A., Ryu, C., & Ghigo, J. (2015). Role of bacterial volatile compounds in bacterial biology. *FEMS Microbiology Reviews*, *39*, 222–233. https://doi.org/10.1093/femsre/fuu013

Avalos, M., Garbeva, P., Raaijmakers, J. M., & van Wezel, G. P. (2019). Production of ammonia as a low-cost and long-distance antibiotic strategy by Streptomyces species. *ISME Journal*. https://doi.org/10.1038/s41396-019-0537-2

Baker, G. C., Smith, J. J., & Cowan, D. A. (2003). Review and re-analysis of domain-specific 16S primers. *Journal of Microbiological Methods*, *55*(3), 541–555. https://doi.org/10.1016/j.mimet.2003.08.009

Balloux, F., Brønstad Brynildsrud, O., van Dorp, L., Shaw, L. P., Chen, H., Harris, K. A., Wang, H., & Eldholm, V. (2018). From Theory to Practice: Translating Whole-Genome Sequencing (WGS) into the Clinic. *Trends in Microbiology*, *26*(12), 1035–1048. https://doi.org/10.1016/j.tim.2018.08.004

Baltz, R. H. (2017). Gifted microbes for genome mining and natural product discovery. *Journal of Industrial Microbiology & Biotechnology*, *44*(4), 573–588. https://doi.org/10.1007/s10295-016-1815-x

Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H.-P., Clement, C., Ouhdouch, Y., & Wezel, G. P. Van. (2015). Taxonomy, Physiology, and Natural Products of Actinobacteria. *American Society for Microbiology*, *80*(1), 1–43. https://doi.org/10.1128/MMBR.00019-15.Address

Becher, P. G., Bengtsson, M., Hansson, B. S., & Witzgall, P. (2010). Flying the fly: Long-range flight behavior of drosophila melanogaster to attractive odors. *Journal of Chemical Ecology*, *36*(6), 599–607. https://doi.org/10.1007/s10886-010-9794-2

Becher, P. G., Verschut, V., Bibb, M. J., Bush, M. J., Molnár, B. P., Barane, E., Al-bassam, M. M., Chandra, G., Song, L., Challis, G. L., Buttner, M. J., & Flärdh, K. (2020). Developmentally regulated volatiles geosmin and 2-methylisoborneol attract a soil arthropod to Streptomyces bacteria promoting spore dispersal. *Nature Microbiology*. https://doi.org/10.1038/s41564-020-0697-x

Bibb, M. J. (2005). Regulation of secondary metabolism in streptomycetes. *Current Opinion in Microbiology*, *8*, 208–215. https://doi.org/10.1016/j.mib.2005.02.016

Blackwood, K. Y. M. S., He, C., Gunton, J., Turenne, C. Y., Wolfe, J., & Kabani, A. M. (2000). Evaluation of recA Sequences for Identification of Mycobacterium Species. *Journal of Clinical Microbiology*, *38*(8), 2846–2852.

Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., Van Wezel, G. P., Medema,

M. H., & Weber, T. (2021). AntiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Research*, *49*(W1), W29–W35. https://doi.org/10.1093/nar/gkab335

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., Medema, M. H., & Weber, T. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Research*, *47*(W1), W81–W87. https://doi.org/10.1093/nar/gkz310

Burton, K. (1968). [105] Determination of DNA concentration with diphenylamine. *Methods in Enzymology*, *12*(PART B), 163–166. https://doi.org/10.1016/0076-6879(67)12127-7

Cane, D. E., He, X., Kobayashi, S., Mura, S. O., & Ikeda, H. (2006). Geosmin Biosynthesis in Streptomyces avermitilis. Molecular Cloning, Expression, and Mechanistic Study of the Germacradienol/Geosmin Synthase. *Journal of Antibiotics*, *59*(8), 471–479.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N., & Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *PNAS*, *108*(SUPPL. 1), 4516–4522. https://doi.org/10.1073/pnas.1000080107

Case, R. J., Boucher, Y., Dahllöf, I., Holmström, C., Doolittle, W. F., & Kjelleberg, S. (2007). Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Applied and Environmental Microbiology*, *73*(1), 278–288. https://doi.org/10.1128/AEM.01177-06

Chater, K. F. (2006). Streptomyces inside-out: a new perspective on the bacteria that provide us with antibiotics. *Philosophical Transactions of the Royal Society B*, *361*, 761–768. https://doi.org/10.1098/rstb.2005.1758

Chaudhary, A., Kaur, P., Singh, B., & Pathania, V. (2009). Chemical Composition of Hydrodistilled and Solvent Volatiles Extracted from Woodchips of Himalayan Cedrus: Cedrus deodara (Roxb.) Loud. *Natural Product Communications*, *4*(9), 1257–1260.

Chaudhary, A., Sood, S., Das, P., Kaur, P., Mahajan, I., Gulati, A., & Singh, B. (2014). Synthesis of novel antimicrobial aryl himachalene derivatives from naturally occurring himachalenes. *Experimental and Clinical Sciences*, *13*, 1216–1225.

Chen, J., Lü, J., He, Z., Zhang, F., Zhang, S., & Zhang, H. (2020). Investigations into the production of volatile compounds in Korla fragrant pears (Pyrus sinkiangensis Yu). *Food Chemistry*, *302*(April 2019), 125337. https://doi.org/10.1016/j.foodchem.2019.125337

Chen, L. H., Tian, K., Wang, G. R., Xu, X. L., He, K. H., Liu, W., & Wu, J. X. (2019). The general odorant receptor GmolOR9 from Grapholita molesta (Lepidoptera: Tortricidae) is mainly tuned to eight host-plant volatiles. *Insect Science*, *6*, 1–11.

https://doi.org/10.1111/1744-7917.12725

Cheng, Z., McCann, S., Faraone, N., Clarke, J. A., Hudson, E. A., Cloonan, K., Hillier, N. K., & Tahlan, K. (2020). Production of plant-associated volatiles by select model and industrially important streptomyces spp. *Microorganisms*, *8*(11), 1–20. https://doi.org/10.3390/microorganisms8111767

Chevrette, M. G., Carlos-shanley, C., Louie, K. B., Bowen, B. P., Northen, T. R., & Currie, C. R. (2019). Taxonomic and Metabolic Incongruence in the Ancient Genus Streptomyces. *Frontiers in Microbiology*, *10*, 1–12. https://doi.org/10.3389/fmicb.2019.02170

Chevrette, M. G., Carlson, C. M., Ortega, H. E., Thomas, C., Ananiev, G. E., Barns, K. J., Book, A. J., Cagnazzo, J., Carlos, C., Flanigan, W., Grubbs, K. J., Horn, H. A., Hoffmann, F. M., Klassen, J. L., Knack, J. J., Lewin, G. R., Mcdonald, B. R., Muller, L., Melo, W. G. P., … Currie, C. R. (2019). The antimicrobial potential of Streptomyces from insect microbiomes. *Nature Communications*, *16*, 1–11. https://doi.org/10.1038/s41467-019-08438-0

Choi, Y. J., & Lee, S. Y. (2013). Microbial production of short-chain alkanes. *Nature*, *502*. https://doi.org/10.1038/nature12536

Chu, S. S., Jiang, G. H., & Liu, Z. L. (2011). Insecticidal compounds from the essential oil of Chinese medicinal herb Atractylodes chinensis. *Pest Management Science*, *67*(10), 1253–1257. https://doi.org/10.1002/ps.2180

Chung, T. Y., Eiserich, J. P., & Shibamoto, T. (1993). Volatile Compounds Isolated from Edible Korean Chamchwi (Aster scaber Thunb). *Journal of Agricultural and Food Chemistry*, *41*(10), 1693–1697. https://doi.org/10.1021/jf00034a033

Ciccarelli, D., Giovanelli, S., & Pistelli, L. (2016). Essential Oils from Anthemis maritima Flowers: Infraspecific Variability along the Adriatic Coast (Italy). *Chemistry and Biodiversity*, *13*(5), 561–570. https://doi.org/10.1002/cbdv.201500184

Citron, C. A., Barra, L., Wink, J., & Dickschat, J. S. (2015). Volatiles from nineteen recently genome sequenced actinomycetes. *Organic and Biomolecular Chemistry*, *13*(9), 2673–2683. https://doi.org/10.1039/c4ob02609h

Citron, C. A., Gleitzmann, J., Laurenzano, G., Pukall, R., & Dickschat, J. S. (2012). Terpenoids are Widespread in Actinomycetes : A Correlation of Secondary Metabolism and Genome Data. *Chembiochem*, *13*, 202–214. https://doi.org/10.1002/cbic.201100641

Citron, C. A., Rabe, P., & Dickschat, J. S. (2012). The Scent of Bacteria: Headspace Analysis for the Discovery of Natural Products. *Journal of Natural Products*, *75*, 1765–1776. https://doi.org/10.1021/np300468h

Cleland, R. (1965). Auxin-induced cell wall loosening in the presence of actinomycin D.

Plant Physiology, *40*(4), 595–600. https://doi.org/10.1104/pp.40.4.595

Cordovez, V., Carrion, V. J., Etalo, D. W., Mumm, R., Zhu, H., van Wezel, G. P., & Raaijmakers, J. M. (2015). Diversity and functions of volatile organic compounds produced by Streptomyces from a disease-suppressive soil. *Frontiers in Microbiology*, *6*, 1–13. https://doi.org/10.3389/fmicb.2015.01081

Crawford, J. M., & Clardy, J. (2011). Bacterial symbionts and natural products. *Chemical Communications*, *47*, 7559–7566. https://doi.org/10.1039/c1cc11574j

Curran, A. M., Rabin, S. I., Prada, P. A., & Furton, K. G. (2005). Comparison of the volatile organic compounds present in human odor using SPME-GC/MS. *Journal of Chemical Ecology*, *31*(7), 1607–1619. https://doi.org/10.1007/s10886-005-5801-4

DeLeo, F. R., Chen, L., Porcella, S. F., Martens, C. A., Kobayashi, S. D., Porter, A. R., Chavda, K. D., Jacobs, M. R., Mathema, B., Olsen, R. J., Bonomo, R. A., Musser, J. M., & Kreiswirth, B. N. (2014). Molecular dissection of the evolution of carbapenem-resistant multilocus sequence type 258 Klebsiella pneumoniae. *Pnas*, *111*(13), 4988–4993. https://doi.org/10.1073/pnas.1321364111

Demain, A. L., & Fang, A. (1995). Emerging concepts of secondary metabolism in Actinomycetes. *Actinomycetologica*, *9*(2), 98–117.

Devulder, G., Perouse de Montclos, M., & Flandrois, J. P. (2005). A multigene approach to phylogenetic analysis using the genus Mycobacterium as a model. *International Journal of Systematic and Evolutionary Microbiology*, *55*, 293–302. https://doi.org/10.1099/ijs.0.63222-0

Dias, M. P., Bastos, M. S., Xavier, V. B., Cassel, E., Astarita, L. V., & Santarém, E. R. (2017). Plant growth and resistance promoted by Streptomyces spp. in tomato. *Plant Physiology and Biochemistry*, *118*, 479–493. https://doi.org/10.1016/j.plaphy.2017.07.017

Dickschat, J. S. (2016). Bacterial terpene cyclases. *Natural Product Reports*, *33*, 87–110. https://doi.org/10.1039/c5np00102a

Dickschat, J. S., Pahirulzaman, K. A. K., Rabe, P., & Klapschinski, T. A. (2014). An Improved Technique for the Rapid Chemical Characterisation of Bacterial Terpene Cyclases. *Chembiochem Communications*, *15*, 810–815. https://doi.org/10.1002/cbic.201300763

Dua, V. K., Alam, M. F., Pandey, A. C., Rai, S., Chopra, A. K., Kaul, V. K., & Dash, A. P. (2008). Insecticidal activity of Valeriana jatamansi (Valerianaceae) against mosquitoes. *Journal of the American Mosquito Control Association*, *24*(2), 315–318. https://doi.org/10.1649/0010-065x(2003)057[0114:antfcg]2.0.co;2

Dweck, H. K. M., Ebrahim, S. A. M., Thoma, M., Mohamed, A. A. M., Keesey, I. W., Trona, F., Lavista-Llanos, S., Svato, A., Sachsea, S., Knadena, M., & Hanssona, B. S. (2015). Pheromones mediating copulation and attraction in Drosophila. *PNAS*,

112(21), E2829–E2835. https://doi.org/10.1073/pnas.1504527112

Dworkin, M. (2006). *The Prokaryotes: A Handbook on the Biology of Bacteria* (M. Dworkin (ed.); 3rd ed.). Springer.

Ehsani, A., Rezaeiyan, A., Hashemi, M., Aminzare, M., Jannat, B., & Afshari, A. (2019). Antibacterial activity and sensory properties of Heracleum persicum essential oil, nisin, and Lactobacillus acidophilus against Listeria monocytogenes in cheese. *Veterinary World*, *12*(1), 90–96. https://doi.org/10.14202/vetworld.2019.90-96

El-Din Mohamed, W. S., & Zaki, D. F. A. (2019). Evaluation of antagonistic actinomycetes isolates as biocontrol agents against wastewater-associated bacteria. *Water Science and Technology*, *79*(12), 2310–2317. https://doi.org/10.2166/wst.2019.231

El-Sayed, A. M. (2007). *The Pherobase: Database of Insect Pheromones and Semiochemicals.* http://www.pherobase.net

Elghwaji, W., El-Sayed, A. M., El-Deeb, K. S., & ElSayed, A. M. (2017). Chemical Composition, Antimicrobial and Antitumor Potentiality of Essential Oil of Ferula tingitana L. Apiaceae Grow in Libya. *Pharmacognosy Magazine*, *13*. https://doi.org/10.4103/pm.pm

Emerson de Lima Procópio, R., Reis da Silva, I., Martins, M. K., Lucio de Azevedo, J., & Magali de Araújo, J. (2012). Antibiotics produced by Streptomyces. *Brazilian Journal of Infectious Diseases*, *16*(5), 466–471. https://doi.org/10.1016/j.bjid.2012.08.014

Enright, M. C., & Spratt, B. G. (1999). Multilocus sequence typing. *Trends in Microbiology*, *7*(12), 482–487.

Erbilgin, N., Gillette, N. E., Mori, S. R., Stein, J. D., Owen, D. R., & Wood, D. L. (2007). Acetophenone as an Anti-attractant for the Western Pine Beetle, Dendroctonus Brevicomis LeConte (Coleoptera: Scolytidae). *Journal of Chemical Ecology*, *33*, 817–823. https://doi.org/10.1007/s10886-007-9267-4

Errard, A., Ulrichs, C., Kühne, S., Mewis, I., Drungowski, M., Schreiner, M., & Baldermann, S. (2015). Single- versus Multiple-Pest Infestation Affects Differently the Biochemistry of Tomato (Solanum lycopersicum 'Ailsa Craig'). *Journal of Agricultural and Food Chemistry*, *63*(46), 10103–10111. https://doi.org/10.1021/acs.jafc.5b03884

Espinoza, J., Urzúa, A., Tampe, J., Parra, L., & Quiroz, A. (2016). Repellent Activity of the Essential Oil from the Heartwood of Pilgerodendron uviferum (D. Don) Florin against Aegorhinus superciliosus (Coleoptera: Curculionidae). *Molecules*, *21*(4), 1–7. https://doi.org/10.3390/molecules21040533

Farrelly, V., Rainey, F. A., & Stackebrandt, E. (1995). Effect of Genome Size and rrn Gene Copy Number on PCR Amplification of 16S rRNA Genes from a Mixture of

Bacterial Species. *American Society for Microbiology*, *61*(7), 2798–2801.

Feng, Y., & Zhang, A. (2017). A Floral Fragrance, Methyl Benzoate, Is An Efficient Green Pesticide. *Scientific Reports*, *7*(42168), 1–9. https://doi.org/10.1038/srep42168

Fox, G. E., Wisotzkey, J. D., & Jurtshuk, JR, P. (1992). How Close Is Close : 1 s rRNA Sequence Identity May Not Be Sufficient To Guarantee Species Identity. *International Journal of Systematic Bacteriology*, *42*(1), 166–170.

Franco, T. A., Xu, P., Brito, N. F., Oliveira, D. S., Wen, X., Moreira, M. F., Unelius, C. R., Leal, W. S., & Melo, A. C. A. (2018). Reverse Chemical Ecology-Based Approach Leading to the Accidental Discovery of Repellents for Rhodnius Prolixus, a Vector of Chagas Diseases Refractory to DEET. *Insect Biochemistry and Molecular Biology*, *103*, 46–52. https://doi.org/10.1016/j.ibmb.2018.10.004

Frandberg, E., Petersson, C., Lundgren, L. N., & Schnurer, J. (2000). Streptomyces halstedii K122 produces the antifungal compounds bafilomycin B1 and C1. *Canadian Journal of Microbiology*, *46*(8), 753–758. https://doi.org/10.1139/cjm-46-8-753

Gabaldón, T., & Koonin, E. V. (2013). Functional and evolutionary implications of gene orthology. *Nature Reviews Genetics*, *14*(5), 360–366. https://doi.org/10.1038/nrg3456

Garbeva, P., Hordijk, C., Gerards, S., & Boer, W. De. (2014). Volatile-mediated interactions between phylogenetically different soil bacteria. *Frontiers in Microbiology*, *5*. https://doi.org/10.3389/fmicb.2014.00289

Gerber, N., & Brunswick, N. (1965). Geosmin , Earthy-Smelling Substance Isolated from Actinomycetes. *American Society for Microbiology*, *13*(6), 935–938.

Gerber, N. N. (1969). A Volatile Metabolite of Actinomycetes, 2-methylisoborneol. *The Journal of Antibiotics*, 508–509.

Gevers, D., Cohan, F. M., Lawrence, J. G., Spratt, B. G., Coenye, T., Feil, E. J., Stackebrandt, E., Peer, Y. Van De, Vandamme, P., Thompson, F. L., & Swings, J. (2005). Opinion: Re-evaluating prokaryotic species. *Nature Reviews Microbiology*, *3*(September), 733–740.

Gil, R., Silva, F. J., Pereto, J., & Moya, A. (2004). Determination of the Core of a Minimal Bacterial Gene Set. *American Society for Microbiology*, *68*(3), 518–537. https://doi.org/10.1128/MMBR.68.3.518

Glaeser, S. P., & Kämpfer, P. (2015). Multilocus sequence analysis (MLSA) in prokaryotic taxonomy. *Systematic and Applied Microbiology*, *38*(4), 237–245. https://doi.org/10.1016/j.syapm.2015.03.007

Gliessman, S. R. (1990). Agroecology: Researching the Ecological Basis for Sustainable Agriculture. In S. R. Gliessman (Ed.), *Angewandte Chemie International Edition,*

6(11), 951–952. (1st ed., Vol. 13, Issue April). Springer-Verlag. https://doi.org/10.1007/978-1-4612-3252-0

González-Torres, P., Rodríguez-Mateos, F., Antón, J., & Gabaldón, T. (2019). Impact of Homologous Recombination on the Evolution of Prokaryotic Core Genomes. *MBio*, *10*(1), 1–17.

Govindarajan, M., Rajeswary, M., & Benelli, G. (2016). Chemical composition, toxicity and non-target effects of Pinus kesiya essential oil: An eco-friendly and novel larvicide against malaria, dengue and lymphatic filariasis mosquito vectors. *Ecotoxicology and Environmental Safety*, *129*, 85–90. https://doi.org/10.1016/j.ecoenv.2016.03.007

Groenhagen, U., Maczka, M., Dickschat, J. S., & Schulz, S. (2014). *Streptopyridines , volatile pyridine alkaloids produced by Streptomyces sp . FORM5*. 1421–1432. https://doi.org/10.3762/bjoc.10.146

Gu, Y., Mo, M., Zhou, J., Zou, C., & Zhang, K. (2007). Evaluation and identification of potential organic nematicidal volatiles from soil bacteria. *Soil Biology & Biochemistry*, *39*, 2567–2575. https://doi.org/10.1016/j.soilbio.2007.05.011

Guo, Y., Zheng, W., Rong, X., & Huang, Y. (2008). A multilocus phylogeny of the Streptomyces griseus 16S rRNA gene clade: Use of multilocus sequence analysis for streptomycete systematics. *International Journal of Systematic and Evolutionary Microbiology*, *58*(1), 149–159. https://doi.org/10.1099/ijs.0.65224-0

Gupta, S., Gupta, C., Garg, A. P., & Prakash, D. (2015). A biotechnological approach to microbial based perfumes and flavours. *Journal of Microbiology & Experimentation*, *2*(1), 11–18. https://doi.org/10.15406/jmen.2015.02.00034

Gust, B., Challis, G. L., Fowler, K., Kieser, T., & Chater, K. F. (2003). PCR-targeted Streptomyces Gene Replacement Identifies a Protein Domain Needed for Biosynthesis of the Sesquiterpene Soil Odor Geosmin. *PNAS*, *100*(4), 1541–1546.

Han, J. H., Cho, M. H., & Kim, S. B. (2012). Ribosomal and protein coding gene based multigene phylogeny on the family Streptomycetaceae. *Systematic and Applied Microbiology*, *35*(1), 1–6. https://doi.org/10.1016/j.syapm.2011.08.007

Hansen, M. C., Tolker-nielsen, T., Givskov, M., & Molin, S. (1998). Biased 16S rDNA PCR amplification caused by interference from DNA flanking the template region. *FEMS Microbiology Ecology*, *26*, 141–149.

Haocheng, L., An, K., Su, S., Yu, Y., Wu, J., Xiao, G., & Xu, Y. (2020). Aromatic Characterization of Mangoes (Mangifera indica L.) Using Solid Phase Extraction Coupled with Gas Chromatography–Mass Spectrometry and Olfactometry and Sensory Analyses. *Foods*, *9*(75), 1–20.

Heinsch, S. C., Hsu, S., Otto-hanson, L., Kinkel, L., & Smanski, M. J. (2019). Complete genome sequences of Streptomyces spp. isolated from disease-suppressive soils.

BMC Genomics, *20*.

Heringa, P. K. (1981). Soils of the Avalon Peninsula, Newfoundland. *Research Branch, Agriculture Canada*, *3*, 117.

Heul, H. U. van der, Bilyk, B. L., McDowall, kenneth J., Seipke, R. F., & Wezel, G. P. Van. (2018). Natural Product Reports Regulation of antibiotic production in Actinobacteria: new perspectives from the post-genomic era. *Natural Product Reports*, *35*, 575–604. https://doi.org/10.1039/c8np00012c

Ho, L. K., Daniel-Ivad, M., Jeedigunta, S. P., Li, J., Iliadi, K. G., Boulianne, G. L., Hurd, T. R., Smibert, C. A., & Nodwell, J. R. (2020). Chemical entrapment and killing of insects by bacteria. *Nature Communications*, *11*(1), 1–9. https://doi.org/10.1038/s41467-020-18462-0

Högnadóttir, Á., & Rouseff, R. L. (2003). Identification of aroma active compounds in orange essence oil using gas chromatography–olfactometry and gas chromatography– mass spectrometry. *Journal of Chromatography A*, *998*(1–2), 201–211. https://doi.org/10.1016/S0021-9673(03)00524-7

Holmes, D. E., Nevin, K. P., & Lovley, D. R. (2004). Comparison of 16S rRNA, nifD, recA, gyrB, rpoB and fusA genes within the family Geobacteraceae fam. nov. *International Journal of Systematic and Evolutionary Microbiology*, *54*, 1591–1599. https://doi.org/10.1099/ijs.0.02958-0

Hu, Y., Chou, W. K. W., Hopson, R., & Cane, D. E. (2011). Genome Mining in Streptomyces Clavuligerus: Expression and Biochemical Characterization of Two New Cryptic Sesquiterpene Synthases. *Cell Press*, *18*(1), 32–37. https://doi.org/10.1016/j.chembiol.2010.11.008

Hummadi, E. H., Dearden, A., Generalovic, T., Clunie, B., Harrott, A., Cetin, Y., Demirbek, M., Khoja, S., Eastwood, D., Dudley, E., Hazir, S., Touray, M., Ulug, D., Hazal Gulsen, S., Cimen, H., & Butt, T. (2020). Volatile organic compounds of Metarhizium brunneum influence the efficacy of entomopathogenic nematodes in insect control. *Biological Control*, *155*(December 2020), 104527. https://doi.org/10.1016/j.biocontrol.2020.104527

Ikawa, M., Sasner, J. J., & Haney, J. F. (2001). Activity of cyanobacterial and algal odor compounds found in lake waters on green alga Chlorella pyrenoidosa growth. *Hydrobiologia*, *443*, 19–22.

Ikeda, H., Kotaki, H., & Omura, S. (1987). Genetic studies of avermectin biosythesis in Streptomyces avermitilis. *Journal of Bacteriology*, *169*(12), 5615–5621. https://doi.org/10.1128/jb.169.12.5615-5621.1987

Inerney, J. O. M. C., Wilkinson, M., Patching, J. W., Embley, T. M., & Powell, R. (1995). Recovery and Phylogenetic Analysis of Novel Archaeal rRNA Sequences from a Deep-Sea Deposit Feeder. *Applied and Environmental Microbiology*, *61*(4), 1646–1648.

Insam, H., & Seewald, M. S. A. (2010). Volatile organic compounds (VOCs) in soils. *Biology and Fertility of Soil*, *46*, 199–213. https://doi.org/10.1007/s00374-010-0442-3

Jáchymová, J., Votruba, J., Víden, I., & Řezanka, T. (2002). Identification of Streptomyces odor spectrum. *Folia Microbiologica*, *47*(1), 37–41. https://doi.org/10.1007/BF02818563

Jain, P. K., & Jain, P. C. (2007). Isolation, characterization and antifungal activity of Streptomyces sampsonii GS 1322. *Indian Journal of Experimental Biology*, *45*(2), 203–206.

Jain, R., Jain, A., Rawat, N., Nair, M., & Gumashta, R. (2016). Feather hydrolysate from Streptomyces sampsonii GS 1322: A potential low cost soil amendment. *Journal of Bioscience and Bioengineering*, *121*(6), 672–677. https://doi.org/10.1016/j.jbiosc.2015.11.003

Jensen, S. E., Westlake, D. W. S., & Wolfe, S. (1983). Partial purification and characterization of isopenicillin N epimerase activity from Streptomyces clavuligerus. *Canadian Journal of Microbiology*, *29*(11), 1526–1531. https://doi.org/10.1139/m83-234

Jia, Q., Chen, X., Köllner, T. G., Rinkel, J., Fu, J., Labbé, J., Xiong, W., Dickschat, J. S., Gershenzon, J., & Chen, F. (2019). terpene synthase Genes originated from Bacteria through Horizontal Gene transfer Contribute to terpenoid Diversity in Fungi. *Scientific Reports*, *9*(January), 1–11. https://doi.org/10.1038/s41598-019-45532-1

Jiang, J., He, X., & Cane, D. E. (2007). Biosynthesis of the Earthy Odorant Geosmin by a Bifunctional Streptomyces Coelicolor Enzyme. *Nature Chemical Biology*, *3*(11), 711–715. https://doi.org/10.1038/jid.2014.371

Jirovetz, L., Buchbauer, G., Ngassoum, M. B., & Geissler, M. (2002). A roma compound analysis of Piper nigrum and Piper guineense essential oils from Cameroon using solid-phase microextraction–gas chromatography, solid-phase microextraction–gas chromatography– mass spectrometry and olfactometry. *Journal of Chromatography A*, *976*(1–2), 265–275. https://doi.org/10.1016/S0021-9673(02)00376-X

Johnson, E. G., Joshi, M. V., Gibson, D. M., & Loria, R. (2007). Cello-oligosaccharides released from host plants induce pathogenicity in scab-causing Streptomyces species. *Physiological and Molecular Plant Pathology*, *71*(1–3), 18–25. https://doi.org/10.1016/j.pmpp.2007.09.003

Johnson, J. S., Spakowicz, D. J., Hong, B., Petersen, L. M., Demkowicz, P., Chen, L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., Sodergren, E., & Weinstock, G. M. (2019). Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications*, *10*(5029), 1–11. https://doi.org/10.1038/s41467-019-13036-1

Jolley, K. A., Bray, J. E., & Maiden, M. C. J. (2018). Open-access bacterial population

genomics: BIGSdb software, the PubMLST.org website and their applications [version 1; peer review: 2 approved]. *Wellcome Open Research*, *3*(124), 1–20. https://doi.org/10.12688/wellcomeopenres.14826.1

Jolley, K. A., & Maiden, M. C. J. (2014). Using MLST to study bacterial variation: Prospects in the genomic era. *Future Microbiology*, *9*(5), 623–630. https://doi.org/10.2217/fmb.14.24

Jones, S. E., Ho, L., Rees, C. A., Hill, J. E., Nodwell, J. R., & Elliot, M. A. (2017). Streptomyces exploration is triggered by fungal interactions and volatile signals. *ELIFE*, *6*, 1–21. https://doi.org/10.7554/eLife.21738

Jones, S. E., Pham, C. A., Zambri, M. P., McKillip, J., Carlson, E. E., & Elliot, M. A. (2019). Streptomyces Volatile Compounds Influence Exploration and Microbial Community Dynamics by Altering Iron Availability. *MBio*, *10*(2), 1–18. https://doi.org/10.1128/mBio.00171-19

Joo, G. J. (2005). Production of an anti-fungal substance for biological control of Phytophthora capsici causing phytophthora blight in red-peppers by Streptomyces halstedii. *Biotechnology Letters*, *27*(3), 201–205. https://doi.org/10.1007/s10529-004-7879-0

Joycharat, N., Thammavong, S., Voravuthikunchai, S. P., Plodpai, P., Mitsuwan, W., Limsuwan, S., & Subhadhirasakul, S. (2014). Chemical constituents and antimicrobial properties of the essential oil and ethanol extract from the stem of Aglaia odorata Lour. *Natural Product Research*, *28*(23), 2169–2172. https://doi.org/10.1080/14786419.2014.924934

Kanchiswamy, C. N., Malnoy, M., & Maffei, M. E. (2015). Bioprospecting bacterial and fungal volatiles for sustainable agriculture. *Trends in Plant Science*, *20*(4), 206–211. https://doi.org/10.1016/j.tplants.2015.01.004

Kataoka, M., Ueda, K., Kudo, T., & Seki, T. (1997). Application of the variable region in 16S rDNA to create an index for rapid species identification in the genus Streptomyces. *FEMS Microbiology Letters*, *151*, 249–255.

Katz, L., & Baltz, R. H. (2016). Natural product discovery: past, present, and future. *Journal of Industrial Microbiology and Biotechnology*, *43*(2–3), 155–176. https://doi.org/10.1007/s10295-015-1723-5

Kelly, S. L., & Kelly, D. E. (2013). Microbial cytochromes P450: biodiversity and biotechnology. Where do cytochromes P450 come from, what do they do and what can they do for us? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1612). https://doi.org/10.1098/rstb.2012.0476

Khamis, A., Raoult, D., & Scola, B. La. (2005). Comparison between rpoB and 16S rRNA Gene Sequencing for Molecular Identification of 168 Clinical Isolates of Corynebacterium. *Journal of Clinical Microbiology*, *43*(4), 1934–1936. https://doi.org/10.1128/JCM.43.4.1934

Khoja, S., Eltayef, K. M., Baxter, I., Myrta, A., Bull, J. C., & Butt, T. (2020). Volatiles of the entomopathogenic fungus, Metarhizium brunneum, attract and kill plant parasitic nematodes. *Biological Control*, *152*(September 2020), 104472. https://doi.org/10.1016/j.biocontrol.2020.104472

Kim, B. J., Kim, C. J., Chun, J., Koh, Y. H., Lee, S. H., Hyun, J. W., Cha, C. Y., & Kook, Y. H. (2004). Phylogenetic analysis of the genera Streptomyces and Kitasatospora based on partial RNA polymerase β-subunit gene (rpoB) sequences. *International Journal of Systematic and Evolutionary Microbiology*, *54*(2), 593–598. https://doi.org/10.1099/ijs.0.02941-0

Kim, D. R., Cho, G., Jeon, C. W., Weller, D. M., Thomashow, L. S., Paulitz, T. C., & Kwak, Y. S. (2019). A mutualistic interaction between Streptomyces bacteria, strawberry plants and pollinating bees. *Nature Communications*, *10*(4802). https://doi.org/10.1038/s41467-019-12785-3

Kirkpatrick, D. M., Leach, H. L., Xu, P., Dong, K., Isaacs, R., & Gut, L. J. (2018). Comparative Antennal and Behavioral Responses of Summer and Winter Morph Drosophila suzukii (Diptera: Drosophilidae) to Ecologically Relevant Volatiles. *Environmental Entomology*, *47*(3), 700–706. https://doi.org/10.1093/ee/nvy046

Klocke, J. A., Darlington, M. V., & Balandrin, M. F. (1987). 1,8-Cineole (Eucalyptol), a Mosquito Feeding and Ovipositional Repellent From Volatile Oil ofHemizonia Fitchii (Asteraceae). *Journal of Chemical Ecology*, *13*(12), 2131–2141. https://doi.org/10.1007/BF01012562

Kolter, R., Siegele, D. A., & Tormo, A. (1993). The stationary phase of the bacterial life cycle. *Annual Review of Microbiology*, *47*, 855–874. https://doi.org/10.1146/annurev.mi.47.100193.004231

Komaki, H. (2021). Reclassification of Streptomyces costaricanus and Streptomyces phaeogriseichromatogenes as later heterotypic synonyms of Streptomyces murinus. *International Journal of Systematic and Evolutionary Microbiology*, *71*(2), 1–6. https://doi.org/10.1099/ijsem.0.004638

Komatsu, M., Tsuda, M., Omura, S., Oikawa, H., & Ikeda, H. (2008). Identification and functional analysis of genes controlling biosynthesis of 2-methylisoborneol. *PNAS*, *105*(21), 7422–7427. https://doi.org/10.1073/pnas.0802312105

Korpi, A., Jarnberg, J., & Pasanen, A.-L. (2009). Microbial volatile organic compounds. *Critical Reviews in Toxicology*, *39*, 139–193. https://doi.org/10.1080/10408440802291497

Kroiss, J., Kaltenpoth, M., Schneider, B., Schwinger, M. G., Hertweck, C., Maddula, R. K., Strohm, E., & Svatos, A. (2010). Symbiotic streptomycetes provide antibiotic combination prophylaxis for wasp offspring. *Nature Chemical Biology*, *6*(4), 261–263. https://doi.org/10.1038/nchembio.331

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular

evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Lehtinen, T., Efimova, E., Santala, S., & Santala, V. (2018). Improved fatty aldehyde and wax ester production by overexpression of fatty acyl-CoA reductases. *Microbial Cell Factories*, *17*(1), 1–10. https://doi.org/10.1186/s12934-018-0869-z

Li, R., Chou, W. K. W., Himmelberger, J. A., Litwin, K. M., Harris, G. G., Cane, D. E., & Christianson, D. W. (2014). Reprogramming the Chemodiversity of Terpenoid Cyclization by Remolding the Active Site Contour of epi-Isozizaene Synthase. *Biochemistry*, *53*, 1155–1168.

Li, X., Li, B., Cai, S., Zhang, Y., Xu, M., Zhang, C., Yuan, B., Xing, K., & Qin, S. (2020). Identification of rhizospheric actinomycete streptomyces lavendulae sps-33 and the inhibitory effect of its volatile organic compounds against ceratocystis fimbriata in postharvest sweet potato (Ipomoea batatas (L.) Lam.). *Microorganisms*, *8*(3). https://doi.org/10.3390/microorganisms8030319

Li, Y., Liu, J., Adekunle, D., Bown, L., Tahlan, K., & Bignell, D. R. D. (2019). TxtH is a key component of the thaxtomin biosynthetic machinery in the potato common scab pathogen Streptomyces scabies. *Molecular Plant Pathology*, *20*(10), 1379–1393. https://doi.org/10.1111/mpp.12843

Lian, Q., Zhang, J., Gan, L., Ma, Q., Zong, Z., & Wang, Y. (2017). The biocontrol efficacy of streptomyces pratensis LMM15 on botrytis cinerea in Tomato. *BioMed Research International*, *2017*. https://doi.org/10.1155/2017/9486794

Light, M., Shutler, D., Cutler, G. C., & Hillier, N. K. (2020). Varroa destructor mite electrophysiological responses to honey bee (Apis mellifera) colony volatiles. *Experimental and Applied Acarology*, *81*(4). https://doi.org/10.1007/s10493-020-00519-w

Lin, X., & Cane, D. E. (2009). Biosynthesis of the Sesquiterpene Antibiotic Albaflavenone in Streptomyces Coelicolor. Mechanism and Stereochemistry of the Enzymatic Formation of Epi-Isozizaene. *Journal of American Chemical Society*, *131*(18), 6332–6333. https://doi.org/10.1021/ja901313v.Biosynthesis

Links, M. G., Dumonceaux, T. J., Hemmingsen, S. M., & Hill, J. E. (2012). The Chaperonin-60 Universal Target Is a Barcode for Bacteria That Enables De Novo Assembly of Metagenomic Sequence Data. *PLoS ONE*, *7*(11). https://doi.org/10.1371/journal.pone.0049755

Linstrom, P. J., & Mallard, W. G. (2001). The NIST Chemistry WebBook: A chemical data resource on the Internet. *Journal of Chemical and Engineering Data*, *46*(5), 1059–1063. https://doi.org/10.1021/je000236i

Liu, K., Chen, Q., Liu, Y., Zhou, X., & Wang, X. (2012). Isolation and Biological Activities of Decanal, Linalool, Valencene, and Octanal From Sweet Orange Oil. *Journal of Food Science*, *77*(11). https://doi.org/10.1111/j.1750-3841.2012.02924.x

Liu, M., Jia, Y., Xie, Y., Zhang, C., Ma, J., Sun, C., & Ju, J. (2019). Identification of the actinomycin D biosynthetic pathway from marine-derived Streptomyces costaricanus SCSIO ZS0073. *Marine Drugs*, *17*(4), 1–13. https://doi.org/10.3390/md17040240

Lofstedt, C., & Bengtsson, M. (1988). Sex Pheromone Biosynthesis of (E,E)-8,10-dodecadienol in Codling mothCydia Pomonella involves E9 Desaturation. *Journal of Chemical Ecology*, *14*(3), 903–915.

Logrieco, A., Mulè, G., Moretti, A., & Bottalico, A. (2002). Toxigenic Fusarium species and mycotoxins associated with maize ear rot in Europe. *European Journal of Plant Pathology*, *108*(7), 597–609. https://doi.org/10.1023/A:1020679029993

Lücker, J., Bowen, P., & Bohlmann, J. (2004). Vitis vinifera terpenoid cyclases: functional identification of two sesquiterpene synthase cDNAs encoding (+)-valencene synthase and (?)-germacrene D synthase and expression of mono- and sesquiterpene synthases in grapevine flowers and berries. *Phytochemistry*, *65*(19), 2649–2659. https://doi.org/10.1016/j.phytochem.2004.08.017

Lyu, A., Yang, L., Wu, M., Zhang, J., & Li, G. (2020). High Efficacy of the Volatile Organic Compounds of Streptomyces yanglinensis 3-10 in Suppression of Aspergillus Contamination on Peanut Kernels. *Frontiers in Microbiology*, *11*, 1–17. https://doi.org/10.3389/fmicb.2020.00142

Madhyasth, M. S., & Bhat, R. V. (1984). Aspergillus parasiticus Growth and Aflatoxin Production on Black and White Pepper and the Inhibitory Action of Their Chemical Constituents. *Applied and Environmental Microbiology*, *48*(2), 376–379.

Maiden, M. C. J., Bygraves, J. A., Feil, E., Morelli, G., Russell, J., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M., & Spratt, B. G. (1998). Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Pnas*, *95*, 3140–3145.

Manter, D. K., Kelsey, R. G., & Karchesy, J. J. (2007). Antimicrobial Activity of Extractable Conifer Heartwood Compounds Toward Phytophthora ramorum. *Journal of Chemical Ecology*, *33*, 2133–2147. https://doi.org/10.1007/s10886-007-9368-0

Marchesi, J. R., Sato, T., Weightman, A. J., Martin, T. A., Fry, J. C., Hiom, S. J., & Wade, W. G. (1998). Design and Evaluation of Useful Bacterium-Specific PCR Primers That Amplify Genes Coding for Bacterial 16S rRNA. *Applied and Environmental Microbiology*, *64*(2), 795–799.

Martens, T., Brinkhoff, T., Simon, M., & Schulz, S. (2005). *Volatiles Released by a Streptomyces Species Isolated from the North Sea*. *2*.

Martín-Sánchez, L., Singh, K. S., Avalos, M., Van Wezel, G. P., Dickschat, J. S., & Garbeva, P. (2019). Phylogenomic analyses and distribution of terpene synthases among Streptomyces. *Beilstein Journal of Organic Chemistry*, *15*, 1181–1193.

https://doi.org/10.3762/bjoc.15.115

Mcmahon, C., Guerin, P. M., & Syed, Z. (2001). 1-Octen-3-ol isolated from bont ticks attracts Amblyomma variegatum. *Journal of Chemical Ecology*, *27*(3), 471–486. https://doi.org/10.1023/A:1010328720035

Mehling, A., Wehmeier, U. F., & Piepersberg, W. (1995). Nucleotide sequences of streptomycete 16S ribosomal DNA: Towards a specific identification system for streptomycetes using PCR. *Microbiology*, *141*(9), 2139–2147. https://doi.org/10.1099/13500872-141-9-2139

Meijerink, J., Braks, M. A. H., & Van Loon, J. J. . (2001). Olfactory receptors on the antennae of the malaria mosquito Anopheles gambiae are sensitive to ammonia and other sweat-borne components. *Journal of Insect Physiology*, *47*, 455–464.

Mekadim, C., Killer, J., Pechar, R., & Mrazek, J. (2018). Variable regions of the glyS, infB and rplB genes usable as novel genetic markers for identification and phylogenetic purposes of genera belonging to the family Propionibacteriaceae. *International Journal of Systematic and Evolutionary Microbiology*, *68*, 2697–2705. https://doi.org/10.1099/ijsem.0.002873

Minaeimoghadam, M., Askarianzadeh, A., Imani, S., Shojaei, M., Larijani, K., & Abbasipour, H. (2017). Identification of chemical compounds of the pheromone in different ages of female adults of the clearwing moth, Paranthrene diaphana Dalla Torre & Strand. *Archives of Phytopathology and Plant Protection*, *50*, 1019–1033. https://doi.org/10.1080/03235408.2017.1411174

Miñana-Galbis, D., Urbizu-Serrano, A., Farfán, M., Fusté, M. C., & Lorén, J. G. (2009). Phylogenetic analysis and identification of Aeromonas species based on sequencing of the cpn60 universal target. *International Journal of Systematic and Evolutionary Microbiology*, *59*(8), 1976–1983. https://doi.org/10.1099/ijs.0.005413-0

Monnin, T., Helft, F., Leroy, C., d'Ettorre, P., & Doums, C. (2018). Chemical Characterization of Young Virgin Queens and Mated Egg-Laying Queens in the Ant Cataglyphis cursor: Random Forest Classification Analysis for Multivariate Datasets. *Journal of Chemical Ecology*, *44*(2), 127–136. https://doi.org/10.1007/s10886-018-0923-7

Morrison III, W. R., Larson, N. L., Brabec, D., & Zhang, A. (2019). Methyl Benzoate as a Putative Alternative, Environmentally Friendly Fumigant for the Control of Stored Product Insects. *Journal of Economic Entomology*, *112*(5), 2458–2468. https://doi.org/10.1093/jee/toz179

Mostafiz, M. M., Jhan, P. K., Shim, J., & Lee, K.-Y. (2018). Methyl benzoate exhibits insecticidal and repellent activities against Bemisia tabaci (Gennadius) (Hemiptera: Aleyrodidae). *PLoS ONE*, 1–14.

Mowat, E., Rajendran, R., Williams, C., Mcculloch, E., Jones, B., Lang, S., & Ramage, G. (2010). Pseudomonasaeruginosaand their small di¡usible extracellular molecules

inhibitAspergillusfumigatus bio¢lmformation. *FEMS Microbiology Letters*, *313*, 96–102. https://doi.org/10.1111/j.1574-6968.2010.02130.x

Mukherjee, S., Stamatis, D., Bertsch, J., Ovchinnikova, G., Sundaramurthi, J. C., Lee, J., Kandimalla, M., Chen, I. M. A., Kyrpides, N. C., & Reddy, T. B. K. (2021). Genomes OnLine Database (GOLD) v.8: Overview and updates. *Nucleic Acids Research*, *49*(D1), D723–D733. https://doi.org/10.1093/nar/gkaa983

Mulyaningsih, S., Youns, M., El-Readi, M. Z., Ashour, M. L., Nibret, E., Sporer, F., Herrmann, F., Reichling, J., & Wink, M. (2010). Biological activity of the essential oil of Kadsura longipedunculata (Schisandraceae) and its major components. *Journal of Pharmacy and Pharmacology*, *62*(8), 1037–1044. https://doi.org/10.1111/j.2042-7158.2010.01119.x

Munita, J. M., & Arias, C. A. (2016). Mechanisms of antibiotic resistance. *American Society for Microbiology*, *6*, 1–24. https://doi.org/10.3389/fmicb.2015.00034

Na, S., Kim, Y. O., Yoon, S., Ha, S., Baek, I., & Chun, J. (2018). UBCG : Up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. *Journal of Microbiology*, *56*(4), 280–285. https://doi.org/10.1007/s12275-018-8014-6

Nagashima, K., Hisada, T., Sato, M., & Mochizuki, J. (2003). Application of new primer-enzyme combinations to terminal restriction fragment length polymorphism profiling of bacterial populations in human feces. *Applied and Environmental Microbiology*, *69*(2), 1251–1262. https://doi.org/10.1128/AEM.69.2.1251-1262.2003

Narayana, K. J. P., Prabhakar, P., Vijayalakshmi, M., Venkateswarlu, Y., & Krishna, P. S. J. (2008). Study on bioactive compounds from Streptomyces sp. ANU 6277. *Polish Journal of Microbiology*, *57*(1), 35–39.

Nei, M., & Kumar, S. (2000). Molecular Evolution and Phylogenetics. In *Genetical Research* (Issue 1). Oxford University Press. https://doi.org/10.1017/s0016672301219405

Ni, J., Robarge, W. P., Xiao, C., & Heber, A. J. (2012). Volatile organic compounds at swine facilities: A critical review. *Chemosphere*, *89*(7), 769–788. https://doi.org/10.1016/j.chemosphere.2012.04.061

Nie, Y., Chi, C., Fang, H., Liang, J., Lu, S., Lai, G., Tang, Y., & Wu, X. (2014). Diverse alkane hydroxylase genes in microorganisms and environments. *Scientific Reports*, *4*(4968), 1–11. https://doi.org/10.1038/srep04968

Ogunbinua, A. O., Flamini, G., Cioni, P. L., Adebayo, M. A., & Ogunwande, I. A. (2009). Constituents of Cajanus cajan (L.) Millsp., Moringa oleifera Lam., Heliotropium indicum L. and Bidens pilosa L. from Nigeria. *Natural Product Communications*, *4*(4), 573–578.

Olaide, O. Y., Tchouassi, D. P., Yusuf, A. A., Pirk, C. W. W., Masiga, D. K., Saini, R. K., & Torto, B. (2019). Zebra Skin Odor Repels the Savannah Tsetse Fly, Glossina

Pallidipes (Diptera: Glossinidae). *PLoS ONE*, 1–18.

Pan, R., Bai, X., Chen, J., Zhang, H., & Wang, H. (2019). Exploring structural diversity of microbe secondary metabolites using OSMAC strategy: A literature review. *Frontiers in Microbiology*, *10*(294), 1–20. https://doi.org/10.3389/fmicb.2019.00294

Paoli, M., Nam, A. M., Castola, V., Casanova, J., & Bighelli, A. (2011). Chemical variability of the wood essential oil of cedrus atlantica manetti from Corsica. *Chemistry and Biodiversity*, *8*(2), 344–351. https://doi.org/10.1002/cbdv.201000094

Parte, A. C., Carbasse, J. S., Meier-Kolthoff, J. P., Reimer, L. C., & Göker, M. (2020). List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. *International Journal of Systematic and Evolutionary Microbiology*, *70*(11), 5607–5612. https://doi.org/10.1099/ijsem.0.004332

Peterson, E., & Kaur, P. (2018). Antibiotic resistance mechanisms in bacteria: Relationships between resistance determinants of antibiotic producers, environmental bacteria, and clinical pathogens. *Frontiers in Microbiology*, *9*(2928), 1–21. https://doi.org/10.3389/fmicb.2018.02928

Pierce, A. M., Pierce Jr., H. D., Oehlschlager, A. C., & Borden, J. H. (1990). 1-Octen-3-ol, attractive semiochemical for foreign grain beetle,Ahasverus advena (Waltl) (Coleoptera: Cucujidae). *Journal of Chemical Ecology*, *17*(3), 567–580.

Pinto-zevallos, D. M., Bezerra, R. H. S., Souza, S. R., & Ambrogi, B. G. (2018). Species- And Density-Dependent Induction of Volatile Organic Compounds by Three Mite Species in Cassava and Their Role in the Attraction of a Natural Enemy. *Experimental and Applied Acarology*, *74*(3), 261–274. https://doi.org/10.1007/s10493-018-0231-5

Pollak, F. C., & Berger, R. G. (1996). Geosmin and related volatiles in bioreactor-cultured Streptomyces citreus CBS 109.60. *Applied and Environmental Microbiology*, *62*(4), 1295–1299. https://doi.org/10.1128/aem.62.4.1295-1299.1996

Pollock, J., Laura, G., Trong, W., & Mick, W. (2018). The Madness of Microbiome : Attempting To Find Consensus. *Applied and Environmental Microbiology*, *84*(7), 1–12.

Pope, T. W., Campbell, C. A. M., Hardie, J., & Wadhams, L. J. (2004). Electroantennogram responses of the three migratory forms of the damson-hopaphid, Phorodon humuli, to aphid pheromones and plant volatiles. *Journal of Insect Physiology*, *50*(11), 1083–1092. https://doi.org/10.1016/j.jinsphys.2004.09.010

Posy, D. C., Mohamed, M. A., Coppel, H. C., & Jeanne, R. L. (1984). Identification of ant repellent allomone produced by social wasp Polistes fuscatus (Hymenoptera: Vespidae). *Journal of Chemical Ecology*, *10*(12), 1799–1807. https://doi.org/10.1007/BF00987362

Qin, Y., Polacek, N., Vesper, O., Staub, E., Einfeldt, E., Wilson, D. N., & Nierhaus, K. H.

(2006). The Highly Conserved LepA Is a Ribosomal Elongation Factor that Back-Translocates the Ribosome. *Cell*, *127*, 721–733. https://doi.org/10.1016/j.cell.2006.09.037

Quaderer, R., Omaru, S., Ikeda, H., & Cane, D. E. (2006). Pentalenolactone Biosynthesis. Molecular Cloning and Assignment of Biochemical Function to PtlI, a Cytochrome P450 of Streptomyces Avermitilis. *Journal of American Chemical Society*, *128*(40), 13036–13037. https://doi.org/10.1021/ja0639214.Pentalenolactone

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), 590–596. https://doi.org/10.1093/nar/gks1219

Rabe, P., Citron, C. A., & Dickschat, J. S. (2013). Volatile Terpenes from Actinomycetes: A Biosynthetic Study Correlating Chemical Analyses to Genome Data. *Chembiochem*, *14*, 2345–2354. https://doi.org/10.1002/cbic.201300329

Rasko, D. A., Altherr, M. R., Han, C. S., & Ravel, J. (2005). Genomics of the Bacillus cereus group of organisms. *FEMS Microbiology Reviews*, *29*, 303–329. https://doi.org/10.1016/j.femsre.2004.12.005

Raza, W., Ling, N., Liu, D., Wei, Z., Huang, Q., & Shen, Q. (2016). Volatile organic compounds produced by Pseudomonas fluorescens WR-1 restrict the growth and virulence traits of Ralstonia solanacearum. *Microbiological Research*, *192*, 103–113. https://doi.org/10.1016/j.micres.2016.05.014

Reller, L. B., Weinstein, M. P., & Petti, C. A. (2007). Detection and Identification of Microorganisms by Gene Amplification and Sequencing. *Clinical Infectious Diseases*, *44*(8), 1108–1114. https://doi.org/10.1086/512818

Reysenbach, A., Wickham, G. S., & Pace, N. R. (1994). Phylogenetic Analysis of the Hyperthermophilic Pink Filament Community in Octopus Spring , Yellowstone National Park. *Applied and Environmental Microbiology*, *60*(6), 2113–2119.

Ribeiro, T. G., Novais, A., Branquinho, R., Machado, E., & Peixe, L. (2015). Phylogeny and Comparative Genomics Unveil Independent Diversification Trajectories of qnrB and Genetic Platforms within Particular Citrobacter Species. *American Society for Microbiology*, *59*(10), 5951–5958. https://doi.org/10.1128/AAC.00027-15

Riclea, R., Aigle, B., Leblond, P., Schoenian, I., Spiteller, D., & Dickschat, J. S. (2012). Volatile Lactones from Streptomycetes Arise via the Antimycin Biosynthetic Pathway. *ChemBioChem*, *13*, 1635–1645. https://doi.org/10.1002/cbic.201200260

Rinkel, J., Rabe, P., Zur Horst, L., & Dickschat, J. S. (2016). A detailed view on 1,8-cineol biosynthesis by Streptomyces clavuligerus. *Beilstein Journal of Organic Chemistry*, *12*, 2317–2324. https://doi.org/10.3762/bjoc.12.225

Rogall, T., Wolters, J., Flohr, T., & Bottger, E. C. (1990). Towards a Phylogeny and

Definition of Species at the Molecular Level within the Genus Mycobacterium. *International Journal of Systematic Bacteriology*, *40*(4), 323–330.

Santos, S. R., & Ochman, H. (2004). Identification and phylogenetic sorting of bacterial lineages with universally conserved genes and proteins. *Environmental Microbiology*, *6*(7), 754–759. https://doi.org/10.1111/j.1462-2920.2004.00617.x

Satyal, P., Hieu, H. V., Chuong, N. T. H., Hung, N. H., Sinh, L. H., Van The, P., Tai, T. A., Hien, V. T., & Setzer, W. N. (2019). Chemical composition, Aedes mosquito larvicidal activity, and repellent activity against Triatoma rubrofasciata of Severinia monophylla leaf essential oil. *Parasitology Research*, *118*(3), 733–742. https://doi.org/10.1007/s00436-019-06212-1

Scheller, U., Zimmer, T., Becher, D., Schauer, F., & Schunck, W. (1998). Oxygenation Cascade in Conversion of N-Alkanes to Alpha,omega-Dioic Acids Catalyzed by Cytochrome P450 52A3. *The Journal of Biological Chemistry*, *273*(49), 32528–32534.

Schirmer, A., Rude, M. A., Li, X., Popova, E., & del Cardayre, S. B. (2010). Microbial Biosynthesis of Alkanes. *Science*, *329*.

Schmidt, R., Cordovez, V., De Boer, W., Raaijmakers, J., & Garbeva, P. (2015). Volatile affairs in microbial interactions. *ISME Journal*, *9*(11), 2329–2335. https://doi.org/10.1038/ismej.2015.42

Schnurer, J., Olsson, J., & Borjesson, T. (1999). Fungal Volatiles as Indicators of Food and Feeds Spoilage. *Fungal Genetics and Biology*, *27*(2–3), 209–217.

Schöller, C. E. G., Gürtler, H., Pedersen, R., Molin, S., & Wilkins, K. (2002). Volatile metabolites from actinomycetes. *Journal of Agricultural and Food Chemistry*, *50*(9), 2615–2621. https://doi.org/10.1021/jf0116754

Schrader, K. K., Harries, M. D., & Page, P. N. (2015). Temperature effects on biomass, geosmin, and 2-methylisoborneol production and cellular activity by Nocardia spp. and Streptomyces spp. isolated from rainbow trout recirculating aquaculture systems. *Journal of Industrial Microbiology and Biotechnology*, *42*(5), 759–767. https://doi.org/10.1007/s10295-015-1600-2

Schulz-bohm, K., Geisen, S., Wubs, E. R. J., Song, C., Boer, W. De, & Garbeva, P. (2017). The prey's scent - volatile organic compound mediated interactions between soil bacteria and their protist predators. *ISME Journal*, *11*, 817–820. https://doi.org/10.1038/ismej.2016.144

Schulz, S., & Dickschat, J. S. (2007). Bacterial volatiles: The smell of small organisms. *Natural Product Reports*, *24*(4), 814–842. https://doi.org/10.1039/b507392h

Shi, W., Li, M., Wei, G., Tian, R., Li, C., Wang, B., Lin, R., Shi, C., Chi, X., Zhou, B., & Gao, Z. (2019). The occurrence of potato common scab correlates with the community composition and function of the geocaulosphere soil microbiome.

Microbiome, *7*(1), 1–18. https://doi.org/10.1186/s40168-019-0629-2

Shirasu, M., Ito, S., Itoigawa, A., Hayakawa, T., Kinoshita, K., Munechika, I., Imai, H., & Touhara, K. (2020). Key Male Glandular Odorants Attracting Female Ring-Tailed Lemurs. *Current Biology*, *30*, 1–8. https://doi.org/10.1016/j.cub.2020.03.037

Shirling, E. B., & Gottlieb, D. (1969). Cooperative description of type cultures of Streptomyces. IV. Species descriptions from the second, third and fourth studies. *International Journal of Systematic Bacteriology*, *19*(4), 391–512. https://doi.org/10.1099/00207713-19-4-391

Shirling, E. B., & Gottlieb, D. (1972). Cooperative Description of Type Strains of Streptomyces V. Additional Descriptions. *International Journal of Systematic Bacteriology*, *22*(4), 265–394. https://doi.org/10.1099/00207713-22-4-265

Siebenberg, S., Bapat, P. M., Lantz, A. E., Gust, B., & Heide, L. (2010). Reducing the variability of antibiotic production in Streptomyces by cultivation in 24-square deepwell plates. *Journal of Bioscience and Bioengineering*, *109*(3), 230–234. https://doi.org/10.1016/j.jbiosc.2009.08.479

Song, J., Lee, S. C., Kang, J. W., Baek, H. J., & Suh, J. W. (2004). Phylogenetic analysis of Streptomyces spp. isolated from potato scab lesions in Korea on the basis of 16S rRNA gene and 16S-23S rDNA internally transcribed spacer sequences. *International Journal of Systematic and Evolutionary Microbiology*, *54*(1), 203–209. https://doi.org/10.1099/ijs.0.02624-0

Soso, S. B., & Koziel, J. A. (2016). Analysis of Odorants in Marking Fluid of Siberian Tiger (Panthera tigris altaica) Using Simultaneous Sensory and Chemical Analysis with Headspace Solid-Phase Microextraction and Multidimensional Gas Chromatography-Mass Spectrometry-Olfactometry. *Molecules*, *21*(7), 1–22. https://doi.org/10.3390/molecules21070834

Stackebrandt, E., Koch, C., Gvozdiak, O., & Schumann, P. (1995). Taxonomic dissection of the genus Micrococcus: Kocuria gen. nov., Nesterenkonia gen. nov., Kytococcus gen. nov., Dermacoccus gen. nov., and Micrococcus cohn 1872 gen. emend. *International Journal of Systematic Bacteriology*, *45*(4), 682–692. https://doi.org/10.1099/00207713-45-4-682

Stensmyr, M. C., Dweck, H. K. M., Farhan, A., Ibba, I., Strutz, A., Mukunda, L., Linz, J., Grabe, V., Steck, K., Lavista-Llanos, S., Wicher, D., Sachse, S., Knaden, M., Becher, P. G., Seki, Y., & Hansson, B. S. (2012). A conserved dedicated olfactory circuit for detecting harmful microbes in drosophila. *Cell*, *151*(6), 1345–1357. https://doi.org/10.1016/j.cell.2012.09.046

Takamatsu, S., Lin, X., Nara, A., Komatsu, M., Cane, D. E., & Ikeda, H. (2011). Characterization of a silent sesquiterpenoid biosynthetic pathway in Streptomyces avermitilis controlling epi-isozizaene albaflavenone biosynthesis and isolation of a new oxidized epi-isozizaene metabolite. *Microbial Biotechnology*, *4*(2), 184–191.

https://doi.org/10.1111/j.1751-7915.2010.00209.x

Takamatsu, S., Xu, L., Fushinobu, S., Shoun, H., Komatsu, M., Cane, D. E., & Ikeda, H. (2011). Pentalenic acid is a shunt metabolite in the biosynthesis of the pentalenolactone family of metabolites: hydroxylation of 1-deoxypentalenic acid mediated by CYP105D7 (SAV_7469) of Streptomyces avermitilis. *The Journal of Antibiotics*, *64*, 65–71. https://doi.org/10.1038/ja.2010.135

Tang, Y., Ellis, N. M., Hopkins, M. K., Smith, D. H., Dodge, D. E., & Persing, D. H. (1998). Comparison of Phenotypic and Genotypic Techniques for Identification of Unusual Aerobic Pathogenic Gram-Negative Bacilli. *Journal of Clinical Microbiology*, *36*(12), 3674–3679.

Teske, A., Wawer, C., Muyzer, G., & Ramsing, N. B. (1996). Distribution of Sulfate-Reducing Bacteria in a Stratified Fjord (Mariager Fjord, Denmark) as Evaluated by Most-ProbableNumber Counts and Denaturing Gradient Gel Electrophoresis of PCR-Amplified Ribosomal DNA Fragments. *Applied and Environmental Microbiology*, *62*(4), 1405–1415.

Thangadurai, D., Anitha, S., Pullaiah, T., Reddy, P. N., & Ramachandraiah, O. S. (2002). Essential oil constituents and in vitro antimicrobial activity of Decalepis hamiltonii roots against foodborne pathogens. *Journal of Agricultural and Food Chemistry*, *50*(11), 3147–3149. https://doi.org/10.1021/jf011541q

Thomas, C. M., & Nielsen, K. M. (2005). Mechanisms of, and Barriers to, Horizontal Gene Transfer between Bacteria. *Nature Reviews Microbiology*, *3*, 711–722. https://doi.org/10.1038/nrmicro1234

Thompson, F. L., Gevers, D., Thompson, C. C., Dawyndt, P., Naser, S., Hoste, B., Munn, C. B., & Swings, J. (2005). Phylogeny and Molecular Identification of Vibrios on the Basis of Multilocus Sequence Analysis. *Applied and Environmental Microbiology*, *71*(9), 5107–5115. https://doi.org/10.1128/AEM.71.9.5107

Tian, X., Zhang, Z., Yang, T., Chen, M., Li, J., Chen, F., Yang, J., Li, W., Zhang, B., Zhang, Z., Wu, J., Zhang, C., Long, L., & Xiao, J. (2016). Comparative Genomics Analysis of Streptomyces Species Reveals Their Adaptation to the Marine Environment and Their Diversity at the Genomic Level. *Frontiers in Microbiology*, *7*, 1–16. https://doi.org/10.3389/fmicb.2016.00998

Toloza, A. C., Zygadlo, J., Cueto, G. M., Biurrun, F., Zerba, E., & Picollo, M. I. (2006). Fumigant and Repellent Properties of Essential Oils and Component Compounds Against Permethrin-Resistant Pediculus humanus capitis (Anoplura: Pediculidae) from Argentina. *Journal of Medical Entomology*, *43*(5), 889–895.

Torres-romero, D., Munoz-Martınez, F., Jimenez, I. A., Castanys, S., Gamarro, F., & Bazzocchi, I. L. (2009). Novel dihydro-beta-agarofuran sesquiterpenes as potent modulators of human P-glycoprotein dependent multidrug resistance. *Organic and Biomolecular Chemistry*, *7*, 5166–5172. https://doi.org/10.1039/b915678j

Tsusaka, T., Makino, B., Ohsawa, R., & Ezura, H. (2020). Evaluation of heritability of β-eudesmol/ hinesol content ratio in Atractylodes lancea De Candolle. *BMC Hereditas*, *157*(1), 1–8. https://doi.org/10.1186/s41065-020-00123-3

Turner, S. L., Li, N., Guda, T., Githure, J., Carde, R. T., & Ray, A. (2011). Ultra-prolonged activation of CO2-sensing neurons disorients mosquitoes. *Nature*, *474*, 87–92. https://doi.org/10.1038/nature10081

Vale, G. A., Hall, D. R., Chamisa, A., & Torr, S. J. (2012). Towards an Early Warning System for Rhodesian Sleeping Sickness in Savannah Areas: Man-Like Traps for Tsetse Flies. *PLoS Neglected Tropical Diseases*, *6*(12). https://doi.org/10.1371/journal.pntd.0001978

Van Hoek, A. H. A. M., Mevius, D., Guerra, B., Mullany, P., Roberts, A. P., & Aarts, H. J. M. (2011). Acquired antibiotic resistance genes: An overview. *Frontiers in Microbiology*, *2*(September), 1–27. https://doi.org/10.3389/fmicb.2011.00203

VanderMolen, K. M., Raja, H. A., El-Elimat, T., & Oberlies, N. H. (2013). Evaluation of culture media for the production of secondary metabolites in a natural products screening program. *AMB Express*, *3*(71), 1–7.

Vernikos, G., Medini, D., Riley, D. R., & Tettlin, H. (2015). Ten years of pan-genome analyses. *Current Opinion in Microbiology*, *23*, 148–154. https://doi.org/10.1016/j.mib.2014.11.016

Vivaldo, G., Masi, E., Taiti, C., Caldarelli, G., & Mancuso, S. (2017). The network of plants volatile organic compounds. *Scientific Reports*, *7*, 1–18. https://doi.org/10.1038/s41598-017-10975-x

Vos, M., Quince, C., Pijl, A. S., de Hollander, M., & Kowalchuk, G. A. (2012). A comparison of rpoB and 16S rRNA as markers in pyrosequencing studies of bacterial diversity. *PLoS ONE*, *7*(2), 1–8. https://doi.org/10.1371/journal.pone.0030600

Wallingford, A. K., Cha, D. H., Linn, C. E., Wolfin, M. S., & Loeb, G. M. (2017). Robust manipulations of pest insect behavior using repellents and practical application for integrated pest management. *Environmental Entomology*, *46*(5), 1041–1050. https://doi.org/10.1093/ee/nvx125

Wang, Changlu, Wang, Z., Qiao, X., Li, Z., Li, F., Chen, M., Wang, Y., Huang, Y., & Cui, H. (2013). Antifungal activity of volatile organic compounds from Streptomyces alboflavus TD-1. *FEMS Microbiology Letters*, *341*(1), 45–51. https://doi.org/10.1111/1574-6968.12088

Wang, Chieh-mei, & Cane, D. E. (2008). Biochemistry and Molecular Genetics of the Biosynthesis of the Earthy Odorant Methylisoborneol in Streptomyces Coelicolor. *Journal of American Chemical Society*, *130*(28), 8908–8909. https://doi.org/10.1021/ja803639g.Biochemistry

Wang, L., Lee, F., Tai, C., Kasai, H., & Lee, F. (2007). Comparison of gyrB gene sequences , 16S rRNA gene sequences and DNA – DNA hybridization in the Bacillus subtilis group. *International Journal of Systematic and Evolutionary Microbiology*, 1846–1850. https://doi.org/10.1099/ijs.0.64685-0

Wang, Z., Wang, C., Li, F., Li, Z., Chen, M., Wang, Y., Qiao, X., & Zhang, H. (2013). Fumigant activity of volatiles from Streptomyces alboflavus TD-1 against Fusarium moniliforme Sheldon. *Journal of Microbiology*, *51*(4), 477–483. https://doi.org/10.1007/s12275-013-2586-y

Wu, D., Jospin, G., & Eisen, J. A. (2013). Systematic Identification of Gene Families for Use as '"Markers"' for Phylogenetic and Phylogeny-Driven Ecological Studies of Bacteria and Archaea and Their Major Subgroups. *PLoS ONE*, *8*(10). https://doi.org/10.1371/journal.pone.0077033

Xing, M., Zheng, L., Deng, Y., Xu, D., Xi, P., Li, M., Kong, G., & Jiang, Z. (2018). Antifungal activity of natural volatile organic compounds against litchi downy blight pathogen peronophythora litchii. *Molecules*, *23*(2), 1–15. https://doi.org/10.3390/molecules23020358

Xu, J., Wang, Y., Xie, S. J., Xu, J., Xiao, J., & Ruan, J. S. (2009). Streptomyces xiamenensis sp. nov., isolated from mangrove sediment. *International Journal of Systematic and Evolutionary Microbiology*, *59*(3), 472–476. https://doi.org/10.1099/ijs.0.000497-0

Xu, P., Ma, W., Han, H., Hou, B., & Jia, S. (2014). Characterization of Naphthalene Degradation by Streptomyces Sp. QWE-5 Isolated From Active Sludge. *Water Science & Technology*, *70*(6), 1129–1134. https://doi.org/10.2166/wst.2014.356

Yang, M., Lu, L., Pang, J., Hu, Y., Guo, Q., Li, Z., Wu, S., Liu, H., & Wang, C. (2019). Biocontrol activity of volatile organic compounds from Streptomyces alboflavus TD-1 against Aspergillus flavus growth and aflatoxin pro- duction. *Journal of Microbiology*, *57*(5), 396–404. https://doi.org/10.1007/s12275-019-8517-9

Zeigler, D. R. (2003). Gene sequences useful for predicting relatedness of whole genomes in bacteria. *International Journal of Systematic and Evolutionary Microbiology*, *53*, 1893–1900. https://doi.org/10.1099/ijs.0.02713-0

Zeng, F., Xu, P., & Leal, W. S. (2019). Odorant Receptors From Culex Quinquefasciatus and Aedes Aegypti Sensitive to Floral Compounds. *Insect Biochemistry and Molecular Biology*, *113*, 1–9. https://doi.org/10.1016/j.ibmb.2019.103213

Zhang, J., Chen, J., Hu, L., Jia, R., Ma, Q., Tang, J., & Wang, Y. (2020). Antagonistic action of Streptomyces pratensis S10 on Fusarium graminearum and its complete genome sequence. *Environmental Microbiology*, *23*(4), 1925–1940. https://doi.org/10.1111/1462-2920.15282

Zhang, K. mei, Shen, Y., Yang, J., Miu, X., Bhowmik, P. C., Zhou, X., Fang, Y. ming, & Xing, B. shan. (2019). The defense system for Bidens pilosa root exudate treatments

in Pteris multifida gametophyte. *Ecotoxicology and Environmental Safety*, *173*(February), 203–213. https://doi.org/10.1016/j.ecoenv.2019.01.097

Zhang, Q., Schlyter, F., Chen, G., & Wang, Y. (2007). Electrophysiological and Behavioral Responses of Ips Subelongatus to Semiochemicals From Its Hosts, Non-Hosts, and Conspecifics in China. *Journal of Chemical Ecology*, *33*, 391–404. https://doi.org/10.1007/s10886-006-9231-8

Zhao, B., Lin, X., Lei, L., Lamb, D. C., Kelly, S. L., Waterman, M. R., & Cane, D. E. (2008). Biosynthesis of the Sesquiterpene Antibiotic Albaflavenone in Streptomyces coelicolor A3 (2). *The Journal of Biological Chemistry*, *283*(13), 8183–8189. https://doi.org/10.1074/jbc.M710421200

Zhao, Y., Xiang, S., Dai, X., & Yang, K. (2013). A simplified diphenylamine colorimetric method for growth quantification. *Applied Microbiology and Biotechnology*, *97*(11), 5069–5077. https://doi.org/10.1007/s00253-013-4893-y

Zhu, X., Xu, B., Kader, A., Song, B., Zhang, Z., Li, F., & Yang, S. (2020). Behavioral Responses of Scolytus schevyrewi (Coleoptera: Curculionidae: Scolytinae) to Volatiles From Apricot Tree (Rosales: Rosaceae). *Environmental Entomology*, *49*(3), 586–592. https://doi.org/10.1093/ee/nvaa027

Appendices

A.2.1: Media preparations

SFM – 1 liter

Soy flour defatted	20 g
Mannitol	20 g
Agar (if making solid medium)	20g

Blend the soy flour for further refinement. If making solid agar, prepare agar and soy flour in one container and mannitol in another. Combine both after autoclave.

The synthetic medium – 1 liter

Glucose	30 g
$(NH_4)_2SO_4$	1 g
$CaCO_3$	2.5 g
NaCl	1 g
Agar (if making solid medium)	20 g

Make solutions of the following salts in dH_2O, then add specified volumes to the desire 1 L of medium being prepared.

Salts	solution volume	amount of solution to make 1 L medium
0.025 g $MgSO_4.7H_2O$	40 ml	2 ml
0.125 g K_2HPO_4	50 ml	12.5 ml
0.125 g $FeSO_4.7H_2O$	50 ml	12.5 ml
0.125 g ZnSO4.H_2O	50 ml	12.5 ml
0.125 g $MnSO_4.4H_2O$	50 ml	12.5 ml

Add all salt solutions before making up the final volume of 1 L. Adjust the pH to 7 with NaOH, then autoclave.

YMS – 1 liter

Yeast extract	4 g
Malt extract	10 g
Soluble starch	4 g

Agar (if making solid medium)	20 g

Adjust the pH to 7.4 with NaOH, then autoclave.

YPD – 1 liter

Yeast extract	10 g
Peptone	20 g
Dextrose	20 g
Agar	20 g

Combine the above ingredients and make up the volume to 1 L with dH_2O, then autoclave.

OBB- 1 liter

1. Weigh 20 g of oat bran (SKU6051S18, Bob's Red Mill Natural Foods, USA) in 500 ml of dH_2O and heat in a microwave until boiling.
2. Once boiling, reduce power to 1, and leave the mixture boil for 15-20 mins
3. Place the mixture on a stir plate to cool with stirring for 45-60 mins. Speeding up cooling may result in reduced and inconsistent thaxtomin production.
4. Filter the mixture through 4 single layers of cheesecloth in a Buchner funnel using a vacuum into a sidearm flask.
5. Scoop the remaining oatmeal into the funnel and rinse with 200-300 ml dH_2O
6. Apply a vacuum until most of the liquid is drained
7. Adjust the pH of the broth to 7.2 with 1M NaOH
8. Add 2 ml/L of trace solution (recipe below)
9. Add dH_2O to a volume of 1 liter
10. Dispense and autoclave for 20 mins

Trace elements for OBB (1 liter)

$ZnCl_2$	40 mg
$FeCl_3.6H_2O$	200 mg
$CuCl_2.2H_2O$	10 mg
$MnCl_2.4H_2O$	10 mg
$Na_2B_4O_7.10H_2O$	10 mg
$(NH_4)_6Mo_7O_{24}.4H_2O)$	10 mg

A.2.3.2.1: *Streptomyces* reference sequences used to make consensus sequence for *rpoB* primer design

The six *Streptomyces* sequences (NCBI accession provided) were used to create a consensus sequence which was then used to run against NCBI non-redundant nucleotide database specific for *Streptomyces*. The top 100 hits were used to create primers used in this study for *rpoB* gene sequencing of the isolates.

>*Streptomyces_coelicolor*_CP042324.1H

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGAAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCACCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCTCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACAGCCAGCGGGTTCAGGATGATGTCGACAGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTG
GCCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGA
TGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGCCCAGCGCGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACGTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCG
AAGGGGTTGATGCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCG
ACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGA
CAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCT
GCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGC
TCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCG
ACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATG
GAGGAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGA
GCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTG
AAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGAGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATC
GGCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCG
TTGCCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGG
GGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAA----------

>*Streptomyces_avermitilis*_AP019621.1

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCAGGGATGCCGGGCT
CGGGAATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG

CGGCGCCATAAGCCTCAAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAAC
GGCTCGCCGGAGCGGCCGTCGAACAGCGGCGCCTTGCCGGTCGGGAGCACCATGCGCTCTCCGTCGCGGTTCGGGATCGTGTGCTGCAGCAGACC
CGCGAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTGGTGCCCGGGGCGACCTGGTCGGCCTCGATGGCCTGCAGGCGCTGC
GCCCAGTCGTCGGCGAGTCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCGAGGGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACACCCTTGT
TGCCGTGGCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGAACCAGCTGGTTCACGCCCGGCGGCAGCTCGT
CGCCCTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGCACCTCGCGCGCCTTCT
CACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACC
TCGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTCGCCGGGCCGTCGGCCAGAACCT
GGCCGGCGATGACCCGGGCGCCCTCGTCCACGACAACCTTCTGGTTGACGGACGTGCCCTGGTTGGAGCGGGAGAACTTGTGCAGGCGGTACGTG
GTGTACGTGCCGTCGTCGTTGGCGGTGGTGACGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCCGACTTGAGCACGTCGCCGGCGTCG
ACGGCGCAGCGGTACTCCATGCCGGTGCCGACGAGCGGCGCCTCGGACTTGATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCAC
CGGGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGGACGCGGGTCTCCTCGAACCGGAAGTCGTCGCCGAGCGGCGCGTTGGCCTGCGC
GATGACGAATCGGTCCTCTTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACGACACCGTCGGTGACCCGGCGGTACGGCGTCTCGACGAAACC
GAACGCGTTGACGCGGCCGTAAGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGGCCGTAGTGC
GACGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGAGTGAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGAGTCGTCATGCGCTCGCGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGC
TCGCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGAGGTGCCGTTGTCGCCAACGGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATCGACGAGATGATGTCCTCGACGGTCAGGATCCCGGCGTCCAGCGGCGCCTCGCCGCCCAGCTTCTTG
TTGACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCAAGAAGCGTCTGCGCGGCCTCGCGGGTCGGGGG
CTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTC
GCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTCAGGAGAACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCGCCCCGGGAGGGGATGACCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCCGGCGAGCGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAA
GTCGCGGTCCTTGCACTCGTCGATGGAGTTCTTCGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGGGACATCGACCCGGAGAAGTCCTCGATCG
GAGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCGCTGTCCAGAGCCGCCTCGACACGAGCCTTCCACGCGTCGTT
GCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCAAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGG
CGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAA----------

>*Streptomyces_clavuligerus*_CP027858.1

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCTTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAG
GCGGCGCCGTATGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATC
GAGTAGGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGA
ACGGCTCGCCGGAGCGGCCGTCGAACATCCTGGCCTTGCCGGAGGGGAGGACCATCCGGTCACCGTCGCGGTTGGGGATCGTGTTCTCGAAGAGA
CCGGCGATCTCGTCCTCGCGGGCGCCGTCGAAGACCGGGGTGGCGACATTGGTGCCGGGCTGGACCTCGTCCGCGCCGATGGACTTCAGCCGCTC
CATCCACTCCTCGTT--GCCC-
TCGACCTTCCAGCCCTGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGGACACCCAGCGGGTTGAGGAT
GATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGAGGATCTTGGAGATGACACCCTTGTTGCCGTGGCGGCCGGCCAGCT
TGTCACCGTCGGTGATCTTGCGCTTCTGGGCGACATAGACGCGGACCAGCTGGTTCACGCCCGGAGGCAGCTCGTCACCCTCCTCGCGGTCGAAGA
CGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTCACCGAAGATCGCGCGGAGCA
GGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTCGGGGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACCTCGGCGCCGATGCGGATGATG
CCGCGCTCGTCGAGGTCGGCCAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCCAGCTTGGTGTCCCGGGCGTCGAC
CTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCGTAGTTGTGGCCCTCCCACGG
CATGAACGCCACGAGGAGGTTCTTGCCGAGCGCCATCTCGCCGTCTTCGGTGGCGGGGCCGTCGGCCAGGACCTGGCTCTCGATGACCCGGTCGC
CCTCGGAGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGAGAACTTGGCCACCCGGTAGGTGGTGAAGGTGCCGTCGTCGTTG
GCGACGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCGGCCTTGATCACGTCACCGGCGTCGACCGCGCAGCGGTACTCCATA
CCGGTACCGACCAGCGGCGCCTCGGCCGTGATCAGCGGAACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCACGGTTGGCGTCGTCGTGCTC
CAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACCTCGGCCGGCAGGACGTAGTCGACCTCG
CCGCCGCGGCGGCGGACCAGCACCCGGGCCTCGGTGAACTGCAGCTCGTCGTTGAGCGCGGCGTTGGCCTGCGCGATGACGAAGCGGTCCTCCTC
GTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGGGTCTCCACGAAGCCGAACGCGTTGACCCGGCCGT
AGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGGCCGTAGTGCGAGGGGTGCACGTCACGGAC

CTCGAAGCCCGCCCGCTCACGGGAGAGACCACCCGGGCCGAGCGCCGACAGACGGCGCTTGTGGGTGAGACCCGACAGCGGGTTGTTCTGGTCCA
TGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGGGGCGTGATCGCCTCGACG
TCCTGGGTGGTCATGCGCTCGCGCACGACACGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTCGCCGACGTTGCGCAGACGA
CGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTGCGAACCGTTCTCACCGATCGTCTCGGTCTCGCCCGCGTGCAGCTTGACCAGG
TACTTGATGGTCGCGATGACATCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTTCACCTTGTAGCGACCGACC
TTGGCCAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCCAGCAGCGTCTGGGCGGCCTCGCGCGTCGGGGGCTCGCCCGGACGGAGCTTGCG
GTAGATGTCGAGGAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATCGACTCGTACTCGCCGAACTCCTCAAGGATCTG
CTCGGTCGTCCAGCCGAGGGCCTTGAGCAGCACCGTGACGGACTGCTTGCGCTTGCGGTCGATACGGACACCGACCATGTCGCGCTTGTCGATCTC
CATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGTGTAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGAGTCGAAGTAGACACCGG
GGGAGCGCACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATCACGAAGGTGCCCTTGTTGGTCATGAGCGGGAAGTCGCCCATGAAGACC
GTCTGAGACTTGATCTCACCGGTCTCGTTGTTCGTGAACTCGGCCGTGACGAAGAGCGGCGCGGCGTAGGTGAAGTCGCGCTCCTTGCACTCGTCG
ATGGAGTTCTTCGGAGGCTCGAAGCGGTGGTCCCGGAAGGTCAGGGACATCGACCCGCTGAAGTCCTCGATCGGGGAGATCTCCTCGAAGATCTC
CTCCAGACCGGACTTCGTGGGGACGTCCTGTCCACTCTCCAGAGCCGCCTCGACCCGAGCCTTCCACGCGGCGTTGCCGAGCAGCCAGTCAAAGCT
CTCGGTCTGGAGCGCGAGGAGGTTCGGAACCTCAAGAGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCAGTGCTGGCGCCGTTGTTCA
TATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>*Streptomyces_venezuelae*_CP029189.1

TCAGACCTCTTCGACGCTGCTCGGCTCACGCCGGGACAGGTCGATACCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCAGAGCGACTGCATTTCCTTGATGAGCACCTTGAAGGACTCGGGAATGCCCGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTAAGCCTCGAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGTGCCTTACCACCCAGCGGCTGCTGCGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGGAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGGCGCGCCTTGCCGGACGGGAGGACCAGGCGGTCACCGTCGCGGTTCGGGATGGTGTGCTCGAAGAGG
CCGGCGAGCTCGTCCTCGCGGGCACCGTCGAACACGGGGGTGGCGACGTTGGTACCGGGCTCGACGCGGTCGGCCCCGATGACCTTCAGACGCTC
GGCCCACTCCTCCGCCAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GGACACCCAGCGGGTTCAGGATGATGTCGACCGGCGTGCCGTCTTCCAGGAAGGGCATGTCCTCGATCGGAAGGATCTTGGAGATGACACCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGGGCGACGTAGACGCGGACCAGCTGGTTCACGCCCGGAGGAAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGGACACCGATGACCTTGCCGATCTCACCGTGAGGGACCTTGAGCGAGGTGTCACGCACCTCGCGGGCCTT
CTCACCGAAGATCGCGCGGAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGA
CGTCCGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCC
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGACCTTCCCAGGGCATGAACGCGACGAGCAGGTTCTTGCCGAGGGCCATTTCGCCCTCTTCGGTCGCGGGACCGTCGGCGAGGAC
CTGGGACTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTCGAGCGGGAGAACTTGGCGACGCGGTACGT
GGTGTACGTGCCGTCGTCGTTGGCGACGGTGATGTAGTCGGCCGAGACCTCCTGGACGACACCGTCCTTCTCCGCACGGATCGAGTCACCGGCGTC
GACCGCACAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCCGCCTTGATGAGCGGAACGGCCTGGCGCATCATGTTCGCGCCCATGAGGG
CACGGTTGGCGTCGTCGTGCTCCAGGAAGGGGATCATCGCGGTCGCGACGGACACCATCTGGCGCGGGGAGACGTCCATGTAGTCGACGTCGTCG
CCGGCGATGTAGTCGATCTCGCCGCCACGACGGCGCACCAGCACGCGGTTCTCGGTGAAGCGCATGTCCTCGGACAGGCCGGCGTTGGCCTGGGC
GATGACGAAGCGGTCTTCCTCGTCGGCCGTGAGGTAGTCGACCTCGTCGGTGACGACACCGTCGATGACCTTGCGGTACGGGGTCTCGACGAAAC
CGAACGCGTTGACGCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGGCCGTAGTG
CGACGGGTGGACGTCTCGGACCTCGAAGCCGGCCCGCTCACGGGACAGACCGCCGGGGCCCAGCGCCGACAGGCGGCGCTTGTGGGTCAGGCCC
GAGAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGT
CTGCGGCGTGATGGCCTCGACGTCCTGCGTGGTCATGCGCTCGCGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGA
GCTCGCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGAACGGCCGGACTCACCGGTGGTCTCGGTC
TCGCCGGCGTGCAGCTTGACCAGGTACTTGATGGTCGCGATGACGTCGTCGGTGGTGAGCACGCCGGCATCGAGCGGCTCGTCGGCGCCGAGCTT
CTTGTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGAGCGGCCTCACGGGTCGG
CGGCTCGCCCGGGCGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGT
ACTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTCAGGAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACG
CCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCG
ATGGAGGAGTCGAAGTAGACACCGGGGGAGCGGACAAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCA
TAAGCGGGAAGTCGCCCATGAAGACCGTCTGAGACTTGATCTCTCCGGTCTCGTTGTTCGTGAACTCGGCCGTGACGAACAGCGGAGCCGCGTACG
TGAAGTCGCGCTCCTTGCACTCGTCGACAGAGTTCTTCGCCGGCTCGAAACGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGA
TCGGCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTCCTGTCCACTCTCAAGCGCCGACTCGACGCGAGACTTCCAGGCGG
CGTTGCCGAGAAGCCAGTCAAAGCTCTCGGTCTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGC
GGGGCGGTGCTGGCACCGTTGTTCGTATT---GGTCGAGGCGTTGCGCGAGGCGGCCAA----------

>*Streptomyces_lividans*_CP009124.1

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT

CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCACCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCGCGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTG
GCCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGA
TGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGCCCAGCGCGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACGTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCG
AAGGGGTTGATGCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCG
ACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGA
CAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCT
GCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGC
TCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCG
ACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATG
GAGGAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGA
GCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTG
AAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGAGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATC
GGCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCG
TTGCCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGG
GGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>*Streptomyces_scabiei*_FN554889.1

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGATGGAGATCGGGTCCGGGAA
CGGCTCACCGGAGCGGCCGTCGAACATCCTCGCCTTACCGGTCGGCTGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGAC
CGGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCCGGGGGCGACGGAGTCGGCGCCGATGACCTGCAGACGCTT
CGCCCAGTCGTCGCCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGCACCTGTCCCGGGTTCATTCGGGACG
GGACGCCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACACCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCG
TCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTC
TCACCGAAGATCGCGCGGAGCAGCCGCTCCTCGGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGAC
CTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGA
GCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCC
TCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACC
TGGCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGCGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCGGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCCAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCCC
CGGGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGTTCTCGGTGAAGCGCATGTCGTCGTCGAGCGTGGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCCTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACGACACCGTCGGTGACCCGGCGGTACGGCGTCTCCACGAAAC
CGAACGCGTTGACGCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTG

CGAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGACAGACCACCGGGACCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCC
GACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGT
CTGGGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACACGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAG
CTCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCGTGCCGTTGTCGGCAGCCGTCTCGGTCTC
GCCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGACGTCCTCGACGGTCAGGACGCCCGCGTCGAGCGGAGCCTCCGCACCCAGCTTCTT
GTTGACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGCCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCCAGGATCTGCTCGGTGGTCCAGCCGAGTGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATACGCACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCCGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTGGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAA
GTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGG
GGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCGTTCTCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGTTG
CCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAGAGATGCGCAGCGGGGC
AGTGCTGGCGGCGGGATTCGTATTCGAGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

A.2.3.2.2: *Streptomyces* reference sequences for *rpoB* primer design

The *Streptomyces rpoB* sequences were used to create the primers used for *rpoB* sequencing in this study. The NCBI accession numbers are provided for the *Streptomyces* from the top 100 hits returned against a search with the consensus sequence (lcl|Query_40537) created using *Streptomyces rpoB* sequences in A.2.3.2.1.

>lcl|Query_40537

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGCTCCGGGAAC
GGCTCACCGGAGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGAC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGACCTGCAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACCGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGAGGATCTTGGAGATGACACCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCCGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTG
GCCCTCGATGATCCGGTCGCCCTCGTCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGAT
GTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCTCCTTCTCGGCCTTGACCACGTCACCGGCGTCGAC
GGCGCAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGCCCAGCGCGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACCACACCGTCGACGACCTTGCGGTACGGCGTCTCGACGAAACCGA
ACGCGTTGACGCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGA
CGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGCGCCGACAGACGACGCTTGTGCGTCAGACCCGAC
AGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTG
CGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCCAGACCCGTACGGACCTGGTTCTGGATCAGCTC
GCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTACCGGCCGTACTCGCCGACCGTCTCGGTCTCACC
GGCGTGCAGCTTCACCAGGTACTTGATGGTCGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTT
CACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTCGGGGGCT
CGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGC
CGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGACC
ATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAG
GAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGCGG
GAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAAGT
CGCGGTCCTTGCACTCGTCGATGCAGTTCTTCGGAGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGC
GAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCGCTCTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCGTTG
CCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGC
GGTGCTGGCACCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>emb|LT629768.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG

CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCACCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCGCGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTG
GCCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGA
TGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGCCCAGCGCGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACGTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCG
AAGGGGTTGATGCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCG
ACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGA
CAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCT
GCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGC
TCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCG
ACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATG
GAGGAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGA
GCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTG
AAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGAGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATC
GGCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCG
TTGCCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGG
GGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP009124.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCACCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCGCGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTG
GCCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGA
TGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGCCCAGCGCGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACGTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCG
AAGGGGTTGATGCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCG
ACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGA

CAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCT
GCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGC
TCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCG
ACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATG
GAGGAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGA
GCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTG
AAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGAGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATC
GGCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCG
TTGCCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGG
GGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP071800.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCACCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCGCGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTG
GCCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGA
TGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGCCCAGCGCGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACGTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCG
AAGGGGTTGATGCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCG
ACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGA
CAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCT
GCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGC
TCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCG
ACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATG
GAGGAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGA
GCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTG
AAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGAGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATC
GGCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCG
TTGCCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGG
GGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP071123.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG

CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCACCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCGCGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTG
GCCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGA
TGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGCCCAGCGCGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACGTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCG
AAGGGGTTGATGCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCG
ACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGA
CAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCT
GCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGC
TCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCG
ACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATG
GAGGAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGA
GCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTG
AAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGAGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATC
GGCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCG
TTGCCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGG
GGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP050522.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGAAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCACCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCTCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACAGCCAGCGGGTTCAGGATGATGTCGACAGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTG
GCCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGA
TGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGCCCAGCGCGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACGTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCG
AAGGGGTTGATGCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCG
ACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGA

CAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCT
GCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGC
TCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCG
ACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATG
GAGGAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGA
GCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTG
AAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGAGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATC
GGCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCG
TTGCCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGG
GGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP042324.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGAAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCACCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCTCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACAGCCAGCGGGTTCAGGATGATGTCGACAGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTG
GCCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGA
TGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGCCCAGCGCGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACGTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCG
AAGGGGTTGATGCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCG
ACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGA
CAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCT
GCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGC
TCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCG
ACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATG
GAGGAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGA
GCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTG
AAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGAGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATC
GGCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCG
TTGCCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGG
GGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>emb|AL939121.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG

CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGAAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCACCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCTCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACAGCCAGCGGGTTCAGGATGATGTCGACAGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTG
GCCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGA
TGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGCCCAGCGCGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACGTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCG
AAGGGGTTGATGCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCG
ACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGA
CAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCT
GCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGC
TCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCG
ACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATG
GAGGAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGA
GCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTG
AAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGAGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATC
GGCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCG
TTGCCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGG
GGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP026730.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCACCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGCCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGT
CGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCGCGGGCCTTCT
CACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCT
GGCCCTCGATGACCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCG
GCGGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCAC
CGGGGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGCCCAGCGCGGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCC
GAAGGGGTTGATGCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGC
GACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCG

ACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGAGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAG
CTCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTC
ACCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAACCGAGGGCCTTGAGCAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCCGGCGAACGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAA
GTCACGGTCCTTGCACTCGTCGATGCTGTTCTTGGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCG
GCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCACGCGTCGT
TGCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGG
GCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP045547.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATCGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGTCGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGGCCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGCCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTG
GCCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGA
TGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCCCCG
GGGACGTAGTCGACCTCTCCGCCACGGCGGCGGACCAGGACGCGGGCCTCGGCGAACCGCAGGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACGTCGTCCGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCGA
ACGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGAC
GGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGACA
GCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGC
GGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTC
GCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGGTCGTCTCGGTCTCACC
GGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCCTCCGCGCCGAGCTTCTTGTT
CACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGGCT
CGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGC
CGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGACC
ATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAG
GAGTCGAAGTAGACACCCGGCGAACGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGCGG
GAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAAGT
CGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGC
GAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCGTTG
CCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGC
GGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP034353.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG

CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGTCGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCGCCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGCCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACCTC
GGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCT
TGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCG
TAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCGAGGACCTGG
CCCTCGATGACGCGGTCGCCCTCGGAGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGAT
GTACGTGCCGTCGTCGTTGGTCGTCGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGGCTTCTCGGCCTTGACCACGTCGCCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATCGCGGTGGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCGCGGCGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCGA
ACGGGTTGATGCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGA
CGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGAC
AGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTG
CGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTC
GCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCACC
GGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTT
CACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGGCT
CGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGC
CGAACTCCTCGAGGATCTGCTCGGTGGTCCAACCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGACC
ATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAG
GAGTCGAAGTAGACGCCGGGCGAACGGACCAGCTGCGAGACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGCG
GGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAAG
TCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGG
GGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCGTT
GCCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGG
GCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP054920.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAA
CGGCTCACCGCTGCGGCCATCGAACAGCCGCGCCTTGCCGGTCGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCGCCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGCCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACCTC
GGCACCGATGCGGATGATGCCGCGCTCGTCGAGATCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCT
TGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCG
TAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCGAGGACCTGG
CCCTCGATGACGCGGTCGCCCTCGGAGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGAT
GTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGGCTTCTCGGCCTTGACCACGTCGCCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCGCGGCGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCGA
ACGGGTTGATGCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGA
CGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGAA

AGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTG
CGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTC
GCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCACC
GGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTT
CACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGGCT
CGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGC
CGAACTCCTCGAGGATCTGCTCGGTGGTCCAACCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGACC
ATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAG
GAGTCGAAGTAGACGCCGGGCGAACGGACCAGCTGCGAGACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGCG
GGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAAG
TCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGG
GGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCGTT
GCCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGG
GCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP074111.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGTCGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCGCCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGCCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCCAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCGAGGACCTG
GCCCTCGATGACGCGGTCGCCCTCGGAGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGA
TGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGGCTTCTCGGCCTTGACCACGTCGCCGGCGTCGG
CGGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCA
CGGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACC
GGGGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACACGGGCCTCGGCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCG
ATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCG
AACGGGTTGATGCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCG
ACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGA
CAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCT
GCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGC
TCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTTG
TTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGG
CTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTC
GCCGAACTCCTCCAGGATCTGCTCGGTGGTCCAACCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACGCCGGGCGAACGGACCAGCTGCGAGACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAG
CGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGA
AGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATC
GGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTC
GTTGCCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCG
GGGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP009754.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGTGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATCGTCAGCAGCTCCTGGAGGGCGTACG

CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGACCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCGCCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGTACCTGTCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGT
CGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCT
CACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACCAGGATGTCACCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGGCCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCGAGGACCTG
GCCCTCGATGACGCGGTCGCCCTCGGAGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGA
TGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGGCTTCTCGGCCTTGACCACGTCGCCGGCGTCGG
CGGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCG
CGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACC
CGGGACGTAGTCGACCTCGCCGCCGCGGCGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCG
ATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCG
AACGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGA
CGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGAC
AGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTG
CGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTC
GCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCACC
GGCGTGCAGCTTCACCAGGTACTTGATCGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTT
CACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCGCGCGTGGGGGGCT
CGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGC
CGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGACC
ATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAG
GAGTCGAAGTAGACGCCGGGCGAACGGACCAGCTGCGAGACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAGCG
GGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAAG
TCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGG
GGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCGTT
GCCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCCAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGG
GCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP012382.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGTGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCACCGGAGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGG
CCGGCCAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGGCCTGCAGACGCTG
GGCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGCCCCGGGTTCATTCGGGACG
GGACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCG
TCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCACCGTGCGGCACCTTCAGGGAGGTGTCGCGCACCTCGCGCGCCTTC
TCACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGAC
CTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGA
GCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCC
TCGTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGGCCGTCGGCCAGGACC
TGGCCCTCGATGACCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTTCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGT
GATGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTC
GGCGGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGG
CGCGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCA
CCGGGGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACACGGGCCTCGGAGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCG
CGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACGTCGTCCGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAAC
CGAACGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCCGGCGTCTCGATCGGGCACATGCGGCCGTAGTGC
GACGGGTGGACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCG

ACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACCCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGC
TCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATCGTGGCGATGACGTCGTCGGTGGTGAGGACGCCGGCGTCCAGGGGCTCGTCGGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCCAGGATCTGCTCGGTGGTCCAACCGAGAGCCTTGAGCAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCCGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCCGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAA
GTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCG
GCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCACGCGTCGT
TGCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGG
GCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP012949.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGTGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCACCGGAGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGAAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGG
CCGGCCAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGGCCTGCAGACGCTG
GGCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGCCCCGGGTTCATTCGGGACG
GGACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCG
TCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCACCGTGCGGCACCTTCAGGGAGGTGTCGCGCACCTCGCGCGCCTTC
TCACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGAC
CTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGA
GCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCC
TCGTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGGCCGTCGGCCAGGACC
TGGCCCTCGATGACCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTTCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGT
GATGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTC
GGCGGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGG
CGCGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCA
CCGGGGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACACGGGCCTCGGAGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCG
CGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACGTCGTCCGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAAC
CGAACGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCCGGCGTCTCGATCGGGCACATGCGGCCGTAGTGC
GACGGGTGGACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACCCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGC
TCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATCGTGGCGATGACGTCGTCGGTGGTGAGGACGCCGGCGTCCAGGGGCTCGTCGGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCCAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTGAGCAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCCGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAA
GTCACGGTCCTTGCACTCGTCGATGCTGTTCTTGGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCG
GCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCACGCGTCGT
TGCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGG
GCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP019724.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG

CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGACC
GGCCAGCTCGTCCTCACGCGCACCGTCGAAGACCGGCGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGGCCTGCAGGCGCTGCG
CCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGG
ACACCCAGCGGGTTCAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTTG
CCGTGCCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGGACCAACTGGTTCACGCCCGGCGGCAGCTCGTCG
CCCTCTTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCTCA
CCGAAGATCGCACGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCTC
GGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCT
TGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCG
TAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTGG
CCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTTCCCTGGTTGAACCGGGAGAACTTGGCCAGGCGGTACGTGATG
TACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGGCG
GCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCACG
GTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCGG
GGACGTAGTCGACCTCGCCGCCACGGCGGCGAACCAGGACGCGGGCCTCGGAGAAGCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGAT
GACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACGTCGTCCGTGACCTGGCCCTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCGAA
CGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGTCCTTCCGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGACG
GGTGGACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGACAG
CGGGTTGTTCTGGTCCATGAACTGGGAGAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCG
GCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTCGC
CGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGGTCGTCTCGGTCTCACCGG
CGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTCAGAACGCCGGCGTCCAGGGGCTCGTCGGCGCCGAGCTTCTTGTTCA
CCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGGCTCG
CCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCG
AACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGACCAT
GTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGA
GTCGAAGTAGACACCCGGCGAGCGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGCGGG
AAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAAGTC
GCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGCG
AGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCGTTGC
CGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGC
GGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP016795.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCGCCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGCCCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATGCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTGACACCGGGGGGAAGCTCGT
CGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGGGAGGTGTCGCGCACCTCGCGCGCCTTCT
CGCCGAAGATCGCGCGGAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGGCCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGGCCGTCGGCCAGGACCT
GGCCCTCGATGACCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGGCTTCTCGGCCTTGACCACGTCGCCGGCGTCG
GCGGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGGCGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCAC
CGGGGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACGCGGGCCTCGGCGAACCGCAGGTCGTCGTTGAGCGTGGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACGTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTAGGGCGTCTCGATGAAGC
CGAACGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGC
GACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCG

ACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTGCGGACCTGGTTCTGGATCAG
CTCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTC
ACCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGATGTCGTCACTGGTGAGCACGCCGGCGTCCAGCGGCTCTTCGGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGAAGCGTCTGCGCCGCCTCGCGCGT-
GGGGGCTCACCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTC
GTACTCGCCGAACTCCTCCAGGATCTGCTCGGTCGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCAC
ACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCCGAGAAGATGTCCTTGTCGGACGTCTTGTC
GATGGAGGAGTCGAAGTAGACGCCCGGCGAACGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTC
ATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTA
CGTGAAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGG-
GGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTT
GGTGGGGACGTCCTGACCTGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCG
CGAGGAGGTTCGGAACCTCGAGGGGCTCCCTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAG
GCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP036534.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCGCCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGCCCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATGCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTGACACCGGGGGGAAGCTCGT
CGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGGGAGGTGTCGCGCACCTCGCGCGCCTTCT
CGCCGAAGATCGCGCGGAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGGCCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGGCCGTCGGCCAGGACCT
GGCCCTCGATGACCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGGCTTCTCGGCCTTGACCACGTCGCCGGCGTCG
GCGGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGGCGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCAC
CGGGGACGTAGTCGACCTCCCCGCCACGGCGGCGGACCAGGACGCGGGCCTCGGCGAACCGCAGGTCGTCGTTGAGCGTGGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACGTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTAGGGCGTCTCGATGAAGC
CGAACGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGC
GACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTGCGGACCTGGTTCTGGATCAG
CTCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTC
ACCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGATGTCGTCACTGGTGAGCACGCCGGCGTCCAGCGGCTCTTCGGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGAAGCGTCTGCGCCGCCTCGCGCGTGGGGG
GCTCACCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCCAGGATCTGCTCGGTCGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCG
ACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCCGAGAAGATGTCCTTGTCGGACGTCTTGTCGATG
GAGGAGTCGAAGTAGACGCCCGGCGAACGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGA
GCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTG
AAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGGGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGAT
CGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCTGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTC
GTTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGGAGGTTCGGAACCTCGAGGGGCTCCCTGATCTTTGCAAAGGAGATGCGCAGCG
GGGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP043317.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT

CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATCGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGGAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGCCCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATGCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTGACACCGGGGGGAAGCTCGT
CACCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGGGAGGTGTCGCGCACCTCGCGCGCCTTCT
CGCCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCGCGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGGCCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGGCCGTCGGCCAGGACCT
GGCCCTCGATGACCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGGCTTCTCGGCCTTGACCACGTCGCCGGCGTCG
GCGGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCAC
CGGGGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACGCGGGCCTCGGCGAACCGCAGGTCGTCGTTGAGCGTGGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACGTCGTCCGTGACCTGGCCCTCGACGACCTTGCGGTAGGGCGTCTCGATGAAGCC
GAACGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGC
GACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTGCGGACCTGGTTCTGGATCAG
CTCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTC
ACCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGATGTCGTCGCTGGTGAGCACGCCGGCGTCCAGCGGCTCTTCGGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGAAGCGTCTGCGCCGCCTCGCGCGTGGGGG
GCTCACCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCCAGGATCTGCTCGGTCGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCG
ACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCCGAGAAGATGTCCTTGTCGGACGTCTTGTCGATG
GAGGAGTCGAAGTAGACGCCCGGCGAACGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGA
GCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTG
AAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGGGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGAT
CGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGGCCTGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTC
GTTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGGAGGTTCGGAACCTCGAGGGGCTCCCTGATCTTTGCAAAGGAGATGCGCAGCG
GGGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP015866.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCGGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATCGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTATGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGTGGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGGAGCAGGCC
GGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACGGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGGCCTGGAGGCGCTGC
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGACCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTCAGGATGATGTCGACGGGCGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGGGAGGTGTCGCGCACCTCGCGCGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGGCCGTCGGCCAGGACCTG
GCCCTCGATGACCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTTCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGAT
GTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGCTCTTGATGAGCGGCACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGGCCTCGGAGAAGCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCGA
AGGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGA

CGGGTGGACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGAC
AGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTG
CGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTC
GCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCACC
GGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTT
GACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTACAGGTTCTCCAGAAGCGTCTGCGCGGCCTCACGCGTGGGCGGCT
CGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGC
CGAACTCCTCCAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTGAGCAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGACCA
TGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCCGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAGG
AGTCGAAGTAGACGCCGGGCGAACGCACCAGCTGCGAGACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAGCGGG
AAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAAGTC
ACGGTCCTTGCACTCGTCGATGCTGTTCTTCGGAGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGGG
AGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCACGCGTCGTTGC
CGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCG
GTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP027297.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCGCTGCGGCCGTCGAACAGCCGGGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTCCAGCAGACC
GGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGCGTGGCCACGTTGGTACGCGGCTCCACCTTGTCGGCGCCGATGTTCTGGAGGCGCTGCG
CCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCCTGGCTGGCGAGCCAACCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGG
ACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTTG
CCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGGACCAGCTGGTTCACGCCCGGGGGCAGCTCGTC
GCCCTCTTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCACCGTGCGGCACCTTCAGGGAGGTGTCGCGCACCTCGCGCGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAAGGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCGAGGACCTG
GCCCTCGATGACCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGAT
GTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTAAGGCGTCTCGATGAAGCCGA
ACGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGAC
GGGTGGACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGACA
GCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGC
GGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGTACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTCG
CCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCACCG
GCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTTG
ACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTCGGGTTGAAGTACAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGGCTC
GCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCC
GAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGACCA
TGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATCGAGG
AGTCGAAGTAGACGCCCGGCGAGCGGACGAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGCGG
GAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGCGCCGCGTACGTGAAGT
CGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGC
GAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCGTTG
CCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGC
GGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP039123.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT

CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCGCTGCGGCCGTCGAACAGCCGGGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTCCAGCAGACC
GGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGCGTGGCCACGTTGGTACGCGGCTCCACCTTGTCGGCGCCGATGTTCTGGAGGCGCTGCG
CCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCCTGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGG
ACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTTG
CCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGGACCAGCTGGTTCACGCCCGGGGGCAGCTCGTC
GCCCTCTTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCACCGTGCGGCACCTTCAGGGAGGTGTCGCGCACCTCGCGCGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAAGGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCGAGGACCTG
GCCCTCGATGACCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGAT
GTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTAAGGCGTCTCGATGAAGCCGA
ACGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGAC
GGGTGGACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGACA
GCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGC
GGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTCG
CCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCACCG
GCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTTG
ACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTCGGGTTGAAGTACAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGGCTC
GCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCC
GAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGACCA
TGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATCGAGG
AGTCGAAGTAGACGCCCGGCGAGCGGACGAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGCGG
GAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGCGCCGCGTACGTGAAGT
CGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGC
GAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCGTTG
CCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGC
GGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP042278.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACT-
GGTTTTACCACCCAGCGGCTGCTGGGTGATCATCGAGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGA
TGTACATGTACCCGACCGAGATCGGGTCCGGGAACGGCTCACCGCTGCGGCCGTCGAACAGCCGGGCCTTGCCGGACGGGAGCACCATGCGCTCG
CCGTCGCGGTTCGGGATGGTGTGCTCCAGCAGACCGGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGCGTGGCCACGTTGGTACGCGGCTC
CACCTTGTCGGCGCCGATGTTCTGGAGGCGCTGCGCCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCCTGGCTGGCGAGCCAGCCGAGGTGGA
TCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGGACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCC
TCGATCGGGTTGATCTTGGAGATGACGCCCTTGTTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACG
CGGACCAGCTGGTTCAC-
GCCGGGGGCAGCTCGTCGCCCTCTTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCACCGTGCGGCACCTTCAGGGAGGTGTCGCG
CACCTCGCGCGCCTTCTCACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGAT
GTCGCCGGCGACGACCTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTG
ATCTCCTCGGGGCCGAGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGA
CAGGATGATCGCGTCCTCGTAGTTGTGACCCTCCCACGGCATGAAGGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGG
ACCGTCGGCGAGGACCTGGCCCTCGATGACCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTT
GGCCAGGCGGTACGTGATGTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGAC
CACGTCGCCGGCGTCGGCGGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGATTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGT
TCGCGCCCATGAGGGCACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATG
TAGTCGACGTCGTCACCGGGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGTTGAGCGT
GGCGTTGGCCTGCGCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTAAG

GCGTCTCGATGAAGCCGAACGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCAC
ATGCGGCCGTAGTGCGACGGGTGGACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCT
TGTGCGTCAGCCCCGACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGG
ATGTTGATCAGGGTCTGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACC
TGGTTCTGGATCAGCTCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCG
ACCGTCTCGGTCTCACCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCC
GCGCCGAGCTTCTTGTTGACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTCGGGTTGAAGTACAGGTTCTCCAGCAGCGTCTGCGCGGCC
TCACGCGTGGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCG
CATGGACTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTC
GATGCGGACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGA
CGTCTTGTCGATCGAGGAGTCGAAGTAGACGCCCGGCGAGCGGACGAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCT
TGTTGGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGC
GCCGCGTACGTGAAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGA
GAAGTCCTCGATCGGCGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGC
CTTCCAGGCGTCGTTGCCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGA
GATGCGCAGCGGGGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGA
CTCAC

>gb|CP043958.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGATGGAGATCGGGTCCGGGAA
CGGCTCACCGGAGCGGCCGTCGAACATCCTCGCCTTACCGGTCGGCTGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGAC
CGGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCCGGGGGCGACGGAGTCGGCGCCGATGACCTGCAGACGCTT
CGCCCAGTCGTCGCCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGCACCTGTCCCGGGTTCATTCGGGACG
GGACGCCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACACCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCG
TCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTC
TCACCGAAGATCGCGCGGAGCAGCCGCTCCTCGGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGAC
CTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGA
GCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCC
TCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACC
TGGCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGCGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCGGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCCAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCCC
CGGGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGTTCTCGGTGAAGCGCATGTCGTCGTCGAGCGTGGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCCTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACGACACCGTCGGTGACCCGGCGGTACGGCGTCTCCACGAAAC
CGAACGCGTTGACGCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTG
CGAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCGCGGGACAGACCACCGGGACCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCC
GACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGT
CTGGGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACACGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAG
CTCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCGTGCCGTTGTCGGCGGCCGTCTCGGTCTC
GCCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGACGTCCTCGACGGTCAGGACGCCCGCGTCGAGCGGAGCCTCCGCACCCAGCTTCTT
GTTGACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGCCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCCAGGATCTGCTCGGTGGTCCAGCCGAGTGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATACGCACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCCGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTGGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAA
GTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGG
GGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCGTTCTCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGTTG
CCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAGAGATGCGCAGCGGGGC
AGTGCTGGCGGCGGGATTCGTATTCGAGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP043957.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGATGGAGATCGGGTCCGGGAA
CGGCTCACCGGAGCGGCCGTCGAACATCCTCGCCTTACCGGTCGGCTGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGAC
CGGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCCGGGGGCGACGGAGTCGGCGCCGATGACCTGCAGACGCTT
CGCCCAGTCGTCGCCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGCACCTGTCCCGGGTTCATTCGGGACG
GGACGCCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACACCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCG
TCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTC
TCACCGAAGATCGCGCGGAGCAGCCGCTCCTCGGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGAC
CTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGA
GCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCC
TCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACC
TGGCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGCGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCGGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCCAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCCC
CGGGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGTTCTCGGTGAAGCGCATGTCGTCGTCGAGCGTGGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCCTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACGACACCGTCGGTGACCCGGCGGTACGGCGTCTCCACGAAAC
CGAACGCGTTGACGCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTG
CGAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCGCGGGACAGACCACCGGGACCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCC
GACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGT
CTGGGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACACGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAG
CTCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCGTGCCGTTGTCGGCGGCCGTCTCGGTCTC
GCCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGACGTCCTCGACGGTCAGGACGCCCGCGTCGAGCGGAGCCTCCGCACCCAGCTTCTT
GTTGACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGCCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCCAGGATCTGCTCGGTGGTCCAGCCGAGTGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATACGCACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCCGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTGGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAA
GTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGG
GGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCGTTCTCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGTTG
CCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAGAGATGCGCAGCGGGGC
AGTGCTGGCGGCGGGATTCGTATTCGAGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP043956.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGATGGAGATCGGGTCCGGGAA
CGGCTCACCGGAGCGGCCGTCGAACATCCTCGCCTTACCGGTCGGCTGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGAC
CGGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCCGGGGGCGACGGAGTCGGCGCCGATGACCTGCAGACGCTT
CGCCCAGTCGTCGCCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGCACCTGTCCCGGGTTCATTCGGGACG
GGACGCCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACACCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCG
TCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTC
TCACCGAAGATCGCGCGGAGCAGCCGCTCCTCGGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGAC
CTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGA
GCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCC
TCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACC
TGGCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGCGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCGGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCCAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCCC
CGGGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGTTCTCGGTGAAGCGCATGTCGTCGTCGAGCGTGGCGTTGGCCTGCGC

GATGACGAAGCGGTCCTCCTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACGACACCGTCGGTGACCCGGCGGTACGGCGTCTCCACGAAAC
CGAACGCGTTGACGCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTG
CGAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCGCGGGACAGACCACCGGGACCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCC
GACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGT
CTGGGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACACGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAG
CTCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCGTGCCGTTGTCGGCGGCCGTCTCGGTCTC
GCCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGACGTCCTCGACGGTCAGGACGCCCGCGTCGAGCGGAGCCTCCGCACCCAGCTTCTT
GTTGACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGCCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCCAGGATCTGCTCGGTGGTCCAGCCGAGTGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATACGCACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCCGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTGGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAA
GTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGG
GGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCGTTCTCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGTTG
CCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAGAGATGCGCAGCGGGGC
AGTGCTGGCGGCGGGATTCGTATTCGAGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>emb|FN554889.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGATGGAGATCGGGTCCGGGAA
CGGCTCACCGGAGCGGCCGTCGAACATCCTCGCCTTACCGGTCGGCTGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGAC
CGGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCCGGGGGCGACGGAGTCGGCGCCGATGACCTGCAGACGCTT
CGCCCAGTCGTCGCCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGCACCTGTCCCGGGTTCATTCGGGACG
GGACGCCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACACCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCG
TCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTC
TCACCGAAGATCGCGCGGAGCAGCCGCTCCTCGGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGAC
CTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGA
GCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCC
TCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACC
TGGCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGCGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCGGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCCAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCCC
CGGGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGTTCTCGGTGAAGCGCATGTCGTCGTCGAGCGTGGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCCTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACGACACCGTCGGTGACCCGGCGGTACGGCGTCTCCACGAAAC
CGAACGCGTTGACGCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTG
CGAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGACAGACCACCGGGACCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCC
GACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGT
CTGGGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACACGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAG
CTCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCGTGCCGTTGTCGGCAGCCGTCTCGGTCTC
GCCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGACGTCCTCGACGGTCAGGACGCCCGCGTCGAGCGGAGCCTCCGCACCCAGCTTCTT
GTTGACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGGG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGCCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCCAGGATCTGCTCGGTGGTCCAGCCGAGTGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATACGCACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCCGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTGGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAA
GTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGG
GGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCGTTCTCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGTTG
CCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAGAGATGCGCAGCGGGGC
AGTGCTGGCGGCGGGATTCGTATTCGAGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|MK607459.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATACCGGGCTC
AGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACGC
GGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCGA
GTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCGCTGCGGCCGTCGAACAGCCGGGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTCCAGCAGACC
GGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGCGTGGCCACGTTGGTACGCGGCTCCACCTTGTCGGCGCCGATGTTCTGGAGGCGCTGCG
CCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCCTGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGG
ACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTTG
CCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGGACCAGCTGGTTCACGCCCGGGGGCAGCTCGTC
GCCCTCTTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCACCGTGCGGCACCTTCAGGGAGGTGTCGCGCACCTCGCGCGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAAGGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCGAGGACCTG
GCCCTCGATGACCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGAT
GTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGGC
GGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGGGACTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGGCCTCGGCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTAAGGCGTCTCGATGAAGCCGA
ACGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGAC
GGGTGGACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGACA
GCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGC
GGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTCG
CCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGACCGTCTCGGTCTCACCG
GCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTTG
ACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTCGGGTTGAAGTACAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGGCTC
GCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCC
GAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTGAGCAGGACGGTAACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGACCA
TGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATCGAGG
AGTCGAAGTAGACGCCCGGCGAGCGGACGAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGCGG
GAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGCGCCGCGTACGTGAAGT
CGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGC
GAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGACTCGACGCGAGCCTTCCAGGCGTCGTTG
CCGAGGAGCCAGTCGAAGCTCTCGGTTTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGC
GGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAA----------

>gb|CP034279.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATACCGAGCTCTTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGAGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAACGCGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAAC
GGCTCGCCGGAGCGGCCGTCGAACAGCCGGGCCTTGCCGGAGGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGAC
CGGCCAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTGGTGCCGGGCTCGACCTGGTCGGCACCGATGGCCTGCAGGCGCTGC
GCCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACGCCCTTGT
TGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGT
CGCCCTCCTCGCGGTCGAAGACGCGCACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCT
CGCCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACC
TCGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCCGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCACCGTTCTCGGTGGCGGGGCCGTCGGCGAGGACCT
GGCCCTCGATCACCCGGTCGCCCTCGGAGACGACGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGAGAACTTGGCGATGCGGTACGTG
GTGTACGTGCCGTCGTCGTTGGCGACGGTGACGTAGTCGGCCGAGACCTCCTGGATGACACCGTCCTTCTCCGCCTTGATGACGTCGCCGGCGTCG
GTGGCACAGCGGTACTCCATGCCGGTGCCGACCAGCGGCGCCTCGC-
CTTGATCAGCGGAACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGCACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGA
CCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCGGGCGCGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGCACGCGG

GCCTCGGCGAACCGGAGCTCGTCGGTCAGCGGCGCGTTGGCCTGCGCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTC
GGTGACCTGACCCTCGACGACCTTGCGGTAAGGCGTCTCGACGAAGCCGAACGCGTTGACGCGGCCGTACGAGGCGAGCGAACCGATCAGACCG
ATGTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGTCCGTAGTGGGACGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGACAG
ACCACCCGGGCCGAGCGCGGACAGACGACGCTTGTGCGTCAGCCCGGACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGA
AGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGCACG
ACGCGCTCCATCCGGGCCAGACCCGTGCGGACCTGGTTCTGAATGAGCTCGCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTC
GGTCTCGACGACGATCTGCGTGCCGTTCTCGCCGATCGTCTCGGTCTCGCCCGCGTGGAGCTTGACCAGGTACTTGATCGTGGCGATGACGTCGTC
GGTGGTGAGTACACCGGCGTCCAGCGGCTCGTCGGCGCCGAGCTTCTTGTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTT
GAAGTAGAGGTTCTCGAGCAGCGTCTGAGCGGCCTCACGGGTCGGCGGCTCGCCCGGGCGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCT
GGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTGGAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTGA
GGAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGAC
GGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGAGTCGAAGTAGACACCCGGCGAGCGGACGAGCTGCGACACCA
CGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCT
CGTTGTTGGTGAACTCGGCCGTGACGAAGAGCGGCGCGGCGTACGTGAAGTCGCGCTCCTTGCACTCGTCGATCGAGTTCTTGGGGGGCTCGAAG
CGGTGGTCGCGGAAAGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGCGAGATCTCCTCGAAGATCTCCTCGAGACCGGACTTGGTGGGGAC
GTCCTGTCCGGACTCCAGAGCCGCCTCGACACGAGCCTTCCACGCGGCGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGGAGGTT
CGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTTGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCG
AGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP013142.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTCAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGCGCGTAGG
CGGCCCCGTACGCCTCGAGTGCCCACACCTCCATCTCGCCGAACCGCTGGCCACCGAACTGAGCCTTACCGCCCAGCGGCTGCTGGGTGATCATGG
AGTACGGACCGGTCGAACGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCCGAGCGGCCGTCGAACAGCCGCGCCTTGCCGCTCGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGACC
CGCGAGCTCGTCCTCGCGCGCACCGTCGAACACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGACCTGGAGCCGCTCGG
CCCACTCCTCGGCGAGGCCGGAGACGTCCCAGCCACGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGCACCTGTCCCGGGTTCATTCGGGACGGC
ACGCCGAGCGGGTTGAGGATGATGTCGACCGGGGTCCCGTCCTCGAGGAACGGCATGTCCTCGATGGGCAGGATCTTGGAGATGACGCCCTTGTT
GCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTGACACCGGGGGGAAGCTCGT
CGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCT
CACCGAAGATCGCGCGCAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCACCCTTCGGGGTCACCTTGCCGACGAGGATGTCGCCGGCGATGACC
TCGGCGCCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCGAGGACCT
GGCCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGAGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGCCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCCGCCTTGACCACGTCACCGGCGTCGA
CGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGGACTTAATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCA
CGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACCTCGTCGCC
GGGGACGTAGTCGACCTCGCCGCCGCGGCGGCGGACCAGGATGCGGTTCTCGACGAAGCGCATGTCGTCGCCGAGCGGCGCGTTGGCCTGCGCG
ATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCG
AACGGGTTGATCCGCCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGA
GGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGAC
AGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTG
CGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTC
GCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGATGATCGTCCGGCCCGACTCGCCGACCGTCTCGGTCTCACC
GGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCGAGCGGCTGGTCGCCGCCGAGCTTCTTGT
TCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGGGTGGGGGGC
TCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCG
CCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTCAGGAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGAC
CATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGA
GGAGTCGAAGTAGACGCCGGGCGAGCGCACCAGCTGGGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAGCG
GGAAGTCGCCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGCGGGGCGGCGTACGTGAAG
TCGCGCTCCTTGCACTCGTCGATGGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGG
GGAGATCTCTTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCGTGCCCGTTCTCGAGAGCCTCCTCGACGCGAGCCTTCCACGCGTCGTT
GCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGG
CGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP054939.1|

TCAGACCTCTTCGACGCTACTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAG
GCGGCGCCGTACGCCTCGAGTGCCCACACCTCCATCTCGCCGAACCGCTGGCCACCGAACTGAGCCTTACCACCCAGCGGCTGCTGGGTGATCATG
GAGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGA
ACGGCTCACCCGAGCGGCCGTCGTACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGA
CCCGCGAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGACCTGGAGCCGCTC
GGCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GCACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCGAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACACCCTTG
TTGCCGTGCCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGCACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGGAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTACCGACGAGGATGTCGCCGGCGATGA
CCTCGGCGCCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCG
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGGGCCATCTCGCCGTTCTGCGTGGCGGGACCGTCGGCGAGGAC
CTGGCCGGTGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACG
TGATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCGGCCTTGACCACGTCGCCGGCGT
CGACCGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGG
GCGCGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACCTCGTC
ACCGGGGACGTAGTCGACCTCGCCGCCGCGGCGGCGGACCAGGATGCGGTTCTCGACGAAGCGCATCTCGTCGTTGAGCGGCGCGTTGGCCTGCG
CGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAG
CCGAACGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCCGGCGTCTCGATCGGGCACATGCGGCCGTAGTG
CGACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCC
GACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGT
CTGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCA
GCTCGCCGACGTTGCGGATACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGATGATCGAGCGGCCGGAGCCGCCGACCGTCTCGGTC
TCACCGGCGTGCAGCTGCACCAGGTACTTGATGGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCCCGCCGAGCTT
CTTGTTCACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCCAGCAGCGTCTGCGCGGCCTCGCGGGTCGG
GGGCTCGCCCGGGCGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGT
ACTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTGAGCAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACA
CCGACCATGTCGCGCTTGTCGATCTCCATCTCGAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCG
ATGCTGGAGTCGAAGTAGACACCGGGCGAACGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCAT
GAGCGGGAAGTCGCCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTAC
GTGAAGTCGCGCTCCTTGCACTCGTCGATAGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTC
GATCGGGGAGATCTCTTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCGTGTCCGCTCTCGAGAGCCTCCTCGACCCGACTCTGCCAGGC
GGTGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTCGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAGGAAATGCGCA
GCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP023697.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAAC
GGCTCGCCGGAACGGCCGTCGAACAGCCTCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGGCC
GGCCAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGACCTGCAGGCGCTCCG
CCCACTCCTCGGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGCACCTGTCCCGGGTTCATCCGGGACGGC
ACGCCGAGGGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGT
CGCCCTCCTCGCGGTCGAAGACGCGCACGCCGATGATCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGGACCTCGCGGGCCTTCT
CACCGAAGATCGCGCGCAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCGCCCTTGGGCGTCACCTTGCCGACGAGGATGTCGCCGGCGACGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGGACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGGCCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCGAGGACCT
GGCCCTCGATGACGCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCACCGGCGTCG
GCGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGGACCTGATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGG
CACGGTTGGCGTCGTCGTGCTCGAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACCTCGTCG
GGGCTGACGTAGTCGACCTCGCCGCCCTTGCGGCGGACCAGCACGCGCGCCTCGGCGAACCGCATGTCGTCACTCAGCGGGGCGTTGGCCTGCGC

GATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACGTCGTCGGTGACCTGGCCGTCGATGACCTTGCGGTACGGGGTCTCGACGAAAC
CGAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGC
GAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGCGCGGACAGACGACGCTTGTGCGTCAGCCCGG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGCACGACGCGCTCCATCCGGGCCAGACCCGTGCGGACCTGGTTCTGGATGAGC
TCGCCGACGCTGCGCAGACGGCGGTTGCCGAAGTGGTCGATGTCGTCCGTCTCGACGACGATCGTCTGACCGCTGTCACCGACGGTCTCGATCTCG
CCGGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGACGTCCTCGACGGTCAGCACGCCCGCGTCGAGCGGCGCGTCGGCACCCAGCTTCTTG
TTGACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGGGTGGGGGG
CTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCTTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTC
GCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTGAGCAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGACGCGGACGCCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGCT
GGAGTCGAAGTAGACACCCGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAGCG
GGAAGTCGCCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGCGCCGCGTACGTGAAG
TCGCGGTCCTTGCACTCGTCGATGGAGTTCTT-
GGAGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGCGAGATCTCCTCGAAGATCTCTTCCAGACCGGA
CTTGGTGGGGACGTCCTGACCGTTCTCCAGAGCGGCCTCGACGCGAGCCTTCCAGGCGTCATTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAG
CGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAAGAGATGCGCAGCGGGGCGGTGCTGGCGGCGTTGTTCGTATTCGCGGTCG
AGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP043959.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTCAGGCACAGGGACTGCATCTCCTTGATGAGGACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGCGCGTAGG
CGGCCCCGTACGCCTCGAGTGCCCACACCTCCATCTCGCCGAACCGCTGGCCACCGAACTGAGCCTTACCACCCAGCGGCTGCTGGGTGATCATGG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTCAGGATGTACATGTAACCGACCGAGATCGGCTCGGGGAAC
GGCTCACCGGAGCGGCCGTCGTACAGCCGCGCCTTACCGGTCGGGAGCACCATGCGCTCTCCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGACC
CGCGAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGACCTGGAGCCGCTCGG
CCCACTCCTCGGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGCACCTGTCCCGGGTTCATTCGGGACGGC
ACGCCGAGCGGGTTGAGGATGATGTCGACCGGGGTCCCGTCCTCGAGGAACGGCATGTCCTCGATGGGCAGGATCTTGGAGATGACGCCCTTGTT
GCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGT
CGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCT
CACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGACGACC
TCGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGGCCGTCGGCGAGGACCT
GGCCCTCGACGATCCGGTCACCCTCGGAGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGAGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCCGCCTTGACCACGTCACCGGCGTCGA
CGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGGACTTGATCAGCGGCACGGCCTGCCGCATCATGTTCGCGCCCATGAGGGCA
CGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACCTCGTCGCC
GGGGACGTAGTCGACCTCGCCGCCGCGGCGGCGGACCAGGATGCGGTTCTCGACGAAGCGCATGTCGTCGCTCAGCGGCGCGTTGGCCTGGGCG
ATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCCACGACCTTGCGGTACGGCGTCTCGATGAAGCCG
AACGGGTTGATCCGCCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCAGGGGTCTCGATCGGGCACATGCGGCCGTAGTGCGA
GGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGAC
AGCGGGTTGTTCTGGTCCATGAACTGAGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTG
CGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTC
GCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGATGATCGAGCGGCCCGACTCGCCGACCGTCTCGGTCTCAC
CGGCGTGCAGCTTCACCAGGTACTTGATCGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCGAGCGGCTGGTCGCCGCCGAGCTTCTTG
TTGACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGGGTGGGGGG
CTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTC
GCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTCAGGAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACGCCGGGCGAGCGGACCAGCTGGGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAG
CGGGAAGTCGCCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGCGGGGCGGCGTACGTGA
AGTCGCGGTCCTTGCACTCGTCGATGGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGATC
GGGGAGATCTCTTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCGTGCCCGTTCTCGAGAGCCTCCTCGACGCGAGCCTTCCAGGCGTCG
TTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGG
GGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP041168.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCTTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGAATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAGCGGGCGTGGAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGATCGAGATCGGGTCCGGGAA
CGGCTCACCGGAGCGGCCGTCGAACAGCCTCGCCTTGCCGGTGGGCTGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGGAGCAGAC
CGGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCCGGGGGCGACGGAGTCGGCGCCGATCGCCTGGAGGCGCTG
CGCCCAGTCCTCGGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GGACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACACCCTTG
TTGCCGTGCCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCG
TCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTC
TCACCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGATGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCACCGTTCTCGGTGGCCGCGCCGTCGGCGAGGACCT
GGCCCTCGATGACGCGGTCGCCCTCGTCGATGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCAC
CGGGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGGACCCGGGCCTCCTCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCG
ATGACGAAGCGGTCCTCCTCGTCGGCCGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCCACGAAACCG
AACGCGTTGACGCGACCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCG
AGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCGA
CAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCT
GCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCACGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCT
CGCCGACGCTGCGCAGGCGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATGGTCGTGCCGCTGTCGCCGACCGTCTCGGTCTCG
CCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCCGCGTCGAGCGGAGCCTCCGCACCCAGCTTCTTG
TTGACCTTGTAACGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTACAGGTTCTCCAGAAGCGTCTGCGCGGCCTCGCGCGTGGGCGG
CTCGCCCGGGCGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATCGACTCGTACTC
GCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTCAGGAGCACGGTGACGGACTGCTTGCGCTTGCGGTCGATACGGACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCCGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTGGGTCATGAG
CGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTCACGAAGAGCGGGGCCGCGTACGTGA
AGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCG
GGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCTGACTCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGT
TCCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAGAGATGCGCAGCGGG
GCTGTGCTGGCGGCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP023407.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAG
GCGGCGCCGTACGCCTCGAGTGCCCACACCTCCATCTCGCCGAACCGCTGGCCACCGAACTGAGCCTTACCACCCAGCGGCTGCTGGGTGATCATG
GAGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGA
ACGGCTCACCCGAGCGGCCGTCGTACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGA
CCCGCGAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGACCTGGAGCCGCTC
GGCCCACTCCTCCGCGAGGCCGGAAACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GCACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCGAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACACCCTTG
TTGCCGTGCCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGCACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGGAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGATGA
CCTCGGCGCCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCG
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGGGCCATCTCGCCGTTCTGCGTGGCGGGACCGTCGGCGAGGAC
CTGGCCGGTGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACG
TGATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCGGCCTTGACCACGTCGCCGGCGT
CGACCGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGG
GCGCGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACCTCGTC
ACCGGGGACGTAGTCGACCTCGCCGCCGCGGCGGCGGACCAGGATGCGGTTCTCGACGAAGCGCATCTCGTCGTTGAGCGGCGCGTTGGCCTGCG

CGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAG
CCGAACGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCCGGCGTTTCGATCGGGCACATGCGGCCGTAGTG
CGACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCC
GACAGCGGGTTGTTCTGGCCCATGAACTGGGACAACTGGCTGGTGCCGAAGAATTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGT
CTGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCA
GCTCGCCGACGTTGCGGATACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGATGATCGAGCGGCCGGAGCCGCCGACCGTCTCGGTC
TCACCGGCGTGCAGCTGCACCAGGTACTTGATGGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCCCGCCGAGCTT
CTTGTTCACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCCAGCAGCGTCTGCGCGGCCTCGCGGGTCGG
GGGCTCGCCCGGGCGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGT
ACTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTGAGCAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACA
CCGACCATGTCGCGCTTGTCGATCTCCATCTCGAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCG
ATGCTGGAGTCGAAGTAGACACCGGGCGAACGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCAT
GAGCGGGAAGTCGCCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTAC
GTGAAGTCGCGCTCCTTGCACTCGTCGATGGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTC
GATCGGGGAGATCTCTTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCGTGTCCGCTCTCGAGAGCCTCCTCGACCCGACTCTGCCAGGC
GGTGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTCGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAGGAAATGCGCA
GCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP032427.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCTTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGAATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGATCGAGATCGGGTCCGGGAAC
GGCTCACCGGAGCGGCCGTCGAACAGCCTCGCCTTGCCGGTGGGCTGCACCATCCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGGAGCAGACC
GGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCCGGGGGCGACGGAGTCGGCGCCGATCGCCTGGAGGCGCTGC
GCCCAGTCCTCGGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACACCCTTGTT
GCCGTGCCGACCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGATGACCTC
GGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCT
TGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCG
TAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCACCGTTCTCGGTGGCCGCGCCGTCGGCGAGGACCTGG
CCCTCGATGACGCGGTCGCCCTCGTCGATGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGAT
GTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGAC
GGCGGAGCGGTACTCCATGCCTGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGGACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGGACCCGGGCCTCCTCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGAT
GACGAAGCGGTCCTCCTCGTCGGCCGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCCACGAAACCGAA
CGCGTTGACGCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGAG
GGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCGACA
GCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGC
GGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTC
GCCGACGCTGCGCAGACGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATGTTCGTGCCGCTGTCACCGACCGTCTCGGTCTCGCC
CGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCCGCGTCGAGCGGAGCTTCCGCACCCAGCTTCTTGTT
GACCTTGTAACGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTACAGGTTCTCCAGAAGCGTCTGCGCGGCCTCGCGCGTGGGCGGCT
CGCCCGGGCGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATCGACTCGTACTCGC
CGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTCAGGAGTACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGACC
ATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCCGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAG
GAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTGGGTCATGAGCG
GGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTCACGAAGAGCGGGGCCGCGTACGTGAAG
TCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGG
GAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCTGACTCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGTTCC
CGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAGAGATGCGCAGCGGGGCT
GTGCTGGCGGCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP042594.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGCGCGTAGG
CGGCGCCGTACGCCTCGAGTGCCCACACCTCCATCTCGCCGAACCGCTGGCCACCGAACTGAGCCTTACCGCCCAGCGGCTGCTGGGTGATCATGG
AGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAA
CGGCTCACCGGAGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGGAGG
CCCGCGAGCTCGTCCTCGCGCGCACCGTCGAACACCGGGGTGGCGACGTTGGTGCCGGGCTCCACCTTGTCGGCGCCGATGGCCTGGAGGCGCTC
GGCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCCCGGCTGGCGAGCCAGCCCAGGTGGATCTCCAGAACCTGTCCCGGGTTCATTCGGGACG
GCACGCCCAGCGGGTTGAGGATGATGTCGACCGGGGTCCCGTCCTCGAGGAACGGCATGTCCTCGATGGGCAGGATCTTGGAGATGACGCCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGCAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCGCCCTTCGGGGTCACCTTGCCGACGAGGATGTCGCCGGCGATGA
CCTCGGCGCCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCG
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCACCGTTCTCGGTGGCCGGACCGTCGGCGAGGAC
CTGGCCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCGAGGCGGTACG
TGATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGT
CGACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGG
GCACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACCTCGTC
ACCGGGGACGTAGTCGACCTCGCCGCCGCGGCGGCGGACCAGGATACGGTTCTCGACGAAGCGCATGTCGTCGTTGAGCGGCGCGTTGGCCTGC
GCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTAAGGCGTCTCGATGAA
GCCGAACGGGTTGATCCGCCCGTAGGAGGCGAGCGAACCGATCAGGCCGATGTTCGGGCCTTCCGGCGTCTCGATCGGGCACATGCGGCCGTAGT
GCGAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCC
CGACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGG
TCTGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCA
GCTCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGATGATCGTCCGGCCCGACTCGCCGACCGTCTCGGTCT
CACCGGCGTGCAGCTTCACCAGGTACTTGATCGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCGAGCGGCTGGTCGCCGCCGAGCTTC
TTGTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGG
GGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGT
ACTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTCAGGAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGCACAC
CGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGAGGTCTTGTCGA
TGGAGGAGTCGAAGTAGACGCCGGGCGAGCGCACCAGCTGGGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATG
AGCGGGAAGTCGCCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGCGGGGCGGCGTACGT
GAAGTCGCGCTCCTTGCACTCGTCGATGGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGA
TCGGGGAGATCTCTTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCGTGCCCGTTCTCGAGAGCCTCCTCGACGCGAGCCTTCCACGCGT
CGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGC
GGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP032229.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCGCTGGACAGCACCTCCACGTTGAGGCAGAGCGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAGG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAACGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGATGGAGATCGGGTCCGGGAA
CGGCTCACCGGAGCGGCCGTCGAACAGCCGGGCCTTGCCGGTCGGCTGCACCATCCGCTCACCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGAC
CCGCCAGCTCGTCCTCGCGCGCGCCGTCGAACACCGGGGTGGCGACGTTCGTGCCGGGGTCGACCTGGTCGGCGCCGATCACCTGGAGGCGCTCG
GCCCAGTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACGCCCTTGT
TGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGGGCGACGTAGACACGGACGAGCTGGTTGACGCCCGGGGGCAGCTCG
TCGCCCTCCTCGCGGTCGAAGACGCGGACACCGATGATCTTGCCGGTCTCGCCGTGCGGCACCTTCAGGGAGGTGTCGCGGACCTCGCGGGCCTTC
TCACCGAAGATGGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCCGCGGTGACC
TCGGCGCCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCCAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAAGGCCACGAGCAGGTTCTTGCCGAGGGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCT
GGCCTTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGACGTGCCCTGGTTGGAGCGCGAGAACTTGTGCAGGCGGTACGTGA
TGTACGTGCCGTCGTCGTTGGCGGTGGTGACGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGA
CGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGCA
CGGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATCGCGGTGGCGACGGACACCATCTGGCGCGGCGACACGTCCATGTAGTCGACGTCGTCACC
GGGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGGACGCGGGCCTCCTCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGA

TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGGGTCTCGACGAAACCGA
ACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGA
GGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCGAC
AGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTG
CGGCGTGATCGCCTCGACGTCCTGAGTGGTCATGCGCTCACGCACGACACGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTC
GCCGACGCTGCGCAGGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGACTGACCGCTGTCGCCGACCGTCTCGGTCTCGC
CGGCGTGCAGCTTGACCAGGTACTTGATCGTCGAGATGACGTCCTCGGTGGTGAGCACACCCGCGTTGAGCGGGGCGTCGGCACCGAGCTTCTTG
TTGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGCGG
CTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTC
GCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAGCCGAGCGCCTTGAGCAGAACGGTGACCGACTGCTTGCGCTTGCGGTCGATGCGGACACCGAC
CATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCCGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGA
GGAGTCGAAGTAGACACCCGGCGAGCGGACGAGCTGGGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGCGCGGCGTACGTGAA
GTCGCGGTCCTTGCACTCGTCGATGGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCG
GGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGAACGTCCTGACCGTTCTCGAGAGCCTTCTCGACGCGAGCCTTCCAGGCGTCATT
GCCGAGCAGCCAGTCGAAGCTCTCGGTCTGGAGCGCAAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGG
GCGGTGCTGGCGGCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP040941.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGCGCGTAGG
CGGCGCCGTACGCCTCGAGTGCCCACACCTCCATCTCGCCGAACCGCTGGCCACCGAACTGAGCCTTACCGCCCAGCGGCTGCTGGGTGATCATGG
AGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAA
CGGCTCACCCGAGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGGAGG
CCCGCGAGCTCGTCCTCGCGCGCACCGTCGAACACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGACCTGGAGGCGCTC
GGCCCACTCCTCGGCGAGGCCGGAGACGTCCCAGCCCCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGAACCTGTCCCGGGTTCATTCGGGACG
GCACGCCCAGCGGGTTGAGGATGATGTCGACCGGGGTCCCGTCCTCGAGGAACGGCATGTCCTCGATGGGCAGGATCTTGGAGATGACGCCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGCAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCGCCCTTCGGGGTCACCTTGCCGACGAGGATGTCGCCGGCGATGA
CCTCGGCGCCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCG
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTGGGTGGCCGGGCCGTCGGCGAGGAC
CTGGCCCTCGACGACACGGTCGCCCTCGGAGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGAGAACTTGGCCAGGCGGTACG
TGATGTACGTGCCGTCGTCGTTGGCCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCCGCCTTGACCACGTCGCCGGCGTC
GACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGGACTTGATCAGCGGCACGGCCTGGCGCATCATGTTCGCGCCCATGAGG
GCGCGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACCTCGTC
ACCGGGGACGTAGTCGACCTCGCCGCCGCGGCGGCGGACCAGGATGCGGTTCTCGACGAAGCGCATGTCGTCGCCGAGCGGCGCGTTGGCCTGC
GCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAA
GCCGAACGGGTTGATCCGCCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCCGGCGTCTCGATCGGGCACATGCGGCCGTAGT
GCGAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCC
CGACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGG
TCTGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCA
GCTCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGATGATCGTCCGGCCCGACTCGCCGACCGTCTCGGTCT
CACCGGCGTGCAGCTTCACCAGGTACTTGATCGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCGAGCGGCTGGTCGCCGCCGAGCTTC
TTGTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGG
GGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGT
ACTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTCAGGAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGCACAC
CGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCGCCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGAGGTCTTGTCGA
TGGAGGAGTCGAAGTAGACGCCGGGCGAGCGCACCAGCTGGGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATG
AGCGGGAAGTCGCCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGCGGGGCGGCGTACGT
GAAGTCGCGCTCCTTGCACTCGTCGATGGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGA
TCGGGGAGATCTCTTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCGTGCCCGTTCTCGAGAGCCTCCTCGACGCGAGCCTTCCACGCGT
CGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGC
GGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP029624.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGCGCGTAGGC
GGCGCCGTACGCCTCGAGTGCCCACACCTCCATCTCGCCGAACCGCTGGCCACCGAACTGAGCCTTACCGCCCAGCGGCTGCTGGGTGATCATGGA
GTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCCGAGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGCAGGACCATGCGGTCGCCGTCGCGGTTCGGGACGGTGTGCTCCAGCAGACC
GGCGAGCTCGTCCTCGCGGGCGCCGTCGAAGACCGGGGGTGGCCACGTTGGTACGGGGCTCCACGCGGTCCGCGCCGATCGCCTGCAGGCGCTGC
GCCCACTCGTCGGCGAGGCCGGAGACGTCCCAGCCCTGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGAACCTGTCCCGGGTTCATTCGGGACGG
CACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTCCCGTCCTCGAGGAACGGCATGTCCTCGATGGGCAGGATCTTGGAGATGACGCCCTTGT
TGCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCG
TCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTC
TCACCGAAGATCGCGCGCAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCACCCTTCGGGGTCACCTTGCCGACGAGGATGTCGCCGGCGATGAC
CTCGGCGCCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGA
GCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCC
TCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTGGGTGGCCGGACCGTCGGCGAGGACC
TGGCCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGAGAACTTGGCCAGGCGGTACGT
GATGTACGTGCCGTCGTCGTTGGCCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCCGCCTTGACCACGTCGCCGGCGTC
GACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGGACTTAATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGG
GCACGGTTGGCGTCGTCGTGCTCGAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACCTCGTC
GCCGGGGACGTAGTCGACCTCGCCGCCGCGGCGGCGGACCAGGATGCGGTTCTCGACGAAGCGCATGTCGTCGCCGAGCGGCGCGTTGGCCTGG
GCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAA
GCCGAACGGGTTGATCCGCCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGT
GCGAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCC
CGACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGG
TCTGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCA
GCTCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGATGATCGAGCGGCCCGACTCGCCGACCGTCTCGGTC
TCACCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCCCGCCGAGCTTC
TTGTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGG
GGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGT
ACTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTCAGGAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGCACAC
CGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGA
TGGAGGAGTCGAAGTAGACGCCGGGCGAGCGCACCAGCTGGGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATG
AGCGGGAAGTCGCCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGCGGGGCGGCGTACGT
GAAGTCGCGCTCCTTGCACTCGTCGATGGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGA
TCGGGGAGATCTCTTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCGTGCCCGTTCTCGAGAGCCTCCTCGACGCGAGCCTTCCACGCGT
CGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGC
GGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP060828.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCTTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCGCCGAATCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACGGAGATCGGGTCCGGGAA
CGGCTCACCGCTACGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTGGGGATGGTGTGGTTCAGCAGAC
CCGCCAGCTCGTCCTCACGGGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGGGCGACCTGGTCGGCGCCGATGACCTGGAGGCGCTG
CGCCCACTCGTCGGCGAGCCCGGAGACGTCCCAGCCACGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GCACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAAATAACACCCTTG
TTGCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTGACGCCCGGGGGAAGCTCG
TCCCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGATCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCGCGGGCCTTC
TCACCGAAGATCGCGCGCAGCAGCCGCTCCTCGGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGACGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCCAGGACCTCTTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGGACGAGGCGCTGCGACAGGATGATCGCGTCTT
CGTAGTTGTGGCCCTCCCACGGCATGAACGCGACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCT
GGCCGGTGATGATCCGGTCGCCCTCGTCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGCGAGAACTTGTGCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTCGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGGGCTTCTCGGCCTTGACGACGTCACCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACCTCGTCAC
CGGGGACGTAGTCGACCTCGCCACCACGACGGCGGACCAGGACGCGGTTCTCGGTGAAGCGCAGTTCGTCGTTGAGCGTGGCGTTGGCCTGCGCG

ATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGGGTCTCGACGAAACC
GAACGCGTTGACGCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGC
GAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCGGGACCAAGGGCCGACAGACGGCGCTTGTGGGTGAGCCCC
GACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGT
CTGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACACGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAG
CTCGCCGACGCTGCGCAGACGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCTGACCGCTGTCGCCGACCGTCTCGGTCTC
GCCCGCGTGCAGCTTGACCAGGTACTTGATCGTCGCGATGACGTCCTCGACGGTCAGCACACCGGCGTCGAGCGGGGCGTCCGCGGCGAGCTTCT
TGTTGACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTACAGGTTCTCGAGGAGCGTCTGCGCGGCCTCACGGGTCGGC
GGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTA
CTCGCCGAACTCCTCCAGGATCTGCTCGGTGGTGTAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGGACGC
CGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGA
TGGAGGAGTCGAAGTAGACACCGGGCGAGCGCACGAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATG
AGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGCGCGGCGTACGT
GAAGTCGCGCTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGAT
CGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGCTCTCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGT
GTTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGGAGCGCGAGCAGGTTCGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCG
GGGCAGTGCTGGCGGCGTTGTTCGTATTCGCGGTCGAGGCGGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP031969.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCTTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGAATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACATCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTATGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAGCGGGCGTGGAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCCTCGCCTTACCGGTCGGCTGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGAC
CCGCGAGCTCGTCCTCACGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCCGGGGGCGACGGCGTCGGCGCCGATGACCTGGAGGCGCTG
CGCCCAGTCCTCGGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GCACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTCCCGTCCTCGAGGAACGGCATGTCCTCGATGGGCAGGATCTTGGAGATGACACCCTTG
TTGCCGTGTCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGGACCAGCTGGTTCACGCCCGGCGGCAGCTCG
TCGCCCTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTCGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTC
TCGCCGAAGATCGCACGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGACGAC
CTCGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGA
GCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCC
TCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACC
TGGCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCCC
CGGGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACCCGGGACTCCTCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCG
ATGACGAAGCGGTCCTCCTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACGACACCGCCGGTGACCTTGCGGTACGGCGTCTCCACGAAACCG
AACGCGTTGACGCGGCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCG
AGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCGA
CAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCT
GCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCACGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCT
CGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGAGGTGCCGCTGTCACCGGTGGTCTCGGTCTCG
CCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCCGCGTCCAGCGGAGCGTCCGCACCCAGCTTCTTG
TTGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCCAGAAGAGTCTGGGCGGCCTCACGCGTCGGGGG
CTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATCGACTCGTACTC
GCCGAACTCCTCCAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTCAGGAGCACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCCGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCCGCCGTAACGAAGAGCGGGGCGGCGTACGTGAA
GTCGCGCTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGG
GGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCAAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGTT
CCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAGAGATGCGCAGCGGGG
CGGTGCTGGCGGCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>emb|LT670819.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGAGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGGAGCTCCTGGAGGGCGTACGC
GGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGACCGCCGAACTGCGCCTTACCACCCAGCGGCTGCTGGGTGATCATCGA
GTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGATGGAGATCGGGTCCGGGAAC
GGCTCACCGGAGCGGCCGTCGAACAGCCGCGCCTTACCGGTCGGCTGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGACC
CGCCAGCTCGTCCTCACGCGCGCCGTCGAAGACCGGGGGTGGCGACGTTGGTGCCGGGGGCGACCGAGTCGGCGCCGATCACCTGCAGACGCTGC
GCCCACTCCTCGCCGAGCCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATACGGGACGG
CACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTCCCGTCCTCGAGGAACGGCATGTCCTCGATGGGCAGGATCTTGGAGATGACACCCTTGT
TGCCGTGGCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACGCGAACCAGCTGGTTCACGCCCGGCGGCAGCTCGT
CGCCCTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCT
CACCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACCT
CGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTG
GCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGAT
GTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGAC
GGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCACCG
GCGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACGCGGGCCTCGGTGAACTGCATGTCGTCGTTCAGCGTGGCGTTGGCCTGCGCGAT
GACGTAGCGGTCCTCCTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCCACGAAACCGAA
CGCGTTGACGCGGCCGTAGGAGGCGAGCGAACCGATCAGGCCGATGTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGTCCGTAGTGGGAC
GGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGCGCCGACAGACGGCGCTTGTGGGTGAGACCCGACA
GCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGC
GGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACTCGCTCCATACGCGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTCG
CCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCTGGCCGTTGTCCGCGGCCGTCTCGGTCTCGCCG
GCGTGCAGCTTGACCAGGTACTTGATCGTCGAGATGACGTCCTCGACGGTCAGGACACCCGCGTCCAGCGGAGCGTCCGCACCCAGCTTCTTGTTG
ACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGGCTC
GCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCC
GAACTCCTCGAGGATCTGCTCGGTCGTCCAGCCGAGAGCCTTCAGGAGAACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGACCAT
GTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGAAGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGA
GTCGAAGTAGACACCCGGCGAACGGACCAGCTGGGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAGCGGGA
AGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTCGTGAACTCGGCCGTCACGAAGAGCGGGGCCGCGTACGTGAAGTCAC
GGTCCTTGCACTCGTCGATGCTGTTCTTAGGCGGCTCGAAACGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGGGAG
ATCTCCTCGAAGATCTCCTCCAGACCGGACTTCGTGGGGACGTCCTGACCGGACTCCAGAGCCGCCTCGACCCGACTCTGCCAGGCGGTGTTCCCGA
GCAGCCAGTCAAAGCTCTCGGTCTGCAACGCGAGCAGGTTGGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTG
CTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP009438.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATACCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATCGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCAAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGAC
CCGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCCGGGGGCGACCTGGTCGGCGCCGATGGCCTGCAGGCGCTG
GGCCCACTCGTCGGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGCACCTGTCCCGGGTTCATTCGGGACG
GCACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTCCCGTCCTCGAGGAACGGCATGTCCTCGATGGGCAGGATCTTGGAGATGACACCCTTG
TTGCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGGGCGACGTAGACGCGCACGAGCTGGTTCACGCCCGGGGGCAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGCACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGATGAC
CTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGA
GCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCC
TCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGGGCCATCTCACCGTTCTGGGTGGCCGGGCCGTCGGCGAGGACC
TGGCCCTCGATGACGCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGT
GATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGGCCTTCTCGGCCTTGACGACGTCGCCGGCGTC
GACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGC-
CTGGATGAGCGGAACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGCACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCG
ACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCCTCGGGGCTGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACGCG

AGCCTCGGAGAAGCGCATGTCGTCGCCGAGCGTGGCGTTGGCCTGCGCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGT
CGGTGACCTGGCCGTCGACGACCTTGCGGTACGGGGTCTCGACGAAACCGAAGGCGTTGACGCGGCCGTAGGAGGCGAGCGAGCCGATCAGACC
GATGTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGGCCGTAGTGCGACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAG
AGACCACCCGGGCCGAGCGCGGACAGACGACGCTTGTGCGTCAGCCCGGACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCC
GAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCGGCGTGATCGCCTCGACGTCCTGCGTCGTCATGCGCTCGCGGA
CGACTCGCTCCATACGCGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGT
CCGTCTCGACGACGACCGTCTCGCCGCTGTCGCCGACCGTCTCGGTCTCGCCGGCGTGCAGCTTCACCAGGTACTTGATCGTCGCGATGATGTCCTC
GACGGTCAGGATGCCCGCGTCCAGCGGGGTGTCCGTACCCAGCTTCTTGTTGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTT
GAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGGGTGGGCGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCT
GGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAGCCGAGCGCCTTGA
GCAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACG
GGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGCTGGAGTCGAAGTAGACACCCGGCGAACGGACCAGCTGCGACACCACG
ACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCG
TTGTTGGTGAACTCGGCCGTGACGAAGAGCGGGGCGGCGTACGTGAAGTCGCGCTCCTTGCACTCGTCGATGGAGTTCTTCGGCGGCTCGAAGCG
GTGGTCGCGGAAGGTCAGCGACATCGACCCGCTGAAGTCCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGT
CCTGACCGTTCTCGAGGGCCTCCTCGACGCGAGCCTTCCACGCGTCGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGGAGCGCGAGAAGGTTCG
GAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAG
GCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP047020.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATACCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGCGACTGCATCTCCTTGATGAGGACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAGGCCTCGAGGGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGCGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGTCCGGTCGAACGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCGGAGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGACC
CGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTACCGGG-
GCGACCTGGTCGGCGCCGATGGCCTGCAGGCGCTGGGCCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGT
GGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGGACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATG
TCCTCGATCGGCAGGATCTTGGAGATGACACCCTTGTTGCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAG
ACGCGGACCAGCTGGTTCACGCCCGGCGGCAGTTCGTCGCCCTCCTCGCGGTCGAAGACGCGCACGCCGATGACCTTGCCGGTCTCGCCGTGCGGC
ACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTCACCGAAGATCGCGCGCAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCACCCTTGGGG
GTCACCTTGCCGACGAGGATGTCACCGGCGATGACCTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGAC
GTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGT
CCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGGGCCATCT
CGCCGTTCTGGGTGGCCGGACCGTCGGCGAGGACCTGGCCCTCGATGACGCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCT
GGTTGGACCGGGAGAACTTGGCCAGACGGTACGTGATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACA
CCGTCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGACCGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGTGCCTCGGACTTGATCAGCGGC
ACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCACGGTTGGCGTCGTCGTCGCTCGAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTG
GCGCGGCGAGACGTCCATGTAGTCCACGTCCTCGGGGGCGACGTAGTCGACCTCGCCACCACGGCGGCGGACCAGGACACGGGCCTCGGCGAAC
CGCATCTCGTCGTTCAGCGGCGCGTTGGCCTGCGCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCACACCG
TCGTCGACCTTGCGGTACGGCGTCTCGACGAAACCGAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTC
GGGCGTCTCGATCGGGCACATACGGCCGTAGTGCGACGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCG
AGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGAC-
GTACCGAAGAACTCCTTGATCGACGCCACGACCGGGCGGATGTTGATCAGGGTCTGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTC
GCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTCGCCGACGCTGCGGATGCGGCGGTTGCCGAAGTGGTCGA
TGTCGTCGGTCTCGACCATGATCGTGCGACCGGACTCGGCGACCGTCTCGGTCTCGCCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGA
CGTCGTCGGTGGTGAGCACGCCCGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCT
TGGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCGCGCGTGGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGC
GTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAG
CGCCTTGAGCAGCACGGTGACCGACTGCTTGCGCTTGCGGTCGATACGGACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCC
CGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGAGTCGAAGTAGACACCGGGGGAACGGACCAGCTGAG
ACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCAC
CGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCGGCGTACGTGAAGTCGCGCTCCTTGCACTCGTCGATGGAGTTCTTCGGCGGC
TCGAAGCGGTGGTCCCGGAACGTCAGGGACATCGACCCGGAGAAGTCCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGT
GGGGACGTCCTGACCGTTCTCCAGAGCCTCCTCGACCCGACTCTGCCAAGCGGTGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAG
CAGGTTCGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAT
TGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP017316.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCCGGGATGCCGGGCT
CCGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACATCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGAGCCTTACCGCCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGGCGCGCCTTGCCGGACGGCAGCATCAGGCGGTCGCCGTCGCGGTTCGGGAGGGTGTGCTCCAGCAGG
CCGGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCGGGGCTCGACGCGGTCGGCGCCGATGGACTGGAGCCGCT
GGGCCCAGTCGTCGGCGAGCCCGGAGACGTCCCAGCCCTGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGAC
GGCACACCGAGCGGGTTGAGGATGATGTCGACGGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACGCCCTT
GTTGCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTAGACGCGGACCAGCTGGTTCACGCCCGGGGGCAGCT
CGTCGCCCTCCTCGCGGTCGAAGACGCGCACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCT
TCTCACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTCGGGGTGACCTTGCCGACCAGGATGTCGCCGGCGACG
ACCTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCC
GAGCTTCGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGT
CCTCGTAGTTGTGGCCCTCCCACGGCATGAACGCCACGAGGAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGGCCGTCGGCGAGGA
CCTGGCCCTCGATGACCCGGTCGCCCTCGGCGACGACGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTAC
GTGATGTACGTGCCGTCGTCGTTGGCGACGGTGATGTAGTCCGCGGAGAGCTCCTGGACGACACCGTCCTTCTCGGCCTTGACCACGTCGCCGGCG
TCGACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGAGCG-
GATCAGCGGCACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGCGCGGTTGGCGTCGTCGTGCTCGAGGAAGGGGATCATGGCGGTCGCGACC
GACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCCCCGGCGACGTAGTCGACCTCGCCGCCGCGGCGGCGCACCAGGACGCGGTT
CTCGGCGAAGCGCATGTCCTCGGTCAGGCCGGCGTTGGCCTGGGCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGG
TGACCCGGCCGTCGATGACCTTGCGGTACGGCGTCTCGACGAAGCCGAAGGCGTTGACACGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGAT
GTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGTCCGTAGTGGGACGGGTGCACGTCTCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGAC
CACCCGGGCCGAGCGCGGACAGACGACGCTTGTGCGTCAGCCCGGACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAG
AACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGGACGAC
GCGCTCCATCCGCGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTCGCCGACGTTGCGCAGGCGGCGGTTGCCGAAGTGGTCGATGTCGTCCG
TCTCGACGACGATCGTCGAGCCGTTCTCACCGACGGTCTCGGTCTCGCCCGCGTGCAGCTTGACCAGGTACTTGATCGTCGCGATGACGTCGTCGAC
GGTCAGGACGCCGGCGTCCAGCGGCTCGTCGGCGCCGAGCTTCTTGTTGACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTCGGGTTGAA
GTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTCGGGGGCTCGCCCGGACGGAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGC
CCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTGCAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTGAGGA
GGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGAGGGG
ATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGAGTCGAAGTAGACACCCGGCGAGCGGACGAGCTGCGACACGACGA
CACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGT
TGTTGGTGAACTCGGCCGTGACGAAGAGCGGCGCGGCGTACGTGAAGTCGCGCTCCTTGCACTCGTCGATGGAGTTCTTCGGGGGCTCGAAGCGG
TGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTC
CTGTCCGTTCTCGAGAGCCGCCTCGACACGAGCCTTCCACGCGGCGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGGAGGTTCGG
AACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTTGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGG
CGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP063808.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCACGCATTTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGAGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACGC
GGCGCCGTATGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCGA
GTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGATCGAGATCGGGTCCGGGAAC
GGCTCACCGGAGCGGCCGTCGAACAGCCTCGCCTTACCGGTCGGCTGCACCATGCGCTCGCCGTCGCGGTTGGGGATGGTGTGGTTCAGCAGACC
CGCGAGCTCGTCCTCGCGCGCACCGTCGAACACCGGGGGTGGCGACGTTGGTGCCGGGGGCGACCTGGTCGGCGCCGATCACCTGGAGGCGCTGC
GCCCAGTCCTCGGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGCACCTGTCCCGGGTTCATTCGGGACGG
CACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTCCCGTCCTCGAGGAACGGCATGTCCTCGATGGGCAGGATCTTGGAGATGACACCCTTGT
TGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTGACACCGGGGGGAAGCTCGT
CGCCCTCCTCGCGGTCGAAGACGCGCACACCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCT
CACCGAAGATCGCGCGCAGCAGGCGCTCCTCGGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGATGACCT
CGGCACCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGGGCCATCTCACCGTTCTCGGTGGCCGCGCCATCGGCGAGGACCTG
GCCCTCGATGATGCGGTCGCCCTCGTTGATGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGAT
GTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGGCTTCTCGGCCTTGACCACGTCACCGGCGTCGAC

GGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACCCGGGCCTCCTCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGAT
GACGAAGCGGTCCTCCTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCCACGAAACCGAA
CGCGTTGACGCGGCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGAG
GGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCGACA
GCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGC
GGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTC
GCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATGTTCGTGCCGCTGTCGCCGACCGTCTCGGTCTCGCC
CGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCCGCGTCCAGCGGAGCCTCCGCACCCAGCTTCTTGTTG
ACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGGGTCGGGGGCTC
GCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATCGACTCGTACTCGCC
GAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTCAGGAGCACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGACCA
TGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCCGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAGG
AGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTGGGTCATGAGCGG
GAAATCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTCACGAAGAGCGGGGCCGCGTACGTGAAGT
CGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGG
GAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCTGACTCCAGAGCCTCCTCGACCCGACTCTGCCAAGCGGTGTTCC
CGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAGAGATGCGCAGCGGGGCG
GTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGG-TCCTTCCGAGGGCTCGGACTCAC

>gb|CP023696.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCCGGGATGCCGGGCT
CCGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACATCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGAGCCTTACCGCCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGGCGCGCCTTGCCGGACGGCAGCATCAGGCGGTCGCCGTCGCGGTTCGGGAGGGTGTGCTCCAGCAGG
CCGGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCGGGGCTCGACGCGGTCGGCGCCGATGGACTGGAGCCGCT
GGGCCCAGTCGTCGGCGAGCCCGGAGACGTCCCAGCCCTGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGAC
GGCACACCGAGCGGGTTGAGGATGATGTCGACGGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACGCCCTT
GTTGCCGTGGCGGCCGGCGAGCTTGTCTCCGTCGGTGATCTTGCGCTTCTGCGCCACGTAGACGCGGACCAGCTGGTTCACGCCCGGGGGCAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGCACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTCGGGGTGACCTTGCCGACCAGGATGTCGCCGGCGACGA
CCTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCG
AGCTTCGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGGCCCTCCCACGGCATGAAGGCCACGAGGAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGGCCGTCGGCGAGGA
CCTGGCCCTCGATGACCCGGTCGCCCTCGGCGACGACGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTAC
GTGATGTACGTGCCGTCGTCGTTGGCGACGGTGATGTAGTCCGCGGAGAGCTCCTGGACGACACCGTCCTTCTCGGCCTTGACCACGTCGCCGGCG
TCGACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGAGCG-
GATCAGCGGCACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGCGCGGTTGGCGTCGTCGTGCTCGAGGAAGGGGATCATGGCGGTCGCGACC
GACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCCCCGGCGACGTAGTCGACCTCGCCGCCGCGGCGGCGCACCAGGACGCGGTT
CTCGGCGAAGCGCATGTCCTCGGTCAGGCCGGCGTTGGCCTGGGCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGG
TGACCCGGCCGTCGATGACCTTGCGGTACGGCGTCTCGACGAAGCCGAAGGCGTTGACGCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGAT
GTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGTCCGTAGTGGGACGGGTGCACGTCTCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGAC
CACCCGGGCCGAGCGCGGACAGACGACGCTTGTGCGTCAGCCCGGACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAG
AACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGGACGAC
GCGCTCCATCCGCGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTCGCCGACGTTGCGCAGGCGGCGGTTGCCGAAGTGGTCGATGTCGTCCG
TCTCGACGACGATCGTCGAGCCGTTCTCACCGACGGTCTCGGTCTCGCCCGCGTGCAGCTTGACCAGGTACTTGATCGTCGCGATGACGTCGTCGAC
GGTCAGGACGCCGGCGTCCAGCGGCTCGTCGGCGCCGAGCTTCTTGTTGACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTCGGGTTGAA
GTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTCGGGGGCTCGCCCGGACGGAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGC
CCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTGCAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTGAGGA
GGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGAGGGG
ATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGAGTCGAAGTAGACACCCGGCGAGCGGACGAGCTGCGACACGACGA
CACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGT
TGTTGGTGAACTCGGCCGTGACGAAGAGCGGCGCGGCGTACGTGAAGTCGCGCTCCTTGCACTCGTCGATGGAGTTCTTCGGGGGCTCGAAGCGG
TGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGCGAGATCTCCTCGAAGATCTCTTCCAGACCGGACTTGGTGGGGACGTC
CTGTCCGTTCTCGAGAGCCGCCTCGACACGAGCCTTCCACGCGGCGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGGAGGTTCGG

AACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTTGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGG
CGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP023689.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGCGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACGC
GGCGCCGTAAGCCTCGAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATGCT
GTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAAGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAACG
GCTCACCCGAGCGGCCGTCGAACAGCGGCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATCGTGTGCTCCAGCAGACCG
GCGAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTGGACGCTGTCGGCGCCGATCGCCTGGAGGCGCTGCG
CCCACTCCTCGCCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGC
ACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACACCCTTGTT
GCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACACCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACAAGGATGTCTCCGGCGACGACCTC
GGCACCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCT
TGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCG
TAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTGG
CCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGATG
TACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGACG
GCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGGACTTGATGAGGGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCACCG
GCGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGCACGCGGTTCTCGGTGAAGCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGAT
GACGAAGCGGTCTTCCTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACCTGACCCTCGACGACCTTGCGGTACGGCGTCTCCACGAAGCCGAA
CGCGTTGACGCGGCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTGGGGCCTTCGGGCGTCTCGATCGGGCACATACGGCCGTAGTGAGAC
GGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCGACA
GCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGC
GGCGTGATCGCCTCGACGTCCTGAGTCGTCATGCGCTCACGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTCG
CCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCTGGCCGTTGTCGGCGGCCGTCTCGGTCTCACCG
GCGTGCAGCTTGACCAGGTACTTGATCGTCGAGATGACGTCCTCGACGGTCAGGACGCCTGCGTCGAGCGGAGCCTCCGCACCCAGCTTCTTGTTG
ACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCGAGGAGCGTCTGCGCCGCCTCGCGCGTCGGGGGCTC
GCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCC
GAACTCCTCCAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTGAGCAGAACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGACCAT
GTCGCGCTTGTCGATCTCCAGCTCCAGCCAGGCGCCCCGGGACGGGATGACCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAGG
AGTCGAAGTAGACACCGGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAACGTGCCCTTGTGCGTCATGAGCGGG
AAGTCGCCCATGAAGACGGTCTGAGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCTGCCGTCACGAAGAGCGGGGCCGCGTACGTGAAGTCG
CGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGGGA
GATCTCCTCGAAGATCTCCTCGAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGCCTCGACCCGACTCTGCCAGGCGGTGTTCCC
GAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCG
GTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|KX503594.1|

GACCTCTTCGACGCTGCTCGGCTCGCGCCGTGACAGGTCGATGCCGAGCTCTTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTCGAT
GGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTCGG
GAATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACACGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACGCGG
CGCCATAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCGAGT
ACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGATGGAGATCGGGTCCGGGAACGG
CTCACCGGAGCGGCCGTCGAACATCCTCGCCTTACCGGTCGGCTGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGGTTCAGCAGACCCGC
GAGCTCGTCCTCACGGGCGCCGTCGAACACCGGGGTGGCGACGTTGGTGCCCGGGGCGACCTGGTCGGCGCCGATGACCTGCAGGCGCTGCGCC
CAGTCCTCGCCGAGCCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGGAC
ACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACACCCTTGTTGCC
GTGTCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGAACCAGCTGGTTCACGCCCGGCGGCAGCTCGTCGCC
CTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTCACC
GAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCTCGG
CACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCTTG
GTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCGTA

GTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTGGCC
CTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGATGTA
CGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACGACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGACGGC
GGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCACGGT
TGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCACCGGGG
ACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACGCGGTTCTCGGTGAAACGCATGTTGTCGTCGAGCGTGGCGTTGGCCTGCGCGATGAC
GAATCGGTCCTCTTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACGACGTCGCCGTCGACCTTGCGGTACGGCGTCTCCACGAAACCGAACGC
GTTGACACGGCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGTCCGTAGTGGGACGGG
TGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCGACAGCG
GGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCGGC
GTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACTCGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAGCTCGCCG
ACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGAGCCGCCGCTGTCGCCGACCGTCTCGGTCTCGCCCGC
GTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCCGCGTCGAGCGGAGCCTCCGCACCCAGCTTCTTGTTGAC
CTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCGAGAAGCGTCTGCGCGGCCTCACGCGTGGGCGGCTCGC
CCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCG
AACTCCTCCAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATACGGACACCGACCAT
GTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATCGAGGA
GTCGAAGTAGACACCCGGCGAACGGACCAGCTGCGACACGACGACGCGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTGGGTCATGAGCGGG
AAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCACCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGCGGGGCCGCGTAGGTGAAGTC
GCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGAG
AGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCGTTCTCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGTTGCC
GAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATCTTTGCGAAAGAGATGCGCAGCGGGGCG
GTGCTGGCGGCGGGGTTCGTATTCGAGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP060825.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATACCGAGCTCCTCCGCGGCGCGGAAGACGTCCTCGTCCGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCAGAGCGACTGCATTTCCTTGATGAGCACCTTGAAGGACTCGGGAATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATCGTCAGCAGCTCCTGGAGTGCGTAGG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAATCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCGA
GTACGGACCGGTCGAACGGGCGTGGAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAAC
GGCTCGCCGGAGCGGCCGTCGAACAGGCGCGCCTTGCCGGTCGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGGAGCAGGC
CGGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCCGGCGGCGACGTCGTCCGCGCCGATGGCCTGGAGACGCTG
CGCCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATGCGGGACG
GGACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACGCCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGGGCCACGTAGACGCGGACCAGCTGGTTGACGCCCGGGGGAAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGCACGCCGATGACCTTGCCGGTCTCGCCGTGGGGCACCTTCAGCGAGGTGTCGCGGACCTCGCGCGCCTT
CTCGCCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTGAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGA
CCTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCC
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGGCCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCATTCTCCGTGGCGGGACCGTCGGCGAGGAC
CTGGCCCTCGATGATCCGGTCGCCCTCGTCGACGACGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGAGAACTTGTGCAGGCGGTACGT
GGTGTACGTGCCGTCGTCGTTGGCGATGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCGGCCTTGAGCACGTCGCCGGCGTC
GGTGGCGCAGCGGTACTCCATGCCGGTGCCCACGAGCGGCGCCTCGGACTTGATGAGCGGCACGGCCTGCCTCATCATGTTCGCGCCCATCAGGG
CGCGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTG-
CCGGGCACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGCACGCGGGACTCGGCGAACCGGAACTCGTCGGTCAGCGGCGCGTTGGCCTGCG
CGATGACGAAGCGGTCCTCCTCGTCGGCGGTGAGGTAGTCGACCTCGTCGCCGACGACACCGTCGACCACCTTGCGGTACGGCGTCTCGATGAAGC
CGAACGCGTTGATCCGGCCGTACGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGACACATGCGGCCGTAGTGC
GACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTGAGACGACGCTTGTGCGTCAGGCCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAG
CTCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCCGTCTCGACCATGATCGAGCGGCCGGACTCGCCGACCGTCTCGGTCTC
ACCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTTCGCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGAGCGGCCTCGCGGGTCGGCG
GCTCGCCCGGGCGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT
CGCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAGCCGAGCGCCTTGAGCAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGAGGGGATGATCTTGGCGGTATAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACGCCCGGCGAGCGCACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTCGTGAACTCGGCGGTGACGAAGAGCGGGGCGGCGTACGTGAA
GTCGCGCTCCTTGCACTCGTCGATGGAGTTCTTCGGGGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGATCG

GGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTCGTGGGGACGTCCTGTCCGCTTTCGAGAGCCGCCTCGACACGAGCCTTCCATGCGGCATT
GCCGAGCAGCCAGTCGAAGCTCTCGGTCTGGAGCGCAAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGG
GCAGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>dbj|AP019621.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCAGGGATGCCGGGCT
CGGGAATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCATAAGCCTCAAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAAC
GGCTCGCCGGAGCGGCCGTCGAACAGCGGCGCCTTGCCGGTCGGGAGCACCATGCGCTCTCCGTCGCGGTTCGGGATCGTGTGCTGCAGCAGACC
CGCGAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTGGTGCCCGGGGCGACCTGGTCGGCCTCGATGGCCTGCAGGCGCTGC
GCCCAGTCGTCGGCGAGTCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCGAGGGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACACCCTTGT
TGCCGTGGCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGAACCAGCTGGTTCACGCCCGGCGGCAGCTCGT
CGCCCTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGCACCTCGCGCGCCTTCT
CACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACC
TCGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTCGCCGGGCCGTCGGCCAGAACCT
GGCCGGCGATGACCCGGGCGCCCTCGTCCACGACAACCTTCTGGTTGACGGACGTGCCCTGGTTGGAGCGGGAGAACTTGTGCAGGCGGTACGTG
GTGTACGTGCCGTCGTCGTTGGCGGTGGTGACGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCCGACTTGAGCACGTCGCCGGCGTCG
ACGGCGCAGCGGTACTCCATGCCGGTGCCGACGAGCGGCGCCTCGGACTTGATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCAC
CGGGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGGACGCGGGTCTCCTCGAACCGGAAGTCGTCGCCGAGCGGCGCGTTGGCCTGCGC
GATGACGAATCGGTCCTCTTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACGACACCGTCGGTGACCCGGCGGTACGGCGTCTCGACGAAACC
GAACGCGTTGACGCGGCCGTAAGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGGCCGTAGTGC
GACGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGAGTGAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGAGTCGTCATGCGCTCGCGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGC
TCGCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGAGGTGCCGTTGTCGCCAACGGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATCGACGAGATGATGTCCTCGACGGTCAGGATCCCGGCGTCCAGCGGCGCCTCGCCGCCCAGCTTCTTG
TTGACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCAAGAAGCGTCTGCGCGGCCTCGCGGGTCGGGGG
CTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTC
GCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTCAGGAGAACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCGCCCCGGGAGGGGATGACCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCCGGCGAGCGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAA
GTCGCGGTCCTTGCACTCGTCGATGGAGTTCTTCGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGGGACATCGACCCGGAGAAGTCCTCGATCG
GAGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCGCTGTCCAGAGCCGCCTCGACACGAGCCTTCCACGCGTCGTT
GCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCAAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGG
CGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP022438.1|

TCAGACCTCTTCGACGCTGCTCGGCTCACGCCGGGACAGGTCGATACCGAGCTCCTCCGCGGCGCGGAAGACGTCCTCGTCCGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGAGAGCACCTCCACGTTGAGGCAGAGCGACTGCATTTCCTTGATGAGCACCTTGAAGGACTCGGGAATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGTTCCTGGAGGGCGTAGG
CCGCTCCGTACGCCTCGAGTGCCCACACTTCCATCTCTCCGAACCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCGA
GTACGGACCGGTCGAACGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAAC
GGCTCGCCGGAGCGGCCGTCGAACAGACGCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATCGTGTGCTGCAGCAGGCC
GGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGGTCGCGACGTTGGTGCCCGGCGCGACGTTGTCCGCGCCGATGGCCTGCAGGCGCTCCG
CCCACTCGTCCGCCAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCGGGGTTCATTCGGGACGGG
ACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTGTTG
CCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGGGCCACGTAGACGCGGACCAGCTGGTTGACGCCCGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGCACACCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCGCGCGCCTTCTC
ACCGAAGATCGCGCGGAGCAGACGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACCT
CGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCCGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC

GTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGGGCCATCTCACCGTTCTCGGTGGCGGGGCCGTCGGCGAGGACCT
GGCCCTCGATCACCCGGTCACCCTCGGAGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGAGAACTTGGCGATGCGGTACGTG
GTGTACGTGCCGTCGTCGTTGGCGACGGTGACGTAGTCGGCCGAGACCTCCTGGATGACACCGTCCTTCTCGGCCTTGATCACGTCACCGGCGTCG
GTGGCACAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGC-
CTTGATGAGCGGAACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGCACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGA
CCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACCTCGGC-
CGGGGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGCACGCGGGCCTCGGTGAAGCGGAGCTCGTCGTTCAGCGTGGCGTTCGCCTGCG
CGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGACCGTCGGTGACCTTGCGGTACGGGGTCTCGACGAAAC
CGAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCAGGGGTCTCGATCGGGCACATGCGGCCGTAGTG
CGACGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGACCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCC
GACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGT
CTGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGA
GCTCGCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTGGGTGCCGGACTCGCCGACCGTCTCGGTCT
CGCCGGCGTGCAGCTTGACGAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCACACCGGCGTCCAGCGGCTCGTCGGCGCCGAGCTTC
TTGTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGAGCGGCCTCACGCGTCGGC
GGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATCGACTCGTAC
TCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTGAGCAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATACGGACACC
GACCATGTCACGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGTGAAGATGTCCTTGTCGGACGTCTTGTCGAT
GGAGGAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGGGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATG
AGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGT
GAAGTCGCGCTCCTTGCACTCGTCGATCGAGTTCTTAGGGGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGA
TCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTCGTGGGGACGTCCTGTCCGCTTTCGAGAGCCGCCTCGACACGAGCCTTCCATGCGG
CATTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGC
GGGGCAGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>dbj|BA000030.4|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCAGGGATGCCGGGCT
CGGGAATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCATAAGCCTCAAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAAC
GGCTCGCCGGAGCGGCCGTCGAACAGCGGCGCCTTGCCGGTCGGGAGCACCATGCGCTCTCCGTCGCGGTTCGGGATCGTGTGCTGCAGCAGACC
CGCGAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTGGTGCCCGGGGCGACCTGGTCGGCCTCGATGGCCTGCAGGCGCTGC
GCCCAGTCGTCGGCGAGTCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCGAGGGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACACCCTTGT
TGCCGTGGCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGAACCAGCTGGTTCACGCCCGGCGGCAGCTCGT
CGCCCTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGCACCTCGCGCGCCTTCT
CACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACC
TCGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTCGCCGGGCCGTCGGCCAGAACCT
GGCCGGCGATGACCCGGGCGCCCTCGTCCACGACAACCTTCTGGTTGACGGACGTGCCCTGGTTGGAGCGGGAGAACTTGTGCAGGCGGTACGTG
GTGTACGTGCCGTCGTCGTTGGCGGTGGTGACGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCCGACTTGAGCACGTCGCCGGCGTCG
ACGGCGCAGCGGTACTCCATGCCGGTGCCGACGAGCGGCGCCTCGGACTTGATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCAC
CGGGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGGACGCGGGTCTCCTCGAACCGGAAGTCGTCGCCGAGCGGCGCGTTGGCCTGCGC
GATGACGAATCGGTCCTCTTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACGACACCGTCGGTGACCCGGCGGTACGGCGTCTCGACGAAACC
GAACGCGTTGACGCGGCCGTAAGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGGCCGTAGTGC
GACGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGAGTGAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGAGTCGTCATGCGCTCGCGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGC
TCGCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGAGGTGCCGTTGTCGCCAACGGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATCGACGAGATGATGTCCTCGACGGTCAGGATCCCGGCGTCCAGCGGCGCCTCGCCGCCCAGCTTCTTG
TTGACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCAAGAAGCGTCTGCGCGGCCTCGCGGGTCGGGGG
CTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTC
GCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTCAGGAGAACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCGCCCCGGGAGGGGATGACCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCCGGCGAGCGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAA

GTCGCGGTCCTTGCACTCGTCGATGGAGTTCTTCGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGGGACATCGACCCGGAGAAGTCCTCGATCG
GAGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCGCTGTCCAGAGCCGCCTCGACACGAGCCTTCCACGCGTCGTT
GCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCAAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGG
CGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP075611.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGCGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTAAGCCTCGAGCGCCCACACCTCCATCTCACCGAAGCGCTGACCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATGC
TGTACGGGCCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACGGAGATCGGGTCCGGGAA
AGGCTCACCCGAGCGACCGTCGAAGAGCCGCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCGCGGTTGGGGATGGTGTGCTGGAGCAGA
CCCGCGAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTGGACGCTGTCGGCGCCGATCGCCTGGAGGCGCT
GCGCCCACTCCTCGCCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGAC
GGCACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACACCCTT
GTTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACC-
GGGGGAAGCTCGTCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGAC
CTCACGGGCCTTCTCACCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACAAGGATGTC
TCCGGCGACGACCTCGGCACCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCT
CCTCGGGGCCGAGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGG
ATGATCGCGTCCTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGT
CGGCCAGGACCTGGCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCA
GGCGGTACGTGATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGT
CGCCGGCGTCGACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGGACTTGATGAGGGGCACGGCCTGACGCATCATGTTCGC
GCCCATGAGGGCGCGGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGT
CCACGTCGTCACCGGCGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGTTCTCGGTGAAGCGCATGTCGTCGTTGAGCGTGGCG
TTGGCCTGCGCGATGACGAAGCGGTCTTCCTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACCTGACCGTCGACGACCTTGCGGTACGGCGTC
TCCACGAAACCGAACGCGTTGACGCGGCCGTAGGAGGCGAGCGAACCGATCAGGCCGATGTTGGGGCCTTCGGGCGTCTCGATCGGGCACATGC
GGCCGTAGTGAGACGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTG
GGTGAGACCCGACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGT
TGATCAGGGTCTGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCACGGACGACGCGCTCCATACGAGCCAGACCCGTACGGACCTGGT
TCTGGATGAGCTCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCTGGCCGTTGTCGGCGGCC
GTCTCGGTCTCACCGGCGTGCAGCTTGACCAGGTACTTGATCGTCGAGATGACGTCCTCGACGGTCAGGACGCCCGCGTCGAGCGGAGCCTCCGCA
CCCAGCTTCTTGTTGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCGAGGAGCGTCTGCGCCGCCTCA
CGCGT-
GGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTC
GTACTCGCCGAACTCCTCAAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTGAGCAGAACGGTCACGGACTGCTTGCGCTTGCGGTCGATACGCAC
ACCGACCATGTCACGCTTGTCGATCTCCAGCTCCAGCCAGGCGCCCCGGGACGGGATGACCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTC
GATGGAGGAGTCGAAGTAGACACCGGGCGAACGGACAAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAACGTGCCCTTGTGCGTC
ATGAGCGGGAAGTCGCCCATGAAGACCGTCTGAGACTTGATCTCACCGGTCTCGTTGTTGGTGAACTCGGCCGTCACGAAGAGCGGGGCCGCATA
CGTGAAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTC
GATCGGGGAGATCTCCTCGAAGATCTCCTCGAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGCCTCGACCCGACTCTGCCAGGC
GGTGTTCCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCA
GCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP054936.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACATCCTCGTCGGTGTCGCGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGAGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAGG
CGGCGCCGTACGCCTCAAGTGCCCACACCTCCATCTCGCCGAACCGCTGGCCGCCGAACTGCGCCTTACCACCCAGCGGCTGCTGGGTGATCATGG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGTCGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGGAGCAGACC
CGCCAGCTCGTCCTCGCGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTGGACGCTGTCGGCGCCGATGACCTGCAGGCGCTGG
GCCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACGCCGAGCGGGTTCAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACACCCTTGT
TGCCGTGCCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCG
TCGCCCTCTTCGCGGTCGAAGAGGCGCACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTC
TCGCCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGAC

GTCGGCGCCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGA
GCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCC
TCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGGGCCATCTCGCCGTTCTCGGTGGCGGGACCGTCGGCGAGGACC
TGGCCGGCGATGACGTGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGCCGGTAGGT
GATGTAGGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGGCCTTCTCGGCCTTGACGACGTCGCCGGCGTC
GACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGGACGGCCTGGCGCATCATGTTCGCGCCCATGAGG
GCGCGGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTC
ACCGGCGACGTAGTCGACCTCGCCGCCGCGGCGGCGCACCAGGACGCGGTTCTCGGAGAAGCGCATGTCGTCGTTGAGGCCGGCGTTGGCCTGC
GCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTACGGGGTCTCGACGAA
ACCGAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGT
GCGAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGGGCCGACAGGCGGCGCTTGTGGGTGAGAC
CCGACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGG
GTCTGCGGCGTGATCGCCTCGACGTCCTGCGTGGTCATCCGCTCGCGGACGACGCGCTCCATACGGGCCAGGCCGGTGCGGACCTGGTTCTGGAT
GAGCTCGCCCACGCTGCGCAGACGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTGCC-
GCCGTTGTCACCGGCGGTCTCGGTCTCGCCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCGGCGTC
CAGCGGCGCGTCCGCGCCCAGCTTCTTGTTGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTGTCCAGCAG
CGTCTGCGCGGCCTCACGCGTGGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCT
CCAGGGTGGCGCGCATCGACTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGTTGGTCCAGCCGAGCGCCTTGAGGAGAACGGTGACGGACTGC
TTGCGCTTGCGGTCGATGCGGACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCGCCCCGGGACGGGATGATCTTGGCGGAGAAG
ATGTCCTTGTCGGACGTCTTGTCGATGGAGGAGTCGAAGTACACGCCCGGCGAACGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGAT
GACGAAGGTGCCCTTGTTGGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGT
GACGAAGAGCGGGGCGGCGAACGTGAAGTCGCGCTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAACGTCAGCG
ACATCGACCCGGAGAAGTCCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCC
GCCTCGACGCGAGCCTTCCACGCGTCGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTG
ATCTTTGCAAAGGAAATGCGCAGCGGGGCGGTGCTGGCGGCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCT
TCCGAGGGCTCGGACTCAC

>gb|CP015849.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCGGCGCGGAAGACATCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGCGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
AGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAGG
CGGCGCCGTACGCCTCAAGGGCCCACACCTCCATCTCGCCGAACCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATGG
AGTACGGGCCGGTCGAACGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAA
CGGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGTCGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGGAGCAGAC
CCGCCAGCTCGTCCTCCCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCCGGGGTCGACCTGGTCGGCGCCGATCGCCTGGAGGCGCTGC
GCCCACTCCTCCGCCAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATCCGGGACGG
GACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGGTTGATCTTGGAGATGACGCCCTTGTT
GCCGTGGCGGCCGGCGAGCTTGTCACCATCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGGACCAGCTGGTTCACGCCCGGCGGCAGCTCGT
CGCCCTCCTCGCGGTCGAAGAGGCGGACACCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGGACCTCACGGGCCTTCT
CGCCGAAGATCGCGCGCAGCAGCCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGGCCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCGAGGACCT
GGCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGAGAACTTGTGCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCGGCGGAGACCTCCTGCACCACACCGGCCTTCTCGGCCTTGACGACGTCACCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGC
GCGGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCAC
CGGCGACGTAGTCGACCTCGCCGCCGCGGCGGCGGACCAGGACGCGGTTCTCGGAGAAGCGCATGTCGTCGTTGAGGCCGGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGGGTCTCGACGAAAC
CGAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGC
GAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCCAGCGCCGACAGACGGCGCTTGTGGGTGAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAG
CTCGCCGACGCTGCGCAGACGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTGCC-
GCCGTTGTCCCCGGCGGTCTCGGTCTCACCGGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGAGGATGTCCTCGACGGTCAGGATGCCGGCGTC
CAGCGGCGCGTCCGCGCCCAGCTTCTTGTTGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCCAGCAG
CGTCTGCGCGGCCTCGCGCGTGGGGGGCTCGCCCGGTCGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTC
CAGGGTGGCGCGCATCGACTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGTTGGTCCAGCCGAGCGCCTTCAGCAGTACGGTGACCGACTGCTT
GCGCTTGCGGTCGATGCGGACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGAT

GTCCTTGTCGGACGTCTTGTCGATGCTCGAGTCGAAGTAGACACCGGGCGAGCGGACCAGCTGGGAGACGACGACACGCTCGGTGCCGTTGATGA
CGAAGGTGCCCTTGTTCGTCATGAGCGGGAAGTCGCCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGA
CGAAGAGCGGGGCGGCGTACGTGAAGTCGCGCTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGAC
ATCGACCCGGAGAAGTCCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCGAGGCCGGACTTGGTGGGGACGTCCTGACCGGATTCCAGAGCCGC
CTCGACGCGAGCCTTCCACGCGTCGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGAT
CTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTGGCAGCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTC
CGAGGGCTCGGACTCAC

>gb|CP056041.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGCGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAGGC
GGCGCCGTAAGCCTCGAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATGCT
GTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAAGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAACG
GCTCACCCGAGCGGCCGTCGAACAGCGGCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATCGTGTGCTCCAGCAGACCG
GCGAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTTCCGGGCTGGACGCTGTCGGCGCCGATCGCCTGGAGGCGCTGCGC
CCACTCCTCGCCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGCA
CACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACACCCTTGTTGC
CGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTCGC
CCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTCAC
CGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACAAGGATGTCTCCGGCGACGACCTCGG
CACCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCTTG
GTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCGTA
GTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTGGCC
CTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTCGCCAGGCGGTACGTGATGTA
CGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGACGGC
GGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGGACTTGATGAGGGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCACGG
TTGGCGTCGTCGTGCTCCAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCACCGGC
GACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGCACGCGGTTCTCGGTGAAGCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGATGA
CGAAGCGGTCTTCCTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCCTCGACGACCTTGCGGTACGGCGTCTCCACGAAACCGAACG
CGTTGACGCGGCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTGGGGCCTTCGGGCGTCTCGATCGGGCACATACGGCCGTAGTGAGACGG
GTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCGACAGC
GGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCGG
CGTGATCGCCTCGACGTCCTGAGTCGTCATGCGCTCACGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTCGCC
GACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCTGGCCGTTGTCGGCGGCCGTCTCGGTCTCACCGG
CGTGCAGCTTGACCAGGTACTTGATCGTCGAGATGACGTCCTCGACGGTCAGGACGCCCGCGTCGAGCGGAGCCTCCGCACCCAGCTTCTTGTTGA
CCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCGAGGAGCGTCTGCGCCGCCTCGCGCGTCGGGGGCTCG
CCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCG
AACTCCTCCAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTGAGCAGAACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGACCATG
TCGCGCTTGTCGATCTCCAGCTCCAGCCAGGCGCCCCGGGACGGGATGACCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGA
GTCGAAGTAGACACCGGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAACGTGCCCTTGTGCGTCATGAGCGGGA
AGTCGCCCATGAAGACGGTCTGAGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCTGCCGTCACGAAGAGCGGGGCCGCGTACGTGAAGTCGC
GGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGGGAG
ATCTCCTCGAAGATCTCCTCGAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGCCTCGACCCGACTCTGCCAGGCGGTGTTCCCG
AGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGT
GCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP047318.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGAGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACGC
GGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCGCCCAGCGGCTGCTGGGTGATCATCGA
GTACGGGCCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCGCTACGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGACC
CGCCAGCTCGTCCTCACGGGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGAGCGACACTGTCGGCGCCGATGGCCTGCAGACGCTGG
GCCCACTCGTCGGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
CACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCGAGGAAGGGCATGTCCTCGATCGGGAGGATCTTGGAGATGACACCCTTGT
TGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTGACACCGGGGGGAAGCTCG

TCGCCCTCCTCGCGGTCGAAGACGCGCACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTC
TCACCGAAGATCGCGCGGAGCAGGCGCTCCTCGGGCGTCAGCTCGGTCTCACCCTTCGGGGTGACCTTGCCGACGAGGATGTCACCGGCGATGAC
CTCGGCGCCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGA
GCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCC
TCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCACCGTTCTCGGTGGCGGGACCGTCGGCGAGGACC
TGGCCCTCGATGATGCGGTCGCCCTCGTGGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGT
GATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGGCCTTCTCGGCCTTGACGACGTCGCCGGCGTC
GACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGC-
CTTGATGAGCGGAACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGCACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGA
CCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCACCGGCGACGTAGTCGACCTCGCCGCCACGACGGCGGACGAGCACGCGG
TTCTCGGCGAAGCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCG
GTGACCTGGCCCTCGACGACCTTGCGGTACGGCGTCTCCACGAAACCGAACGCGTTGACGCGGCCGTAGGAGGCGAGCGAACCGATCAGACCGAT
GTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGGCCGTAGTGCGACGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGCGACAGAC
CACCGGGACCCAGCGCCGAGAGACGACGCTTGTGCGTCAGCCCGGACAGCGGGTTGTTCTGGTCCATGAACTGCGAGAGCTGGCTGGTGCCGAAG
AACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCCCGGACGAC
GCGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAGCTCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGG
TCTCGACGACGATGTTCGTGCCGTTGTCCCCGGCGGTCTCGGTCTCGCCGGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGAC
GGTCAGGATGCCCGCGTCCAGCGGGGCGTCCGCACCCAGCTTCTTGTTGACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTGGGGTTGAA
GTACAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGC
CCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTCCAGGATCTGCTCGGTGGTCCAACCGAGGGCCTTGAGGA
GGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCGCCCCGGGACGGGA
TGATCTTGGCAGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGAGTCGAAGTAGACACCGGGCGAACGGACGAGCTGCGACACCACGACA
CGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTG
TTGGTGAACTCGGCCGTGACGAAGAGCGGGGCCGCGTACGTGAAGTCGCGCTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAGCGGTG
GTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCT
GTCCGTTCTCCAGAGCCTCCTCGACCCGACTCTGCCAAGCGGTGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAA
CCTCGAGAGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCG
GCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP022310.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTTTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGCGCGTAGGC
GGCGCCGTACGCCTCGAGTGCCCACACCTCCATCTCGCCGAACCGCTGGCCACCGAACTGAGCCTTACCACCCAGCGGCTGCTGGGTGATCATGGA
GTACGGACCGGTCGAACGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTAGCCGACGGAGATCGGGTCCGGGAAC
GGCTCACCCGAGCGACCGTCGAACAGCCGCGCCTTGCCGGACGGGAGGACCATGCGGTCGCCGTCGCGGTTCGGGATGGTGTGCTCCAGCAGACC
GGCGAGCTCGTCCTCGCGGGCGCCGTCGAAGACCGGGGTGGCCACGTTGGTACGGGGCTCGACCCGGTCCGCGCCGATCGCCTGGAGGCGCTGC
GCCCACTCGTCGGCGAGGCCGGAGACGTCCCAGCCCTGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
CACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACGCCCTTGTT
GCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGGACCAGCTGGTTCACGCCCGGCGGCAGCTCGT
CGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCT
CACCGAAGATCGCGCGCAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCACCCTTCGGGGTCACCTTGCCGACGAGGATGTCGCCGGCGATGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTGGGTGGCCGGACCGTCGGCGAGGACCT
GGCCCTCGATGACGCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAATTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGGAGCGGATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGG
CACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCG
CCGGGGACGTAGTCGACCTCGCCGCCGCGGCGGCGGACCAGGATGCGGTTCTCGACGAAGCGCATCTCGTCGTTGAGCGTGGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCTTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCC
GAACGGGTTGATCCGGCCGTACGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGC
GAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGCGCGTTGAGACGACGCTTGTGCGTCAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCAG
CTCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACCATGATCGAGCGACCGGACTCGCCGACCGTCTCGGTCTC
ACCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCCCGCCGAGCTTCTT
GTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCGAGCAGCGTCTGCGCGGCCTCGCGGGTGGGCG
GCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACT

CGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTCAGGAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACGCCG
ACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATG
CTGGAGTCGAAGTAGACGCCCGGCGAGCGCACCAGCTGGGAGACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAG
CGGGAAGTCGCCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCGGCGTACGTG
AAGTCGCGCTCCTTGCACTCGTCGATGGAGTTCTTAGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGAT
CGGGGAGATCTCTTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGCCCGTTCTCGAGAGCCTCCTCGACGCGAGCCTTCCACGCGTC
GTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCG
GGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP034463.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACATCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATTTCCTTGATGAGCACCTTGAAGGACTCGGGAATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTATGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGAGCCTTACCGCCCAGCGGCTGCTGGGTGATCATGG
AGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAA
CGGCTCACCGCTACGGCCGTCGAACAGCCGCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGGTTCAGCAGAC
CCGCGAGCTCGTCCTCACGCGCACCGTCGAACACCGGGGTGGCGACGTTGGTGCCGGGGGCGACCTGGTCGGCACCGATGTTCTGGAGGCGCTGC
GCCCACTCGTCGGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGAACCTGTCCCGGGTTCATTCGGGACGG
CACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTCCCGTCCTCGAGGAACGGCATGTCCTCGATGGGCAGGATCTTGGAGATGACACCCTTGT
TGCCGTGGCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGAACCAGCTGGTTCACGCCCGGCGGCAGCTCGT
CGCCTTCCTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGGGAGGTGTCACGGACCTCGCGCGCCTTCT
CACCGAAGATCGCGCGGAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACC
TCGGCACCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAATGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCT
GGCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGCCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGGACTTAATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCAC
CGGGGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACGCGGTTCTCGGTGAAGCGCATGCCCTCGTCGAGGGTGGCGTTGGCCTGCGC
GATGACGAATCGGTCCTCTTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACGACACCGTCGACGACCTTGCGGTACGGCGTCTCCACGAAACC
GAACGCGTTGACGCGGCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATACGGCCGTAGTGA
GACGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGACAGACCACCAGGGCCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACTCGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAGC
TCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATGTTCGCGCCGCTGTCGCCGACCGTCTCGGTCTCG
CCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGACGCCCGCGTCGAGCGGAGCCTCCGCGCCCAGCTTCTTG
TTGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAACGCTTCGGGTTGAAGTAGAGGTTCTCCAGAAGCGTCTGCGCGGCCTCACGCGTGGGGGG
CTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGTGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTC
GCCGAACTCCTCCAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTGAGCAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGAC
CATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATCGA
GGAGTCGAAGTAGACACCCGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTGGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCAGCCGTCACGAAGAGCGGGGCCGCGTACGTGAA
GTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGG
GGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTCGTGGGGACGTCCTGACCGGACTCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGTT
CCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTCGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAGGAAATGCGCAGCGGGG
CGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP029601.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCACAGGGACTGCATCTCCTTGATGAGGACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGCGCGTAGG
CGGCGCCGTACGCCTCGAGTGCCCACACCTCCATCTCGCCGAACCGCTGGCCACCGAACTGAGCCTTACCGCCCAGCGGCTGCTGGGTGATCATGG
AGTACGGACCGGTCGAACGCGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGCAGGACCATGCGGTCGCCGTCGCGGTTCGGGACGGTGTGGTCGAGCAGAC
CGGCGAGCTCGTCCTCGCGGGCGCCGTCGAAGACCGGGGTGGCCACGTTGGTACGGGGCTCCACGCGGTCCGCGCCGATCGCCTGGAGGCGCTG
CGCCCACTCGTCGGCGAGACCGGAGACGTCCCAGCCCTGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATCCGGGACG
GGACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTCCCGTCCTCGAGGAACGGCATGTCCTCAATGGGCAGGATCTTGGAGATGACGCCCTTG

TTGCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTGACACCGGGGGGAAGCTC
GTCGCCCTCCTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGCAGCAGACGCTCCTCGGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGATGA
CCTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCG
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCACCGTTCTGGGTGGCCGGACCGTCGGCGAGGAC
CTGGCCCTCGATGATCCGGTCGCCCTCGGCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCGAGGCGGTACG
TGATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGT
CGACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGCTCTTGATCAGCGGCACGGCCTGGCGCATCATGTTCGCGCCCATGAGG
GCACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACCTCGTC
ACCGGGGACGTAGTCGACCTCGCCGCCGCGGCGGCGGACCAGGATGCGGTTCTCGGCGAAGCGCATGTCGTCGCTGAGCGGCGCGTTGGCCTGC
GCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAA
GCCGAACGGGTTGATCCGCCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCCGGCGTCTCGATCGGGCACATGCGGCCGTAGT
GCGAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGGGCGTTGAGACGGCGCTTGTGCGTCAGCCC
CGACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGG
TCTGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATCA
GCTCGCCGACGTTGCGGATGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGATGATCGAGCGGCCCGACTCGCCGACCGTCTCGGTC
TCCCCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGATGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCCCCGCCGAGCTTC
TTGTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGGGTGGG
GGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGT
ACTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTCAGGAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGCACAC
CGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGAGGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGA
TGGAGGAGTCGAAGTAGACGCCGGGCGAGCGGACCAGCTGGGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCAT
GAGCGGGAAGTCGCCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCGGCGTAC
GTGAAGTCGCGCTCCTTGCACTCGTCGATGGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTC
GATCGGGGAGATCTCTTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGCCCGTTCTCGAGAGCCTCCTCGACGCGAGCCTTCCAGGC
GTCGTTCCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCA
GCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|KX503547.1|

GACCTCTTCGACGCTGCTCGGCTCGCGCCTCGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTCGAT
GGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTCGG
GGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAACTCCTGGAGGGCGTACGCGG
CGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCGAGT
ACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAACGG
CTCGCCGGAGCGGCCGTCGAACAGGCTCGCCTTGCCGGTCGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGCAGCAGGCCGG
CCAGCTCGTCCTCACGCGCACCGTCGAAGACGGGAGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATGATCTGCAGGCGCTCCGCC
CATTCCTCGGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGAACCTGTCCCGGGTTCATTCGGGACGGCAC
ACCGAGGGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCATGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTGTTGC
CGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTCG
CCCTCCTCGCGGTCGAAGACGCGCACGCCGATGATCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTCG
CCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGACGACCTC
GGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCT
TGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGATGTCGTCCTGGACGAGGCGCTGCGACAGGATGATCGCGTCCTCG
TAGTTGTGGCCTTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCGAGGACCTGG
CCCTCGATGACGCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTCGAGCGGGCGAACTTGGCCAGGCGGTACGTGATG
TACGTGCCGTCGTCGTTGGTCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGGCG
GCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGGAGCGGATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACCTCGTCGGG
GCTGACGTAGTCGACCTCGCCGCCCTTGCGGCGGACCAGTACACGGGCCTCGGCGAACCGCATGTCGTCACTCAGCGGGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACGTCGTCGGTGACCTGGCCGTCGATCACCTTGCGGTACGGGGTCTCGACGAAACCG
AACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGGCCGATGTTCGGGCCTTCGGGCGTCTCGATCGGACACATGCGGCCGTAGTGCG
AGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGCGCGGACAGACGACGCTTGTGCGTCAGCCCGGA
CAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCT
GCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCACGCACGACGCGCTCCATACGAGCCAGACCCGTACGGACCTGGTTCTGGATGAGCT
CGCCGACGCTGCGCAGGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATGTTCTGACCGCTGTCACCGACGGTCTCGATCTCGC
CGGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGACGTCCTCGACGGTCAGCACGCCCGCGTCGAGCGGCGCGTCGGCACCCAGCTTCTTGT
TGACCTTGTAACGGCCGACCTTGGCCAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGGGTGGGGGGC

TCGCCCGGACGCAGCTTGCGGTAGATGTCCAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCG
CCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTGAGCAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGACGCGGACGCCGAC
CATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGA
CGAGTCGAAGTAGACACCCGGCGAACGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGCG
GGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTAGGTGAAG
TCGCGGTCCTTGCACTCGTCGATGCTGTTCTTGGGAGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGG
GGAGATCTCCTCGAAGATCTCTTCCAGTCCGGACTTCCTGGGGACGTCCTGACCGTTCTCCAGAGCGTCCTCGACGCGAGTCTTCCAGGCGTCGTTG
CCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAAGAGATGCGCAGCGGGGC
GGTGCTGGCAATGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP029788.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATACCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATTTCCTTGATGAGCACCTTGAAGGACTCGGGAATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAG
GCGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATC
GAGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGA
ACGGCTCACCGCTACGGCCGTCGAACAGTCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGA
CCCGCGAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCGCCTGGAGGCGCTG
CGCCCACTCGTCCGCCAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GCACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACACCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGGACCAGCTGGTTCACACCCGGCGGCAGCTCG
TCGCCCTCCTCGCGGTCGAAGACGCGCACACCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTC
TCACCGAAGATGGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGATGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTGGGTGGCCGGGCCGTCGGCGAGCACCT
GGCCCGTGATGACGCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCGGCCTTGACCACGTCGCCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGGACTTAATCAGCGGCACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCCTCGG
GGGCGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACGCGGGCCTCCTCGAACCGCATGTCGGCGGTGAGCGGCGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTTCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGGGTCTCGACGAAGC
CGAACGCGTTGACGCGGCCGTAGGTGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGGCCGTAGTG
CGACGGGTGGACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGCGCGGACAGACGACGCTTGTGCGTCAGCCCG
GACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGT
CTGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATCCGCTCGCGGACGACACGCTCCATACGGGCCAGACCCGTGCGGACCTGGTTCTGGATGAG
CTCGCCGACGCTGCGCAGGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGACCGGCTGGCCGTTGTCGCCGACGGTCTCGGTCT
CACCGGCGTGCAGCTTCACCAGGTACTTGATCGTCGCGATGATGTCCTCGACGGTCAGGATCCCGGCGTCCAGCGGGGTGTCCGTGCCCAGCTTCT
TGTTGACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGG
GGCTCGCCCGGGCGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTA
CTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTGAGCAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGGACACC
GACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGAT
GCTGGAGTCGAAGTAGACACCGGGGGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAACGTGCCCTTGTTCGTCATGA
GCGGGAAGTCGCCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGCGGCGCGGCGTAGGT
GAAGTCGCGCTCCTTGCACTCGTCGATCGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCCGAGAAGTCCTCGAT
CGGGGAGATCTCCTCGAAGATCTCCTCGAGGCCGGACTTGGTGGGGACGTCCTGACCGTTCTCGAGAGCCTCCTCGACCCGACTCTGCCAGGCGGT
GTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTCGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCG
GGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>dbj|AP023439.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAAC
GGCTCGCCGGAGCGGCCGTCGAACAGCGGCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGGTTCAGCAGACC
GGCCAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTGGTGCCCGGCGCGACCTGGTCGGCCTCGATGGCCTGCAGGCGCTGG
GCCCAGTCGTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGACCCGGGTTCATTCGGGACG

GCACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATGGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGA
CCTCGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCG
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTGGGTGGCCGGACCGTCGGCCAGGAC
CTGGCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGCCGGTACGT
GATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCAGAGACCTCCTGGACCACACCGGGCTTCTCGGCCTTGACCACGTCGCCGGCGTC
GACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGGACTTAATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGG
GCACGGTTGGCGTCGTCGTGCTCGAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCCTC
GCCGGTGACGTAGTCGACCTCGCCGCCCTTACGGCGGACGAGCACCCGGGCCTCGGCGAAGCGGAGGTCGTCCGTCAGCGC-
GCGTTGGCCTGCGCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCCTCGAACACCTTGCGGTACGG
GGTCTCCACGAAACCGAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACA
TGCGGCCGTAGTGCGAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTT
GTGGGTGAGACCCGACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTACCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGA
TGTTGATCAGGGTCTGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATCCGCTCGCGGACGACACGCTCCATACGCGCCAGACCCGTACGGACCT
GGTTCTGGATGAGCTCGCCGACGCTGCGCAGGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATGTTCGTACCGCTGTCGCCG
ACCGTCTCGGTCTCGCCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCCGCGTCCAGCGGAGCGTCC
GCACCCAGCTTCTTGTTGACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTTTCGAGCAGCGTCTGCGCGGCC
TCACGCGTGGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCG
CATCGACTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTGAGCAGGACGGTGACGGACTGCTTGCGCTTGCGGTC
GATGCGCACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGA
CGTCTTGTCGATGGAGGAGTCGAAGTAGACACCGGGCGAACGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCT
TGTTCGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTCGTGAACTCAGCGGTGACGAAGAGCGGG
GCCGCGTACGTGAAGTCGCGGTCCTTGCACTCGTCGATGGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGCT
GAAGTCCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCGAGCCCGGACTTGGTGGGGACGTCCTGACCGTTCTCGAGAGCCTCCTCGACCCGACT
CTGCCAGGCGGTGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGCAGGTTCGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAG
AGATGCGCAGCGGGGCAGTGCTGGCGGCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGG
ACTCAC

>gb|FJ438820.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATACCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATTTCCTTGATGAGCACCTTGAAGGACTCGGGAATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAG
GCGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATC
GAGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGA
ACGGCTCACCGCTACGGCCGTCGAACAGTCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGA
CCCGCGAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCGCCGATCGCCTGGAGGCGCTG
CGCCCACTCGTCCGCCAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GCACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACACCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGGACCAGCTGGTTCACACCCGGCGGCAGCTCG
TCGCCCTCCTCGCGGTCGAAGACGCGCACACCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTC
TCACCGAAGATGGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGATGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTGGGTGGCCGGGCCGTCGGCGAGCACCT
GGCCCGTGATGACGCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCGGCCTTGACCACGTCGCCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGGACTTAATCAGCGGCACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCCTCGG
GGGCGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACGCGGGCCTCCTCGAACCGCATGTCGGCGGTGAGCGGCGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTTCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGGGTCTCGACGAAGC
CGAACGCGTTGACGCGGCCGTAGGTGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGGCCGTAGTG
CGACGGGTGGACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGCGCGGACAGACGACGCTTGTGCGTCAGCCCG
GACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGT
CTGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATCCGCTCGCGGACGACACGCTCCATACGGGCCAGACCCGTGCGGACCTGGTTCTGGATGAG
CTCGCCGACGCTGCGCAGGCGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGACCGGCTGGCCGTTGTCGCCGACGGTCTCGGTCT
CACCGGCGTGCAGCTTCACCAGGTACTTGATCGTCGCGATGATGTCCTCGACGGTCAGGATCCCGGCGTCCAGCGGGGTGTCCGTGCCCAGCTTCT

TGTTGACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGG
GGCTCGCCCGGGCGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTA
CTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTGAGCAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGGACACC
GACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGAT
GCTGGAGTCGAAGTAGACACCGGGGGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAACGTGCCCTTGTTCGTCATGA
GCGGGAAGTCGCCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGCGGCGCGGCGTAGGT
GAAGTCGCGCTCCTTGCACTCGTCGATCGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCCGAGAAGTCCTCGAT
CGGGGAGATCTCCTCGAAGATCTCCTCGAGGCCGGACTTGGTGGGGACGTCCTGACCGTTCTCGAGAGCCTCCTCGACCCGACTCTGCCAGGCGGT
GTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTCGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCG
GGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP029195.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATCGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGAACGCGCGTGGAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGACGGAGATCGGGTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCCTCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGGTTGAGCAGAC
CCGCGAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGGTGGCGACGTTGGTGCCGGGGGCGACCTTGTCGGCGCCGATCGCCTGGAGGCGCTG
CGCCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GGACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGGGCGACGTAGACGCGGACCAGCTGGTTCACGCCCGGGGGAAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGATCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGA
CCTCGGCACCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCCGGGCCG
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTCTTCGGTGGCGGGACCGTCGGCGAGGAC
CTGGCCCTCGATCACGCGGTCGCCCTCGTCGACGACGACCTTCTGGTTGACCGAGGTGCCCTGGTTCGAGCGGGAGAACTTGTGCAGGCGGTACGT
GATGTACGTGCCGTCGTCGTTGGCGACCGTGATGTAGTCGGCCGACAGCTCCTGGACGACGCCGTCCTTCTCGGCCTTGAGCACGTCGCCGGCGTC
GGTGGCGCAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCCGACTTAATCAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGG
CACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGACG
GCTC-
GACGTAGTCGACCTCGCCGCCACGCTTGCGGACCAGGACGCGGTTCTCGGAGAAGCGGAGCTCCTCGTCGAGCGCGGCGTTGGCCTGGGCGATGA
CGAAGCGGTCCTCCTCGTCGGCGGTGACGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCGAAC
GCGTTGACGCGGCCGTACGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGTCCGTAGTGGGACG
GGTGCACGTCTCGGACCTCGAAGCCGGCCCGCTCACGGGACAGACCACCAGGGCCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCCGACAGC
GGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCGG
CGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTCGCC
GACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCT-
CGTGCCGTTCTCGCCGATCGTCTCGGTCTCGCCCGCGTGGAGCTTGACCAGGTACTTGATCGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGC
GTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTTCACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAG
CAGCGTCTGAGCGGCCTCACGGGTCGGCGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCT
TCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTGGAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTGAGGAGGACGGTGACGGAC
TGCTTGCGCTTGCGGTCGATGCGGACACCGACCAGGTCGCGCTTGTCGATCTCCATCTCGAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGA
GAAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGAGTCGAAGTAGACACCGGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGT
TGATGACGAAGGTGCCCTTGTTCGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGG
CGGTGACGAAGAGCGGGGCGCCGTACGTGAAGTCACGGTCCTTGCACTCGTCGATCGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTC
AGCGACATCGACCCGGAGAAGTCCTCGATCGGCGAGATCTCCTCGAAGATCTCTTCCAGGCCGGACTTCGTGGGGACGTCCTGTCCCGACTCAAGA
GCAGCCTCGACGCGAGCCTTCCAGGCGGCGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGGAGGTTCGGAACCTCGAGGGGCTC
CTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTTGCGCCGTTGTTCGTATT---
GGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP045096.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGGACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGAATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGATCGAGATCGGGTCCGGGAAC

GGCTCACCGGAGCGGCCGTCGAACATCCTCGCCTTGCCGGTCGGCTGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGACC
GGCCAGTTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTTCCGGGGGCGACGGAGTCGGCGCCGATGACCTGCAGACGCTTC
GCCCAGTCGTCGCCGAGACCGGAGACGTCCCAGCCGCGACTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATGCGGGACG
GCACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACACCCTTG
TTGCCGTGCCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGGGCGACATAGACACGGACCAGCTGGTTGACGCCGGGCGGCAGCTC
GTCGCCCTCTTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGGACCTTCAGCGATGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGACGAC
CTCGGCACCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGA
GCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGTACGAGGCGCTGCGACAGGATGATCGCGTCC
TCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGGGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACC
TGGCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCCAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCGTCAC
CGGGGACGTAGTCGACCTCGCCGCCACGACGGCGGACCAGGACGCGGTTCTCGGTGAAGCGCATGTCGTCGTCGAGCGTGGCGTTGGCCTGCGC
GATGACGAATCGGTCCTCTTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACGACACCGTCGGTGACCCGGCGGTAGGGCGTCTCCACGAAACC
GAACGCGTTGACGCGCCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCGGGGGTCTCGATCGGACACATACGGCCGTAGTGC
GAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGAGACAGACCACCGGGACCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATCCGCTCGCGGACGACTCGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAGC
TCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCTGGCCGTTGTCCGCGGCCGTCTCGGTCTCG
CCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGACGTCCTCGACGGTCAGGACGCCCGCGTCGAGCGGAGCCTCCGCACCCAGCTTCTTG
TTGACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTACAGGTTCTCCAGAAGCGTCTGCGCGGCCTCACGCGTGGGGGG
CTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGCCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTC
GCCGAACTCCTCCAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGAGGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCGGGGGAACGGACCAGCTGCGACACGACGACGCGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTGGGTCATGAG
CGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTAAGTGA
AGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGATC
GGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCGTTCTCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTG
TTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGCGGCTCCTTGATCTTTGCGAAAGAGATGCGCAGCGG
GGCAGTGCTGGCGGCTGGGTTCGTATTCGAGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP029043.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCACGCATTTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGCGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACGC
GGCGCCGTAGGCCTCGAGGGCCCACACCTCCATCTCGCCGAACCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCGA
GTACGGGCCGGTCGAACGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGATGGAGATCGGGTCCGGGAAC
GGCTCACCGGAGCGGCCGTCGAACAGCCTCGCCTTACCGGTCGGCTGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGGAGCAGACC
CGCGAGCTCGTCCTCGCGCGCACCGTCGAAGACGGGGGTCGCGACGTTCGTGCCCGGGGCGACCTGGTCGGCGCCGATCACCTGCAGGCGCTGCG
CCCAGTCCTCGGCGAGGCCCGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGC
ACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTGTTG
CCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGCACACCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGGGTCACCTTGCCGACGAGGATGTCGCCGGCGATGACCTC
GGCGCCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCT
TGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCG
TAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGGGCCATCTCACCGTTCTCGGTGGCCGGACCGTCGGCGAGGACCTGG
CCCTCGACGATCCGGTCGCCCTCGTGGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGAT
GTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCTGCCTTCTCGGCCTTGACGACGTCGCCGGCGTCGAC
CGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGC-
CTTGATGAGCGGAACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGCACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGA
CCGACACCATCTGGCGCGGGGAGACGTCCATGTAGTCGACGTCCTCGCCGCCGACGTAGTCGACCTCGCCGCCACGCCGGCGGACCAGGACGCGG
GCCTCCTCGAAGCGGAGGTCGTCCGTCAGCGGCGCGTTGGCCTGCGCGATGACGAAGCGGTCCTCCTCGTCGGCCGTCAGGTAGTCGACCTCGTC
GGTGACCTGGCCGTCGATGACCTTGCGGTACGGCGTCTCGACGAAGCCGAACGCGTTGACCCGGCCGTACGAGGCGAGCGAACCGATCAGACCGA
TGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGACGGGTGGACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGACAGA
CCACCCGGGCCGAGCGCGGACAGACGACGCTTGTGCGTCAGCCCGGACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAA
GAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCACGCACGA

CGCGCTCCATACGAGCCAGACCCGTACGGACCTGGTTCTGGATGAGCTCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCG
GTCTCGACGACGATCT-
CTGACCGCTGTCGCCGACCGTCTCGGTCTCGCCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCCGC
GTCCAGCGGAGCGTCCGCACCCAGCTTCTTGTTGACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCGAG
CAGCGTCTGCGCGGCCTCACGGGTCGGCGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCT
TCTCCAGGGTGGCGCGCATCGACTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTCAGGAGGACGGTGACGGACT
GCTTGCGCTTGCGGTCGATGCGCACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGA
AGATGTCCTTGTCGGACGTCTTGTCGATGGAGGAGTCGAAGTAGACGCCCGGCGAGCGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTG
ATGACGAAGGTGCCCTTGTTGGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTCGTGAACTCGGCG
GTGACGAAGAGCGGGGCCGCGTACGTGAAGTCGCGCTCCTTGCACTCGTCGATGCTGTTCTTCGGGGGCTCGAAACGGTGGTCGCGGAACGTCAG
CGACATCGACCCGGAGAAGTCCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCGAGGCCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAG
CCGCCTCGACCCGACTCTGCCAGGCGGTGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGCAGGTTCGGAACCTCGAGGGGCTCC
TTGATCTTTGCAAAAGAGATGCGCAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGT
CCTTCCGAGGGCTCGGACTCAC

>gb|CP021978.1|

TCAGACCTCTTCGACGCTGCTCGGCTCACGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGAATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCGGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCGGCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGGTTCAGCAGAC
CGGCCAGCTCGTCCTCACGCGCACCGTCGAAGACAGGGGTGGCGACGTTGGTGCCGGGGGCGACCTGGTCGGCCTCGATGGCCTGCAGGCGCTG
CGCCCAGTCGTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GCACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGCAGCAGGCGCTCCTCGGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGA
CCTCGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCG
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTGGGTGGCCGGACCGTCGGCCAGGAC
CTGGCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGT
GATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGGGCTTCTCGGCCTTGACCACGTCGCCGGCGTC
GACCGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGGACTTAATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGG
CACGGTTGGCGTCGTCGTGCTCGAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCCTCG
CCGGTGACGTAGTCGACCTCGCCGCCCTTACGGCGGACCAGGACTCGTGCCTCGGCGAAGCGGAGGTCGTCCGTCAGCGGCGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGACCCTCGAAGACCTTGCGGTACGGGGTCTCGACGAAACC
GAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGGCCGTAGTGC
GAGGGGTGCACGTCACGGACTTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGGGCCGACAGACGGCGCTTGTGGGTGAGACCCG
AGAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATCCGCTCACGGACGACACGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAGC
TCGCCGACGCTACGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATGTTCGTACCGCTGTCGCCGACCGTCTCGGTCTCG
CCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCCGCGTCCAGCGGAGCGTCCGCGCCCAGCTTCTTG
TTGACCTTGTAACGGCCGACCTTGGCCAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGG
CTCACCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTC
GCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTGAGCAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTCGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACACCGGGCGAACGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTACCGGTCATGAG
CGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGA
AGTCGCGGTCCTTGCACTCGTCGATGGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGCTGAAGTCCTCGATC
GGGGAGATCTCCTCGAAGATCTCCTCGAGCCCGGACTTGGTGGGGACGTCCTGACCGTTCTCGAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTG
TTGCCGAGGAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGCAGGTTCGGAACCTCAAGAGGCTCCTTGATCTTTGCAAAAGAGATGCGCAGCGG
GGCAGTGCTGGCGGCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP017248.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGCGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACGC

GGCGCCGTAGGCCTCGAGGGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCGA
GTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCGGAGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGACC
CGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCCGGG-
GCGACCTGGTCGGCGCCGATGGCCTGCAGGCGCTGGGCCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGT
GGATCTCCAGGACCTGCCCCGGGTTCATTCGGGACGGCACGCCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATG
TCCTCGATCGGCAGGATCTTGGAGATGACACCCTTGTTGCCGTGACGGCCGGCGAGCTTGTCGCCGTCGGTGATCTTGCGCTTCTGCGCGACGTAG
ACGCGGACCAGCTGGTTCACGCCCGGCGGCAGCTCGTCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGG
CACCTTCAGCGAGGTGTCACGGACCTCACGTGCCTTCTCCCCGAAGATCGCGCGCAGCAGACGCTCCTCGGGGGTCAGCTCGGTCTCACCCTTCGG
GGTGACCTTGCCGACGAGGATGTCACCGGCGATGACCTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGA
CGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCG
TCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGGGCCATC
TCGCCGTTCTGGGTGGCCGGGCCGTCGGCCAGCACCTGGCCGGTGACGACGCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCC
TGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCAC
ACCGTCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGACCGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGTGCCTCGGACTTAATCAGCGG
CACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCACGGTTGGCGTCGTCGTGCTCGAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCT
GGCGCGGCGAGACGTCCATGTAGTCCACGTCCTCGGGGCCGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGCACGCGGGCCTCGGCGAA
CCGCATCTCGTCGTTCAGCGGCGCGTTGGCCTGCGCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGCTGACCTGGCC
GTCGACGACCTTGCGGTACGGCGTCTCGATGAAACCGAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTT
CCGGCGTCTCGATCGGGCACATACGGCCGTAGTGCGACGGGTGAACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCC
GAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGA
TGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGTGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATA
CGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTCACCGACGCTGCGGATACGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACCAT
GATCGTGCGACCGGTCTCGGTCACCGTCTCGGTCTCGCCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCCTCGGTGGTGAGCAC
GCCGGCGTCCAGCGGCTCGTCCGCACCGAGCTTCTTGTTGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTT
CTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGT
GGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAGCCGAGCGCCTTCAGCAGCACGGTGA
CCGACTGCTTGCGCTTGCGGTCGATGCGGACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGG
CGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGTGGAGTCGAAGTAGACACCGGGGGAACGGACCAGCTGGGACACCACGACACGCTCGGT
GCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCACCGGTCTCGTTATTGGTGAA
CTCGGCGGTGACGAAGAGCGGGGCGGCGTACGTGAAGTCGCGCTCCTTGCACTCGTCGATGGAGTTCTTCGGCGGCTCGAAGCGGTGGTCCCGGA
ACGTCAGGGACATCGACCCGGAGAAGTCCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGTTCT
CCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGCAGGTTCGGAACCTCGAGA
GGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAG
GGGGTCCTTCCGAGGGCTCGGACTCAC

>emb|HE971709.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCGGCGCGGAAGACATCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACTCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACGC
GGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAACCGCTGGCCACCGAACTGAGCCTTACCACCCAGCGGCTGCTGGGTGATCATGGA
GTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAACG
GCTCACCGCTACGGCCGTCGAACAGCCGCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGGTTCAGCAGACCC
GCGAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGTCGCGACGTTCGTGCCCGGCGCGACCTGGTCGGCGCCGATCGCCTGGAGGCGCTGCGC
CCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGGA
CACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACGCCCTTGTTG
CCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGAACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGATGACCTC
GGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCT
TGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCG
TAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGCGCCGTCGGCGAGGACCTGG
CCCTCGATGACGCGGTCGCCCTCGTCGATGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGAT
GTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGAC
GGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGGACTTAATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCCTCGCCA
CCGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACGCGGCCCTCGGTGAACCGCATGTCGTCATTGAGCGTGGCGTTGGCCTGCGCGAT
GACGAACCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGAGACCTGACCGTCGACGACCTTGCGGTACGGCGTCTCGACGAAGCCGA
ACGCGTTGACGCGACCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATACGGCCGTAGTGAGA

CGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGACCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCGAC
AGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTG
CGGCGTGATCGCCTCGACGTCCTGAGTCGTCATGCGCTCACGGACGACGCGCTCCATACGAGCCAGACCCGTACGGACCTGGTTCTGGATGAGCTC
GCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCTGGCCGTTGTCCGCGGCCGTCTCGGTCTCACC
GGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGACGTCCTCGACGGTCAGGATGCCCGCGTCCAGCGGAGCCTCCGCACCCAGCTTCTTGTT
GACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGGCT
CGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGC
CGAACTCCTCCAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTGAGCAGAACGGTCACGGACTGCTTGCGCTTGCGGTCGATACGCACACCGACCA
TGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCCGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAGG
AGTCGAAGTAGACACCGGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTGCGTCATGAGCGG
GAAGTCGCCCATGAAGACCGTCTGAGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTCACGAAGAGCGGGGCCGCGTACGTGAAGT
CACGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGGG
AGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCTGACTCCAGAGCCGCCTCGACCCGACTCTGCCAGGCGGTGTTCCC
GAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCG
GTGCTTGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP029196.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATCGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGAACGCGCGTGGAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGACGGAGATCGGGTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCCTCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGGTTGAGCAGAC
CCGCGAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGGTGGCGACGTTGGTGCCGGGGGCGACCTTGTCGGCGCCGATCGCCTGGAGGCGCTG
CGCCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GGACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCGAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGGGCGACGTAGACGCGGACCAGCTGGTTCACGCCCGGGGGAAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGATCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGA
CCTCGGCACCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCCGGGCCG
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTCTTCGGTGGCGGGACCGTCGGCGAGGAC
CTGGCCCTCGATCACGCGGTCGCCCTCGTCGACGACGACCTTCTGGTTGACCGAGGTGCCCTGGTTCGAGCGGGAGAACTTGTGCAGGCGGTACGT
GATGTACGTGCCGTCGTCGTTGGCGACCGTGATGTAGTCGGCCGACAGCTCCTGGACGACGCCGTCCTTCTCGGCCTTGAGCACGTCGCCGGCGTC
GGTGGCGCAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCCGACTTAATCAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGG
CACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGACG
GCTC-
GACGTAGTCGACCTCGCCGCCACGCTTGCGGACCAGGACGCGGTTCTCGGAGAAGCGGAGCTCCTCGTCGAGCGCGGCGTTGGCCTGGGCGATGA
CGAAGCGGTCCTCCTCGTCGGCGGTGACGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCGAAC
GCGTTGACGCGGCCGTACGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCCGGGGTCTCGATCGGGCACATGCGTCCGTAGTGGGACG
GGTGCACGTCTCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCAGGGCCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCCGACAG
CGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCG
GCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCACGCACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTCGC
CGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCT-
CGTGCCGTTCTCGCCGATCGTCTCGGTCTCGCCCGCGTGGAGCTTGACCAGGTACTTGATCGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGC
GTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTTCACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAG
CAGCGTCTGAGCGGCCTCACGGGTCGGCGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCT
TCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTGGAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTGAGGAGGACGGTGACGGAC
TGCTTGCGCTTGCGGTCGATGCGGACACCGACCAGGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAG
AAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGAGTCGAAGTAGACACCGGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTT
GATGACGAAGGTGCCCTTGTTCGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGC
GGTGACGAAGAGCGGGGCGCCGTACGTGAAGTCACGGTCCTTGCACTCGTCGATCGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCA
GCGACATCGACCCGGAGAAGTCCTCGATCGGCGAGATCTCCTCGAAGATCTCTTCCAGGCCGGACTTCGTGGGGACGTCCTGTCCCGACTCAAGGG
CAGCCTCGACGCGAGCCTTCCAGGCGGCGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGGAGGTTCGGAACCTCGAGGGGCTCC
TTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTTGCGCCGTTGTTCGTATT---
GGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP029197.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATCGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGAACGCGCGTGGAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGACGGAGATCGGGTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCCTCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGGTTGAGCAGAC
CCGCGAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGGTGGCGACGTTGGTGCCGGGGGCGACCTTGTCGGCGCCGATCGCCTGGAGGCGCTG
CGCCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GGACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCGAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGGGCGACGTAGACGCGGACCAGCTGGTTCACGCCCGGGGGAAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGATCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGA
CCTCGGCACCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCCGGACCG
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTCTTCGGTGGCGGGACCGTCGGCGAGGAC
CTGGCCCTCGATCACGCGGTCGCCCTCGTCGACGACGACCTTCTGGTTGACCGAGGTGCCCTGGTTCGAGCGGGAGAACTTGTGCAGGCGGTACGT
GATGTACGTGCCGTCGTCGTTGGCGACCGTGATGTAGTCGGCCGACAGCTCCTGGACGACGCCGTCCTTCTCGGCCTTGAGCACGTCGCCGGCGTC
GGTGGCGCAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCCGACTTAATCAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGG
CACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGACG
GCTC-
GACGTAGTCGACCTCGCCGCCACGCTTGCGGACCAGGACGCGGTTCTCGGAGAAGCGGAGCTCCTCGTCGAGCGCGGCGTTGGCCTGGGCGATGA
CGAAGCGGTCCTCCTCGTCGGCGGTGACGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCGAAC
GCGTTGACGCGGCCGTACGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCCGGGGTCTCGATCGGGCACATGCGTCCGTAGTGGGACG
GGTGCACGTCTCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCAGGGCCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCCGACAG
CGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCG
GCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTCGC
CGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCT-
CGTGCCGTTCTCGCCGATCGTCTCGGTCTCGCCCGCGTGGAGCTTGACCAGGTACTTGATCGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGC
GTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTTCACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAG
CAGCGTCTGAGCGGCCTCACGGGTCGGCGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCT
TCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTGGAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTGAGGAGGACGGTGACGGAC
TGCTTGCGCTTGCGGTCGATGCGGACACCGACCAGGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAG
AAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGAGTCGAAGTAGACACCGGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTT
GATGACGAAGGTGCCCTTGTTCGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGC
GGTGACGAAGAGCGGGGCGCCGTACGTGAAGTCACGGTCCTTGCACTCGTCGATCGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCA
GCGACATCGACCCGGAGAAGTCCTCGATCGGCGAGATCTCCTCGAAGATCTCTTCCAGGCCGGACTTCGTGGGGACGTCCTGTCCCGACTCAAGGG
CAGCCTCGACGCGAGCCTTCCAGGCGGCGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGGAGGTTCGGAACCTCGAGGGGCTCC
TTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTTGCGCCGTTGTTCGTATT---
GGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP018074.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATCGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGAACGCGCGTGGAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGACGGAGATCGGGTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCCTCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGGTTGAGCAGAC
CCGCGAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGGTGGCGACGTTGGTGCCGGGGGCGACCTTGTCGGCGCCGATCGCCTGGAGGCGCTG
CGCCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GGACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCGAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGGGCGACGTAGACGCGGACCAGCTGGTTCACGCCCGGGGGAAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGATCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGA
CCTCGGCACCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCCGGACCG
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTCTTCGGTGGCGGGACCGTCGGCGAGGAC
CTGGCCCTCGATCACGCGGTCGCCCTCGTCGACGACGACCTTCTGGTTGACCGAGGTGCCCTGGTTCGAGCGGGAGAACTTGTGCAGGCGGTACGT
GATGTACGTGCCGTCGTCGTTGGCGACCGTGATGTAGTCGGCCGACAGCTCCTGGACGACGCCGTCCTTCTCGGCCTTGAGCACGTCGCCGGCGTC
GGTGGCGCAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCCGACTTAATCAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGG

CACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGACG
GCTC-
GACGTAGTCGACCTCGCCGCCACGCTTGCGGACCAGGACGCGGTTCTCGGAGAAGCGGAGCTCCTCGTCGAGCGCGGCGTTGGCCTGGGCGATGA
CGAAGCGGTCCTCCTCGTCGGCGGTGACGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCGAAC
GCGTTGACGCGGCCGTACGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCCGGGGTCTCGATCGGGCACATGCGTCCGTAGTGGGACG
GGTGCACGTCTCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCAGGGCCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCCGACAG
CGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCG
GCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTCGC
CGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCT-
CGTGCCGTTCTCGCCGATCGTCTCGGTCTCGCCCGCGTGGAGCTTGACCAGGTACTTGATCGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGC
GTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTTCACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAG
CAGCGTCTGAGCGGCCTCACGGGTCGGCGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCT
TCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTGGAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTGAGGAGGACGGTGACGGAC
TGCTTGCGCTTGCGGTCGATGCGGACACCGACCAGGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAG
AAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGAGTCGAAGTAGACACCGGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTT
GATGACGAAGGTGCCCTTGTTCGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGC
GGTGACGAAGAGCGGGGCGCCGTACGTGAAGTCACGGTCCTTGCACTCGTCGATCGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCA
GCGACATCGACCCGGAGAAGTCCTCGATCGGCGAGATCTCCTCGAAGATCTCTTCCAGGCCGGACTTCGTGGGGACGTCCTGTCCCGACTCAAGGG
CAGCCTCGACGCGAGCCTTCCAGGCGGCGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGGAGGTTCGGAACCTCGAGGGGCTCC
TTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTTGCGCCGTTGTTCGTATT---
GGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>emb|FR845719.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATCGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGGCCGGTCGAACGCGCGTGGAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGACGGAGATCGGGTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCCTCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGGTTGAGCAGAC
CCGCGAGCTCGTCCTCACGCGCACCGTCGAAGACCGGGGGTGGCGACGTTGGTGCCGGGGGCGACCTTGTCGGCGCCGATCGCCTGGAGGCGCTG
CGCCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GGACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCGAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTG
TTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGGGCGACGTAGACGCGGACCAGCTGGTTCACGCCCGGGGGAAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGATCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGA
CCTCGGCACCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCCGGACCG
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTCTTCGGTGGCGGGACCGTCGGCGAGGAC
CTGGCCCTCGATCACGCGGTCGCCCTCGTCGACGACGACCTTCTGGTTGACCGAGGTGCCCTGGTTCGAGCGGGAGAACTTGTGCAGGCGGTACGT
GATGTACGTGCCGTCGTCGTTGGCGACCGTGATGTAGTCGGCCGACAGCTCCTGGACGACGCCGTCCTTCTCGGCCTTGAGCACGTCGCCGGCGTC
GGTGGCGCAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCCGACTTAATCAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGGG
CACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGACG
GCTC-
GACGTAGTCGACCTCGCCGCCACGCTTGCGGACCAGGACGCGGTTCTCGGAGAAGCGGAGCTCCTCGTCGAGCGCGGCGTTGGCCTGGGCGATGA
CGAAGCGGTCCTCCTCGTCGGCGGTGACGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCGAAC
GCGTTGACGCGGCCGTACGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCCGGGGTCTCGATCGGGCACATGCGTCCGTAGTGGGACG
GGTGCACGTCTCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCAGGGCCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCCGACAG
CGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCG
GCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTCGC
CGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCT-
CGTGCCGTTCTCGCCGATCGTCTCGGTCTCGCCCGCGTGGAGCTTGACCAGGTACTTGATCGTGGCGATGACGTCGTCGGTGGTGAGCACGCCGGC
GTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTTCACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAG
CAGCGTCTGAGCGGCCTCACGGGTCGGCGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCT
TCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTGGAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTGAGGAGGACGGTGACGGAC
TGCTTGCGCTTGCGGTCGATGCGGACACCGACCAGGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAG
AAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGAGTCGAAGTAGACACCGGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTT
GATGACGAAGGTGCCCTTGTTCGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGC
GGTGACGAAGAGCGGGGCGCCGTACGTGAAGTCACGGTCCTTGCACTCGTCGATCGAGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCA
GCGACATCGACCCGGAGAAGTCCTCGATCGGCGAGATCTCCTCGAAGATCTCTTCCAGGCCGGACTTCGTGGGGACGTCCTGTCCCGACTCAAGGG

CAGCCTCGACGCGAGCCTTCCAGGCGGCGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGGAGGTTCGGAACCTCGAGGGGCTCC
TTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTTGCGCCGTTGTTCGTATT---
GGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP029188.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATACCGAGCTCCTCCGCGGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAG
GCGGCGCCGTACGCCTCGAGCGCCCACACCTCCATCTCGCCGAACCGCTGGCCACCGAACTGAGCCTTACCACCCAGCGGCTGCTGGGTGATCATC
GAGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGA
ACGGCTCGCCGGAGCGGCCGTCGAACAGGCGCGCCTTGCCGGACGGGAGGACCATGCGGTCGCCGTCGCGGTTCGGGATCGTGTGGTCCAGCAG
ACCCGCCAGCTCGTCCTCACGGGCGCCGTCGAAGACGGGGGTGGCGACGTTGGTGCGGGGGTCGACCTGGTCGGCGCCGATGGCCTGCAGGCGC
TGGGCCCACTCGTCGGCGAGGCCGGAGACGTCCCAGCCCTGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGA
CGGGACGCCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACGCCCT
TGTTGCCGTGGCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCT
CGTCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGGACCTCGCGGGCCT
TCTCGCCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTGGGGGTGACCTTGCCGACCAGGATGTCGCCGGCGACG
ACCTCGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCCGGGCC
GAGCTTGGTGTCGCGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGT
CCTCGTAGTTGTGGCCTTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGGCCGTCGGCGAGGA
CCTGGCCCTCGATGACCCGGTCGCCCTCGGAGACGATGACCTTCTGGTTCACCGAGGTGCCCTGGTTGGAGCGGGAGAACTTCGCGACGCGGTAC
GTGGTGTACGTGCCGTCGTCGTTGGCGACGGTGACGTAGTCGGCCGAGACCTCCTGGATCACACCGTCCTTCTCCGCCTTGATGACGTCGCCGGCG
TCGGTGGCGCAGCGGTACTCCATGCCGGTGCCGACGAGCGGCGCCTCGC-
CTTGATCAGCGGAACGGCCTGGCGCATCATGTTCGCGCCCATCAGGGCGCGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCCA
CCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTC--
CTCGGGCGACGTAGTCGACCTCGCCGCCGCGGCGGCGGACGAGCACGCGGGCCTCGGAGAAGCGCATGTCGTCGCCGAGCGCGGCGTTGGCCTG
CGCGATGACGAAGCGGTCCTCCTCGTCGGCGGTGAGGTAGTCCACCTGGTCGGTGACCTGGCCGCCCTCGACCTTGCGGTACGGCGTCTCGACGA
AGCCGAACGCGTTCACCCGGCCGTAGGTGGCGAGCGAACCGATCAGGCCGATGTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGTCCGTAG
TGGGACGGGTGCACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGACAGACCGCCGGGGGCCCAGCGCCGACAGGCGGCGCTTGTGGGTCAGGC
CCGACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGG
GTCTGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTCTGGATC
AGCTCGCCGACGTTGCGCAGACGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTGGGTGCCGCTCTCGCCGACCGTCTCGGTC
TCGCCGGCGTGCAGCTTGACCAGGTACTTGATGGTCGCGATGACGTCGTCGGTGGTGAGCACGCCGGCGTCCAGCGGCTCGTCGGCGCCGAGCTT
CTTGTTGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTACAGGTTCTCCAGCAGCGTCTGCGCGGCCTCGCGCGTCGG
GGGCTCGCCCGGGCGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATCGACTCGT
ACTCGCCGAACTCCTCCAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTCAGGAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGCACAC
CGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGA
TGGAGGAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATG
AGCGGGAAGTCGCCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGCGGGGCGGCGTACGT
GAAGTCGCGCTCCTTGCACTCGTCGATCGAGTTCTTCGGAGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGA
TCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCGGACTCCAGAGCCGCCTCGACACGAGCCTTCCATGCGG
CGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGC
GGGGCGGTGCTTGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP013743.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGCGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGTTCCTGGAGGGCGTACGC
GGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGCGCCTTACCACCCAGCGGCTGCTGGGTGATCATCGA
GTACGGGCCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGATCGAGATCGGGTCCGGGAAC
GGCTCACCGCTACGGCCGTCGAACAGCCGCGCCTTGCCGGTCGGCTGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGACC
CGCGAGCTCGTCCTCGCGCGCACCGTCGAAGACGGGGGTCGCGACGTTCGTGCCCGGGTCGACCTGGTCGGCGCCGATCACCTGGAGGCGCTGCG
CCCACTCCTCGCCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGG
ACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACACCCTTGTT
GCCGTGCCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGT
CGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCT
CACCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTCGGGGTGACCTTGCCGACGAGGATGTCGCCGGCGATGACCT
CGGCGCCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC

TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGGGCCATCTCGCCGTTCTGCGTGGCGGGACCGTCGGCGAGGACCT
GGCCGGTGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGCGAGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCGGCCTTGACCACGTCACCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATCAGCGGCACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCAC
CGGCGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGCACGCGGTTCTCGGCGAACCGCATGTCGTCGGAGAGCGTGGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGGGTCTCGACGAAACC
GAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCAGGGGTCTCGATCGGGCACATGCGGCCGTAGTGC
GAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGAGTGAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGCGTGGTCATGCGCTCGCGGACGACACGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGC
TCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCGTGCCGTTGTCCCCGGCGGTCTCGGTCTCA
CCGGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCCGCGTCCAGCGGAGCGTCCGCACCCAGCTTCTTG
TTGACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGCGG
CTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATCGACTCGTACTC
GCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTCAGAAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATACGGACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGAGGGGATGATCTTGGCCGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGG
AGGAGTCGAAGTAGACGCCGGGCGAGCGGACCAGCTGGGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAG
CGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCCGCCGTCACGAAGAGCGGGGCGGCGTACGTGA
AGTCGCGCTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCG
GGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCTCCTCGACCCGACTCTGCCAAGCGGTGT
TGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGG
GCGGTGCTGGCGCCGTTGTTCATATTCGCGTTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP011497.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATTCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGCGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACGC
GGCGCCGTAGGCCTCGAGTGCCCACACCTCCATCTCGCCGAACCGCTGGCCACCGAACTGCGCCTTACCACCCAGCGGCTGCTGGGTGATCATCGA
GTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGACC
CGCCAGCTCGTCCTCACGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCCGGG-
GCGACCTGGTCGGCGCCGATGGCCTGCAGGCGCTGGGCCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGT
GGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGCACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTTCCGTCCTCGAGGAACGGCATG
TCCTCGATGGGCAGGATCTTGGAGATGACACCCTTGTTGCCGTGCCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAG
ACGCGGACCAGCTGGTTCACGCCCGGCGGCAGCTCGTCGCCCTCCTCGCGGTCGAAGACGCGCACGCCGATGACCTTGCCGGTCTCGCCGTGCGG
CACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTCACCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGGGTCAGCTCCGTCTCACCCTTCGG
GGTCACCTTGCCGACGAGGATGTCGCCGGCGATGACCTCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGA
CGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCG
TCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGGGCCATC
TCGCCGTTCTGGGTGGCCGGGCCGTCGGCCAGCACCTGGCCGGTGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCC
TGGTTGGACCGGGAGAACTTGGCCAGGCGGTACGTGATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCAC
ACCGTCCTTCTCGGCCTTGACCACGTCACCGGCGTCGACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGGACTTAATCAGCG
GCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCACGGTTGGCGTCGTCGTGCTCCAGGAAGGGGATCATGGCGGTCGCGACCGACACCATC
TGGCGCGGCGAGACGTCCATGTAGTCCACGTCCTCGGGGCCGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGCACGCGGGCCTCGGCGA
ACCGCATCTCGTCGTTCAGCGGCGCGTTGGCCTGCGCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGGC
CCTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCGAACGCGTTGACGCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCT
TCCGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGC
CGAGGGCGTTCAGACGACGCTTGTGCGTCAGCCCCGACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGCGAGGTGCCGAAGAACTCCTTG
ATCGACGCCACGACCGGGCGGATGTTGATCAGGGTCTGCGGCGTGATCGCCTCGACGTCCTGCGTGGTCATGCGCTCGCGCACGACACGCTCCATA
CGGGCGAGACCCGTACGGACCTGGTTCTGGATCAGCTCACCGACGCTGCGGATACGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACCAT
GATCGTGCGGCCGGACTCGGCCACCGTCTCGGTCTCGCCGGCGTGCAGCTTCACCAGGTACTTGATGGTGGCGATGACGTCGTCGGTGGTGAGCA
CGCCGGCGTCCAGCGGCTCGTCCGCACCGAGCTTCTTGTTGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTACAGGT
TCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGGCTCGCCCGGGCGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTG
TGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAGCCGAGCGCCTTCAGCAGCACGGTG
ACCGACTGCTTGCGCTTGCGGTCGATGCGGACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCGCCCCGGGACGGGATGATCTTG
GCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGTGGAGTCGAAGTAGACACCGGGGGAACGGACCAGCTGGGACACCACGACACGCTCGG
TGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCACCGGTCTCGTTGTTGGTGA

ACTCGGCGGTGACGAAGAGCGGGGCGGCGTACGTGAAGTCGCGCTCCTTGCACTCGTCGATGGAGTTCTTCGGCGGCTCGAAGCGGTGGTCCCGG
AACGTCAGGGACATCGACCCGGAGAAGTCCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGTTC
TCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGCAGGTTCGGAACCTCGAGA
GGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTGGCGCCGTTGTTCATATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAG
GGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP040244.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATACCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CAGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATCGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGACGCGCCTTGCCGGAGGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATCGTGTGCTGCAGCAGA
CCCGCGAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTGGTGCCCGGGGCGACCTGGTCGGCGCCGATGGCCTGCAGGCGCT
TCGCCCAGTCGTCCGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGAC
GGGACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCGAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACACCCTT
GTTGCCGTGACGGCCGGCGAGCTTGTCACCATCGGTGATCTTGCGCTTCTGGGCGACGTAGACGCGGACCAGCTGGTTCACGCCCGGGGGAAGCT
CGTCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCT
TCTCGCCGAAGATCGCGCGGAGGAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACG
ACCTCGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCCGGGCC
GAGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGT
CCTCGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCACCGTTCTCGGTGGCCGGACCGTCGGCGAGGA
CCTGGCCCTCGATGACCCGGTCGCCCTCGGAGACGACGACCTTCTGGTTGACCGAGGTGCCCTGGTTCGAGCGGGAGAACTTGGCCATCCGGTACG
TGGTGTACGTGCCGTCGTCGTTGGCGACGGTGATGTAGTCGGCCGAGACCTCCTGGACGACACCGTCCTTGTCGGCCTTCGGCACGTCACCGGCGT
CGGTGGCACAGCGGTACTCCATGCCGGTACCGACGAGCGGGGCCTCGGACTTAATCAGCGGCACCGCCTGACGCATCATGTTCGCGCCCATGAGG
GCACGGTTGGCGTCGTCGTGCTCGAGGAAGGGGATCATGGCCGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACCTCGTC
GGC-
GGGACGTAGTCGACCTCGCCACCACGGCGGCGGACCAGGACGCGCGGCTCGGTGAAGCGGAGCTCGTCGCTCAGGGTCGCGTTCGCCTGGGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTGATGTAGTCGACCTCGTCGGTGACCTGACCGTCGACGACCTTGCGGTACGGCGTCTCGATGAAGCCGA
ACGCGTTGACGCGGCCGTACGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCGGGGGTCTCGATCGGGCACATGCGTCCGTAGTGGGAC
GGGTGCACGTCTCGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCGACA
GCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGC
GGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGCACGACGCGCTCCATACGAGCCAGACCCGTACGGACCTGGTTCTGGATGAGCTCG
CCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTCCCGGCCCGACTCGCCGATCGTCTCGGTCTCGCCC
GCGTGAAGCTTCACGAGGTACTTGATCGTGGCGATGATGTCGTCGGTGGTCAGGACACCGGCGTCCAGCGGCTCGTCCGCGCCGAGCTTCTTGTTC
ACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGAGCGGCCTCACGGGTCGGCGGCTC
GCCCGGACGGAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGC
CGAACTCCTGCAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTCAGGAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGACC
ATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCCGAGAAGATGTCCTTGTCGGACGTCTTGTCGATCGAG
GAGTCGAAGTAGACGCCCGGCGAACGGACCAGCTGCGAGACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAGCG
GGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGCGGGGCGCCGTACGTGAAG
TCACGGTCCTTGCACTCGTCGATCGAGTTCTTCGGGGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGG
AGAGATCTCCTCGAAGATCTCCTCCAGACCCGACTTCCTGGGGACGTCCTGTCCCGACTCGAGAGCTGCCTCGACGCGAGCCTTCCAGGCGGCATTG
CCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGGAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGC
GGTGCTGGCGCCGTTGTTCGTATT---GGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP023703.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCTTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCATAAGCCTCAAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAACGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCACCCGAGCGGCCGTCGAACAGCGGGGCCTTGCCGGTCGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGCAGCAGA
CCGGTCAGCTCGTCCTCGCGGGCACCGTCGAACACCGGGGTGGCGACGTTGGTACCGGGCGCGACGCTGTCGGCACCGATGA-
CTGGAGCCGCTGGGCCCAGTCCTGCGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGAACCTGTCCCGGGT
TCATTCGGGACGGCACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAG
ATGACACCCTTGTTGCCGTGGCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGAACCAGCTGGTTCACGCCCG
GCGGCAGCTCGTCGCCCTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCACCGTGCGGCACCTTCAGCGAGGTGTCGCGCACTT

CGCGCGCCTTCTCACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGC
CGGCGACGACCTCGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCC
TCGGGGCCGAGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGAT
GATCGCGTCCTCGTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTGGCCGGTCCGTCG
GCGAGGACCTGACCGGTGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTCACCGAGGTTCCCTGGTTGGACCGGGAGAACTTGGCCAGG
CGGTACGTGATGTACGTGCCGTCGTCGTTGGCCGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCGGCCCTGACCACGTCG
CCGGCGTCGACGGCGGAGCGGTACTCCATGCCGGTGCCGACCAGCGGCGCCTCGCTCTTGATGAGCGGGACGGCCTGACGCATCATGTTCGCGCC
CATGAGGGCACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGGGAGACGTCCATGTAGTCG
ACGTCGTCACCGGGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGGACGCGGGCCTCCTCGAACTGGAACTCGTCACTGAGCGGCGCGTT
GGCCTGCGCGATGACGAACCGGTCCTCTTCGTCGGCCGTGAGGTAGTCGACCTCGTCGGTGACGACACCGCCGGTGACCTTGCGGTACGGGGTCT
CGATGAAACCGAACGGGTTGATCCGGCCGTACGAGGCGAGCGAACCGATCAGGCCGATGTTCGGGCCTTCCGGCGTCTCGATCGGGCACATGCGG
CCGTAGTGCGACGGGTGAACGTCGCGGACCTCGAAGCCGGCCCGCTCACGGGACAGACCACCCGGGCCGAGCGCGTTGAGACGACGCTTGTGCG
TCAGACCCGACAGCGGGTTGTTCTGGTCCATGAACTGGGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTG
ATCAGGGTCTGCGGCGTGATCGCCTCGACGTCCTGAGTGGTCATGCGCTCACGCACGACCCGCTCCATACGGGCGAGACCCGTACGGACCTGGTTC
TGGATCAGCTCGCCGACGCTGCGGATACGACGGTTGCCGAAGTGGTCGATGTCATCGGTCTCGACCATGATCGAGCGACCCGACTCACCGAGCGTC
TCGGTCTCACCGGCGTGCAGCTTCACCAGGTACTTGATCGTGGCGATGATGTCGTCGGTGGTGAGCACACCGGCGTCCAGCGGCTCTTCCGCGCCG
AGCTTCTTGTTCACCTTGTAGCGGCCGACCTTCGCGAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGAGCAGCCTCGCGG
GTCGGCGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATCGA
CTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTCAGGAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCG
GACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTT
GTCGATGGAGGAGTCGAAGTAGACGCCGGGCGAGCGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTG
GTCATGAGCGGGAAATCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTCGTGAACTCGGCCGTCACGAAGAGCGGGGCCGC
GTACGTGAAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGT
CCTCGATCGGGGAGATCTCCTCGAAGATCTCTTCGAGGCCGGACTTGGTGGGGACGTCCTGACCGTTCTCGAGAGCCTCCTCGACGCGAGCCTTCC
ACGCTTCGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGGGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATG
CGCAGCGGGGCAGTGCTGGCAGCATTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCA
C

>gb|CP033073.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGCGACAGGTCGATACCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATTTCCTTGATGAGCACCTTGAAGGACTCGGGAATGCCGGGTT
CCGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTGAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAGGCCTCGAGTGCCCACACCTCCATCTCGCCGAACCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAA
CGGCTCACCGCTACGGCCGTCGAACAGCCGCGCCTTGCCGGTCGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGGAGCAGAC
CCGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTGGCGACGTTGGTGCCGGGGTCGACCTGGTCGGCGCCGATGACCTGCAGGCGCTG
GGCCCACTCGTCGGCGAGACCGGAGACGTCCCAGCCTCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATGCGGGACG
GCACACCCAGCGGGTTCAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACACCCTTGT
TGCCGTGCCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGGACCAACTGGTTCACGCCCGGCGGCAGCTCGT
CGCCCTCCTCGCGGTCGAAGAGGCGCACACCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCT
CACCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGACGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGGGCCATCTCACCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCT
GGCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTG
ATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCG
ACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGGACTTGATCAGCGGCACGGCCTGCCGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCCTCGG
GGCCGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGCACGCGGGCCTCGGCGAACCGCATGTCGTCACCGAGCGTGGCGTTGGCCTGCGC
GATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCGTCGCGGACCTTGCGGTACGGGGTCTCGACGAAAC
CGAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGGCCGATGTTCGGGCCTTCCGGCGTCTCGATCGGGCACATGCGGCCGTAGTGC
GAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCCAGCGCCGACAGACGGCGCTTGTGGGTGAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGACAACTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTGGTCATGCGCTCGCGGACGACGCGCTCCATACGCGCCAGGCCCGTACGGACCTGGTTCTGGATGAG
CTCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGAGTCGCCGTTGTCACCGGTGGTCTCGGACTC
GCCGGCGTGGAGCTGCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCCGCGTCGAGCGGGGCGTCCGCACCCAGCTTCT
TGTTGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTAAAGGTTCTCCAGCAGCGTCTGCGCGGCCTCACGCGTGGGG
GGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATCGACTCGTAC
TCGCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAGCCGAGCGCCTTGAGCAGCACGGTGACCGACTGCTTGCGCTTGCGGTCGATACGGACACCG

ACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGAAGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATG
CTGGAATCGAAGTAGACACCGGGGGAACGGACCAACTGGGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAG
CGGGAAGTCACCCATGAAGACGGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGCGGGGCGGCGTACGTGA
AGTCGCGCTCCTTGCACTCGTCGATGGAGTTCTTGGGAGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATC
GGGGAGATCTCCTCGAAGATCTCTTCGAGGCCGGACTTGGTGGGGACGTCCTGACCGGACTCCAGAGCCGCCTCGACGCGAGCCTTCCAGGCGTC
GTTGCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCGAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAAATGCGCAGCG
GGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP028834.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGCGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTACACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAGGC
GGCGCCGTAGGCCTCCAGCGCCCACACCTCCATCTCGCCGAACCGCTGGCCGCCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCGA
GTACGGACCGGTGGAGCGGGCGTGCAGCTTGTCGTCGACAAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAAC
GGCTCACCGCTGCGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGCAGCAGACC
CGCGAGCTCGTCCTCGCGCGCACCGTCGAAGACGGGGGTCGCGACGTTCGTGCCCGGGTCGACCTGGTCGGCGCCGATCGCCTGCAGGCGCTGCG
CCCACTCCTC-
GCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGGACACCCAGCG
GGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTGGAGATGACGCCCTTGTTGCCGTGCCGG
CCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGTGGAAGCTCGTCGCCCTCCTCGC
GGTCGAAGACGCGCACACCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTCACCGAAGATCG
CGCGCAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCGCCCTTGGGGGTCACCTTGCCGACGAGGATGTCGCCGGCGATGACCTCGGCACCGATG
CGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCTTGGTGTCACG
GGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCGTAGTTGTGAC
CCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGGGCCATCTCGCCCTCCTGCGTGGCGGGACCGTCGGCGAGGACCTGGCCCTCGACG
AGGCGGTCGCCCTCGCTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGATGTACGTGCC
GTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGACCGCAGAGCG
GTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATCAGGGCACGGTTGGCGT
CGTCGTGCTCCAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGGGAGACGTCCATGTAGTCCACGTCGTCACCGGCGACGTAG
TCGACCTCGCCGCCACGGCGGCGGACCAGGACGCGGTTCTCGGCGAAGCGCAGGTCGTCCGTCAGCGTGGCGTTGGCCTGCGCGATGACGAAGC
GGTCCTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGACGAAACCGAACGCGTTGA
CCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGAGGGGTGCAC
GTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGGCGGCGCTTGTGGGTGAGACCCGACAGCGGGTTG
TTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCGGCGTGAT
CGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACTCGCTCCATACGCGCCAGGCCCGTACGGACCTGGTTCTGGATGAGCTCGCCGACGCT
GCGCAGACGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCTGGCCGTTGTCGGCGGTCGTCTCGATCTCGCCGGCGTGCA
GCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCCGCGTCCAGCGGAGCGTCCGCACCCAGCTTCTTGTTGACCTTGTA
GCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGGCTCGCCCGGA
CGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATCGACTCGTACTCGCCGAACTCC
TCGAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTCAGGAGCACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGACCATGTCGCG
CTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTCGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGAGGAGTCGAA
GTAGACACCCGGCGAACGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGCGGGAAGTCGC
CCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGCGGAGCCGCGTACGTGAAGTCGCGCTCCT
TGCACTCGTCGATCGAGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGGGAGATCTCCT
CGAAGATCTCTTCGAGGCCGGACTTGGTGGGGACATCCTGACCGGATTCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGTTGCCGAGCAGCC
AGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAGAGATGCGCAGCGGGGCGGTGCTGGCG
CCGTTGTTCGTATTCGCGGTCGAGGCATTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP034539.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCTGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGCGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACGC
GGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATGCT
GTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCGCTACGGCCGTCGAACAGCCGCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCCCGGTTCGGGATGGTGTGCTGCAGCAGACC
CGCGAGCTCGTCCTCGCGCGCACCGTCGAAGACGGGGGTCGCGACGTTCGTGCCCGGGTCGACCTGGTCGGCGCCGATCGCCTGGAGGCGCTGCG
CCCACTCCTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGC
ACACCCAGCGGGTTGAGGATGATGTCGACCGGAGTTCCGTCCTCGAGGAACGGCATGTCCTCGATGGGCAGGATCTTCGAGATGACACCCTTGTTG

CCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTACACGCGCACCAGCTGGTTGACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTC
ACCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTCGGGGTCACCTTGCCGACGAGGATGTCACCGGCGATGACCTC
GGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCT
TGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTCG
TAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCACCGTTCTCGGTGGCCGGACCGTCGGCGAGGACCTGG
CCCTCGATGATGCGGTCGCCCTCGAGGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAACGGGCGAACTTGGCCAGGCGGTACGTGAT
GTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGGCCTTCTCGGCCTTGACGACGTCGCCGGCGTCGAC
TGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGC-
CTTGATGAGCGGAACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGCACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGA
CCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCCTCGCCGCCGACGTAGTCGACCTCGCCGCCACGGCGGCGAACCAGGACGCGG
GCCTCGGCGAACCGCAGGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGATGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTC
GGTGACCTGGCCGTCGACGACCTTGCGGTACGGCGTCTCGACGAAGCCGAACGCGTTGACCCGGCCGTAGGAGGCCAGCGAGCCGATCAGACCG
ATGTTCGGGCCTTCGGGCGTCTCGATCGGACACATACGGCCGTAGTGCGACGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAG
ACCACCCGGGCCAAGAGCCGACAGGCGGCGCTTGTGGGTGAGACCCGACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGA
AGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACG
ACACGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAGCTCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTC
GGTCTCGACGACGATCGTCTGGCCGTTGTCCGCGGCCGTCTCGGTCTCACCGGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGACGTCCTCG
ACGGTCAGGATGCCCGCGTCCAGCGGAGCGTCCGCACCCAGCTTCTTGTTGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTG
AAGTACAGGTTCTCCAGAAGCGTCTGCGCGGCCTCACGCGTGGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTG
GCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTCCAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTGAG
CAGAACGGTCACGGACTGCTTGCGCTTGCGGTCGATGCGCACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGG
GATGATCTTGGCCGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGCTGGAGTCGAAGTAGACACCGGGCGAACGGACCAGCTGCGACACCACGA
CACGCTCGGTGCCGTTGATGACGAACGTGCCCTTGTTCGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGT
TGTTGGTGAACTCAGCCGTCACGAAGAGCGGAGCGGCGTACGTGAAGTCGCGCTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAGCGGT
GGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACATCC
TGACCGGACTCCAGAGCCGCCTCGACCCGACTCTGCCAGGCGGTGTTGCCGAGCAGCCAGTCAAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGG
AACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGG
CGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|KX503595.1|

GACCTCTTCGACGCTGCTCGGCTCGCGCCGTGACAGGTCGATGCCGAGCTCTTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTCGAT
GGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTCGG
GAATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACACGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACGCGG
CGCCATAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATGGAGT
ACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAACGG
CTCGCCGGAGCGGCCGTCGAACATCCTCGCCTTACCGGTCGGCTGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGGTTGAGCAGACCCGC
GAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTCGTGCCCGGGTCGACCTGGTCGGCACCGATGACCTGCAGGCGCTTCGCCC
AGTCCTCGCCGAGCCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGGACA
CCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTGTTGCC
GTGCCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGGGCGACGTACACCCGGACCAGCTGGTTGACGCCGGGCGGCAGCTCGTCGC
CCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTCAC
CGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCTCG
GCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCTT
GGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGTACGAGGCGCTGCGACAGGATGATCGCGTCCTCGT
AGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTGGC
CCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGATGT
ACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACGACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGACCG
CGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCACGG
TTGGCGTCGTCGTGCTCCAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCACCGGG
AACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACGCGGTTCTCGGTGAAACGCATGTTGTCGTCGAGCAGGGCGTTGGCCTGCGCGATGA
CGAAGCGGTCCTCTTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACGACGTCGCCGTCGACCTTGCGGTACGGCGTCTCCACGAAACCGAACG
CGTTGACGCGGCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGTCCGTAGTGGGACGG
GTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGACAGACCACCGGGACCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCCGACAGC
GGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCGG
CGTGATCGCCTCGACGTCCTGGGTCGTCATCCGCTCGCGGACGACTCGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAGCTCGCC
GACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTGCCG-
CCGCTGTCGCCGACCGTCTCGGTCTCGCCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCCGCGTTGA

GCGGAGCCTCCCCACCCAGCTTCTTGTTGACCTTGTAACGGCCGACCTTCGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCAAGCAGCG
TCTGCGCGGCCTCACGCGTGGGCGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCA
GGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTCCAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTGACGGACTGCTTG
CGCTTGCGGTCGATACGGACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGAGGAGAAGATG
TCCTTGTCGGACGTCTTGTCGATCGAGGAGTCGAAGTAGACGCCCGGCGAACGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGAC
GAAGGTGCCCTTGTGGGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCACCGGTCTCGTTGTTGGTGAACTCGGCCGTGAC
GAAGAGCGGGGCCGCGAAGGTGAAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAGCGGTGGTCACGGAAGGTCAGGGAC
ATCGACCCGGAGAAGTCCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCGTTCTCCAGAGCCTCCT
CGACCCGACTCTGCCAGGCGGTGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATC
TTTGCGAAAGAGATGCGCAGCGGGGCGGTGCTGGCGGCGGGGTTCGTATTCGAGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCC
GAGGGCTCGGACTCAC

>gb|CP050975.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCACGCATTTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCAGAGAGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGACCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCGGCGCCTTGCCGGTCGGGAGCACCATGCGCTCGCCGTCACGGTTCGGGATCGTGTGGTTCAGCAGAC
CGGCCAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTGGTACCGGGGGCGACCACGTCGGCCTCGATGGCCTGCAGACGCTGC
GCCCAGTCCTCCCCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCGAGGGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTGT
TGCCGTGGCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGAACCAGCTGGTTCACGCCCGGCGGCAGCTCGT
CGCCCTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGCACCTCGCGCGCCTTCT
CACCGAAGATCGCGCGGAGCAGGCGCTCCTCGGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACC
TCGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTCGCCGGGCCGTCGGCCAGTACCT
GGCCGGCGACGACCCGGGCGCCCTCGTCCACGACAACCTTCTGGTTGACGGACGTGCCCTGGTTGGAGCGGGAGAACTTGTGCAGGCGGTACGTG
GTGTACGTGCCGTCGTCGTTGGCGGTGGTGACGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCCGACTTGAGCACGTCGCCGGCGTCG
ACGGCGCAGCGGTACTCCATGCCGGTGCCGACGAGCGGCGCCTCGGACTTGATGAGCGGAACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCCTCACC
GGGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGGACGCGGGACTCCTCGAAGCGGAGCTCGTCGGTCAGCGGCGCGTTGGCCTGCGCG
ATGACGAACCGGTCCTCTTCGTCGGCCGTCAGGTAGTCGACGTCGTCGGTGACGACACCGCCCGTGACCTTGCGGTACGGGGTCTCGACGAAACCG
AACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCAGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGA
GGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCGAGCGCCGACAGACGGCGCTTGTGGGTGAGACCCGAC
AGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTG
CGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGCACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTC
GCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGAGGTGCCGTTGTTACCAACGGTCTCGGTCTCACC
GGCGTGCAGCTTCACCAGGTACTTGATCGACGAGATGATGTCCTCGACGGTCAGGATGCCGGCGTCCAGCGGCTCG-CCGC-CC-
AGCTTCTTGTTGACCTTGTACCGGCCGACCTTGGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCCAGAAGCGTCTGCGCGGCCTCACGG
GTCGGCGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGA
CTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTCAGGAGAACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCG
GACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCGCCCCGGGACGGGATGACCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTT
GTCGATGGAGGAGTCGAAGTAGACGCCCGGCGAGCGGACCAGCTGGGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTG
GTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGC
GTACGTGAAGTCGCGGTCCTTGCACTCGTCGATGGAGTTCTTCGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGCGACATCGACCCGGAGAAGT
CCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTCGTGGGGACGTCCTGTCCGCTGTCCAGAGCCGCCTCGACACGAGCCTTCCA
CGCTTCGTTTCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCAAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCG
CAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP022744.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCATCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGATGGAGATCGGGTCCGGGAA
CGGCTCACCGCTACGGCCGTCGAACAGCCGCGCCTTACCGGTCGGCTGCACCATGCGGTCGCCGTCGCGGTTCGGGATGGTGTGGTTCAGCAGAC

CGGCGAGCTCGTCCTCACGCGCACCGTCGAACACCGGGGGTGGCGACGTTGGTGCCGGCGTCGACCTGGTCGGCGCCGATCACCTGGAGGCGCTGC
GCCCACTCCTCGCCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
CACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTGTT
GCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGCACCTCGCGCGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCGGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGACGACCT
CGGCACCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTGGGTGGCCGGACCGTCGGCCAGGACCTG
GCCATCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGAT
GTACGTACCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGAC
CGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGCGCCTCGGACTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCACCG
GCGACGTAGTCGACCTCGCCACCACGACGACGGACGAGCACGCGGTTCTCGGCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGAT
GACGAATCGGTCCTCTTCGTCGGCCGTGAGGTAGTCGACCTCGTCGGTGACGACACCAGCGGTGACCTTGCGGTACGGCGTCTCCACGAAACCGAA
CGCGTTGACGCGGCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCAGGGGTCTCGATCGGGCACATGCGGCCGTAGTGGGAC
GGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGCGACAGACCACCGGGGCCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCCGACA
GCGGGTTGTTCTGGTCCATGAACTGCGAGAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGC
GGCGTGATCGCCTCGACGTCCTGGGTCGTCATCCGCTCACGGACGACTCGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAGCTCG
CCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGAGCCACCGCTGTCGCCGACCGTCTCCGTCTCGCCC
GCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGACCCCGGCGTTCAGCGGCTCG-CCGC-CC-
AGCTTCTTGTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGAGTTGAAATAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGC
GTCGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATCGA
CTCGTACTCGCCGAACTCCTCCAGGATCTGCTCGGTGGTCCACCCGAGAGCCTTGAGCAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATACG
GACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTT
GTCGATCGAGGAGTCGAAGTAGACGCCCGGCGAACGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGCCG
GTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGTGGGGCGGC
GTACGTGAAGTCGCGCTCCTTGCACTCGTCGATGCTGTTCTTCGGGGGTTCGAAACGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGT
CCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCACTCTCCAGAGCCTCCTCGACCCGACTCTGCCA
AGCGGTGTTGCCGAGGAGCCAGTCAAAGCTCTCGGTCTGCAACGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAGAGATGC
GCAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|KX503607.1|

GACCTCTTCGACGCTGCTCGGCTCGCGCCGTGACAGGTCGATGCCGAGCTCTTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTCGAT
GGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTCGG
GAATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACACGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACGCGG
CGCCATAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATGGAGT
ACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAACGG
CTCGCCGGAGCGGCCGTCGAACATCCTCGCCTTACCGGTCGGCTGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGGTTGAGCAGACCCGC
GAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTCGTGCCCGGGTCGACCTGGTCGGCACCGATGACCTGCAGGCGCTTCGCCC
AGTCCTCGCCGAGCCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGGGACA
CCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTGTTGCC
GTGCCGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGGGCGACGTACACCCGGACCAGCTGGTTGACGCCGGGCGGCAGCTCGTCGC
CCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCTCAC
CGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACCTCG
GCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGCTT
GGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGTACGAGGCGCTGCGACAGGATGATCGCGTCCTCGT
AGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTGGC
CCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGATGT
ACGTGCCGTCGTCGTTGGTGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACGACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGACCG
CGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGCTCTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCACGG
TTGGCGTCGTCGTGCTCCAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCACCGGG
AACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACGCGGTTCTCGGTGAAACGCATGTTGTCGTCGAGCAGGGCGTTGGCCTGCGCGATGA
CGAAGCGGTCCTCTTCGTCGGCCGTCAGGTAGTCGACCTCGTCGGTGACGACGTCGCCGTCGACCTTGCGGTACGGCGTCTCCACGAAGCCGAACG
CGTTGACGCGGCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGTCCGTAGTGGGACGG
GTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGACAGACCACCGGGACCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCCGACAGC
GGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGCGG
CGTGATCGCCTCGACGTCCTGGGTCGTCATCCGCTCGCGGACGACTCGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAGCTCGCC

GACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCTGCCG-
CCGCTGTCGCCGACCGTCTCGGTCTCGCCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCCGCGTTGA
GCGGAGCCTCCCCACCCAGCTTCTTGTTGACCTTGTAACGGCCGACCTTCGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCAAGCAGCG
TCTGCGCGGCCTCACGCGTGGGCGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCA
GGGTGGCGCGCATGGACTCGTACTCGCCGAACTCCTCCAGGATCTGCTCGGTGGTCCAGCCGAGGGCCTTGAGGAGGACGGTGACGGACTGCTTG
CGCTTGCGGTCGATACGGACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGAGGAGAAGATG
TCCTTGTCGGACGTCTTGTCGATCGAGGAGTCGAAGTAGACGCCCGGCGAACGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGAC
GAAGGTGCCCTTGTGGGTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCACCGGTCTCGTTGTTGGTGAACTCGGCCGTGAC
GAAGAGCGGGGCCGCGAAGGTGAAGTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAGCGGTGGTCACGGAAGGTCAGGGAC
ATCGACCCGGAGAAGTCCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCGTTCTCCAGAGCCTCCT
CGACCCGACTCTGCCAGGCGGTGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATC
TTTGCGAAAGAGATGCGCAGCGGGGCGGTGCTGGCGGCGGGGTTCGTATTCGAGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCC
GAGGGCTCGGACTCAC

>gb|CP016438.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCATCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGATGGAGATCGGGTCCGGGAA
CGGCTCACCGCTACGGCCGTCGAACAGCCGCGCCTTACCGGTCGGCTGCACCATGCGGTCGCCGTCGCGGTTCGGGATGGTGTGGTTCAGCAGAC
CGGCGAGCTCGTCCTCACGCGCACCGTCGAACACCGGGGGTGGCGACGTTGGTGCCGGCGTCGACCTGGTCGGCGCCGATCACCTGGAGGCGCTGC
GCCCACTCCTCGCCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
CACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTGTT
GCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGCACCTCGCGCGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCGGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGACGACCT
CGGCACCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTGGGTGGCCGGACCGTCGGCCAGGACCTG
GCCATCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGAT
GTACGTACCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGAC
CGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGCGCCTCGGACTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCACCG
GCGACGTAGTCGACCTCGCCACCACGACGACGGACGAGCACGCGGTTCTCGGCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGAT
GACGAATCGGTCCTCTTCGTCGGCCGTGAGGTAGTCGACCTCGTCGGTGACGACACCAGCGGTGACCTTGCGGTACGGCGTCTCCACGAAACCGAA
CGCGTTGACGCGGCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCAGGGGTCTCGATCGGGCACATGCGGCCGTAGTGGGAC
GGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGCGACAGACCACCGGGGCCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCCGACA
GCGGGTTGTTCTGGTCCATGAACTGCGAGAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGC
GGCGTGATCGCCTCGACGTCCTGGGTCGTCATCCGCTCACGGACGACTCGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAGCTCG
CCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGAGCCACCGCTGTCGCCGACCGTCTCCGTCTCGCCC
GCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGACCCCGGCGTTCAGCGGCTCG-CCGC-CC-
AGCTTCTTGTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGAGTTGAAATAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGC
GTCGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATCGA
CTCGTACTCGCCGAACTCCTCCAGGATCTGCTCGGTGGTCCACCCGAGAGCCTTGAGCAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATACG
GACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTT
GTCGATCGAGGAGTCGAAGTAGACGCCCGGCGAACGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGCCG
GTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGTGGGGCGGC
GTACGTGAAGTCGCGCTCCTTGCACTCGTCGATGCTGTTCTTCGGGGGTTCGAAACGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGT
CCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCACTCTCCAGAGCCTCCTCGACCCGACTCTGCCA
AGCGGTGTTGCCGAGGAGCCAGTCAAAGCTCTCGGTCTGCAACGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAGAGATGC
GCAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP046024.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCATCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG

AGTACGGACCGGTCGACCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGATGGAGATCGGGTCCGGGAA
CGGCTCACCGCTACGGCCGTCGAACAGCCGCGCCTTACCGGTCGGCTGCACCATGCGGTCGCCGTCGCGGTTCGGGATGGTGTGGTTCAGCAGAC
CGGCGAGCTCGTCCTCACGCGCACCGTCGAACACCGGGGGTGGCGACGTTGGTGCCGGCGTCGACCTGGTCGGCGCCGATCACCTGGAGGCGCTGC
GCCCACTCCTCGCCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
CACACCCAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTGTT
GCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGTC
GCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGCACCTCGCGCGCCTTCTC
ACCGAAGATCGCGCGGAGCAGGCGCTCCTCGGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGACGACCT
CGGCACCGATCCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTGGGTGGCCGGACCGTCGGCCAGGACCTG
GCCATCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGAT
GTACGTACCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCGGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGAC
CGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGCGCCTCGGACTTGATGAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCACCG
GCGACGTAGTCGACCTCGCCACCACGACGACGGACGAGCACGCGGTTCTCGGCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCGCGAT
GACGAATCGGTCCTCTTCGTCGGCCGTGAGGTAGTCGACCTCGTCGGTGACGACACCAGCGGTGACCTTGCGGTACGGCGTCTCCACGAAACCGAA
CGCGTTGACGCGGCCGTAGGAGGCGAGCGAACCGATCAGACCGATGTTCGGGCCTTCAGGGGTCTCGATCGGGCACATGCGGCCGTAGTGGGAC
GGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGCGACAGACCACCGGGGCCCAGCGCCGAGAGACGACGCTTGTGCGTCAGACCCGACA
GCGGGTTGTTCTGGTCCATGAACTGCGAGAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTGC
GGCGTGATCGCCTCGACGTCCTGGGTCGTCATCCGCTCACGGACGACTCGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAGCTCG
CCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGAGCCACCGCTGTCGCCGACCGTCTCCGTCTCGCCC
GCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGACCCCGGCGTTCAGCGGCTCG-CCGC-CC-
AGCTTCTTGTTCACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGAGTTGAAATAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGC
GTCGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATCGA
CTCGTACTCGCCGAACTCCTCCAGGATCTGCTCGGTGGTCCACCCGAGAGCCTTGAGCAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATACG
GACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTT
GTCGATCGAGGAGTCGAAGTAGACGCCCGGCGAACGGACCAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGCCG
GTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGTGGGGCGGC
GTACGTGAAGTCGCGCTCCTTGCACTCGTCGATGCTGTTCTTCGGGGGTTCGAAACGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGT
CCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGTCCACTCTCCAGAGCCTCCTCGACCCGACTCTGCCA
AGCGGTGTTGCCGAGGAGCCAGTCAAAGCTCTCGGTCTGCAACGCGAGCAGGTTGGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAGAGATGC
GCAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP045643.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCACGCATTTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCAGAGAGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGTTCCTGGAGGGCGTAAG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCGGCGCCTTGCCGGTCGGGAGCACCATGCGCTCGCCGTCACGGTTCGGGATCGTGTGCTGCAGCAGAC
CCGCGAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTGGTGCCCGGGACGACCTGGTCGGCGCCGATGGCCTGCAGGCGCTG
CGCCCATTCCTCGCCGAGTCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACG
GGACACCGAGGGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTG
TTGCCGTGGCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGAACCAGCTGGTTCACGCCCGGCGGCAGCTCG
TCGCCCTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGCACCTCGCGCGCCTTCT
CACCGAAGATCGCGCGGAGCAGACGCTCCTCGGGGGTCAGCTCGGTCTCGCCCTTGGGCGTGACCTTGCCGACGAGGATGTCGCCGGCGACGACC
TCGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTCGCCGGGCCGTCGGCCAGAACCT
GGCCGGCGATGACCCGGGCGCCCTCGTCGACGACAACCTTCTGGTTGACGGACGTGCCCTGGTTGGAGCGGGAGAACTTGTGCAGGCGGTACGTG
GTGTACGTGCCGTCGTCGTTGGCGGTGGTGACGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCCGACTTGAGCACGTCGCCGGCGTCG
ACGGCGCAGCGGTACTCCATGCCGGTGCCGACGAGCGGTGCCTCGGACTTGATGAGCGGAACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCCAGGAACGGGATCATCGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCCTCACC
GGGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGAACGCGGGACTCCTCGAAGCGGAGCTCGTCGGTCAGCGGCGCGTTGGCCTGCGCG
ATGACGAACCGGTCCTCTTCGTCGGCCGTCAGGTAGTCGACGTCGTCGGTGACGATGCCGCCCGTGACCTTGCGGTACGGCGTCTCGACGAAACCG
AACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCAGGGGTCTCGATCGGGCACATGCGGCCGTAGTGCG
AGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCGA
CAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCT

GCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGCACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCT
CGCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGAGGTGCCGTTGTCGCCAACGGTCTCGGTCTCAC
CGGCGTGCAGCTTCACCAGGTACTTGATCGACGAGATGATGTCCTCGACGGTCAGGATGCCGGCGTCCAGCGGCTCG-CCGC-CC-
AGCTTCTTGTTGACCTTGTAGCGGCCGACCTTGGCCAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGG
GTCGGGGGCTCACCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGA
CTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGAGCCTTCAGGAGAACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCG
GACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCGCCCCGGGAGGGGATGACCTTCGCGGAGAAGATGTCCTTGTCGGACGTCTT
GTCGATGGAGGAGTCGAAGTAGACACCCGGCGAGCGGACCAGCTGGGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTG
GTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGC
GTACGTGAAGTCGCGGTCCTTGCACTCGTCGATGGAGTTCTTCGGGGGCTCGAAGCGGTGGTCGCGGAACGTCAGGGACATCGACCCGGAGAAGT
CCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTCGTGGGGACGTCCTGTCCGCTGTCCAGAGCCGCCTCGACACGAGCCTTCCA
CGCTTCGTTTCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCAAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGCG
CAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP022161.1|

TCAGACCTCTTCGACGCTGCTCGGCTCACGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACGTCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCATCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAACGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTGAGGATGTACATGTACCCGACCGAGATCGGGTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGGCGCGCCTTGCCGGAGGGGAGCACCATCCGCTCACCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGA
CCGGCCAGCTCGTCCTCGCGGGCACCGTCGAAGACCGGGGTCGCGACGTTGGTGCCGGGGGCGACGGAGTCGGCGCCGATCGCCTGGAGGCGCT
GCGCCCACTCGTCCGCGAGGCCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGAC
GGGACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTCCCGTCCTCGAGGAACGGCATGTCCTCGATGGGCAGGATCTTGGAGATGACACCCTT
GTTGCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCCACGTACACGCGCACCAGCTGGTTGACACCGGGGGGAAGCTC
GTCGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTT
CTCACCGAAGATCGCGCGGAGCAGGCGCTCCTCGGGCGTCAGCTCGGTCTCACCCTTCGGGGTGACCTTGCCGACGAGGATGTCACCGGCGATGA
CCTCGGCACCGATACGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCG
AGCTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGGACGAGGCGCTGCGACAGGATGATCGCGTC
CTCGTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGCGCCATCTCACCGTTCTGGGTGGCCGGACCGTCGGCGAGGAC
CTGGCCCTCGATGATGCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGT
GATGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCAGAGACCTCCTGGACCACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTC
GACGGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGGACTTAATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGG
GCACGGTTGGCGTCGTCGTGCTCGAGGAAGGGGATCATGGCCGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTC
ACCGGCGACGTAGTCGACCTCGCCACCACGACGACGAACGAGGACACGGTTCTCGGCGAACCGCATGTCGTCGTTGAGCGTGGCGTTGGCCTGCG
CGATGACGAACCGGTCCTCTTCGTCGGCCGTCAGGTAGTCCACCTCGTCGGTGACCTGACCGCCTTCGACCTTGCGGTACGGCGTCTCCACGAAACC
GAACGCGTTGACGCGGCCGTAGGAGGCGAGCGAACCGATCAGGCCGATGTTCGGACCTTCGGGGGTCTCAATCGGGCACATACGGCCGTAGTGC
GACGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGCGACAGACCACCGGGACCCAGTGCCGAGAGACGACGCTTGTGCGTCAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGAGAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCACGGACGACACGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAGC
TCGCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATGTTGCCGCCGCTGTCGCCGACCGTCTCGGTCTCG
CCCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCCGCGTCGAGCGGGGCGTCCGCACCCAGCTTCTTG
TTGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTACAGGTTCTCGAGGAGCGTCTGCGCGGCCTCACGCGTGGGGGG
CTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGCCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTC
GCCGAACTCCTCCAGGATCTGCTCGGTCGTCCAACCGAGGGCCTTGAGGAGGACGGTCACGGACTGCTTGCGCTTGCGGTCGATACGGACACCGA
CCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCGCCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATCG
AGGAGTCGAAGTAGACACCGGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAACGTGCCCTTGCCCGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCCCCGGTCTCGTTGTTGGTGAACTCGGCCGTGACGAAGAGCGGGGCGGCGTACGTGAA
GTCGCGCTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAGCGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGATCGG
GGAGATCTCCTCGAAGATCTCCTCCAGGCCGGACTTGGTGGGGACGTCCTGTCCGTTCTCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGTTG
CCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGCAGGTTCGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAGGAGATGCGCAGCGGGGC
GGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP075896.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCAAGCTCCTCCGCAGCGCGGAAGACATCCTCGTCGGTGTCGCGCATCTC
GATGGACATGCCGTCCGAGGACAGCACCTCCACGTTGAGGCACAGGGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGAATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGACCGGTGACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTACG
CGGCGCCGTAAGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG

AGTACGGACCGGTCGAACGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAAC
GGCTCGCCGGAGCGGCCGTCGAACAGCGGCGCCTTACCGGTCGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATCGTGTGGTTCAGCAGACC
CGCCAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTGGTGCCGGGCTCGACCTTGTCGGCCTCGATGGCCTGGAGGCGCTGCG
CCCAGTCGTCCGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGACCCGGGTTCATTCGGGACGGC
ACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACGCCCTTGTT
GCCGTGACGGCCGGCGAGCTTGTCACCGTCGGTGATCTTGCGCTTCTGCGCGACGTAGACGCGCACCAGCTGGTTCACACCGGGGGGAAGCTCGT
CGCCCTCCTCGCGGTCGAAGACGCGGACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCT
CACCGAAGATCGCGCGCAGCAGGCGCTCCTCGGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGATGACCT
CGGCGCCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCCAGCGCCATCTCGCCCTGCTGGGTGGCCGGACCGTCCGCGAGGATCTG
GCCCTCGATGATCCGGTCGCCCTCGCCGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGA
TGTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACCACACCCGGCTTCTCGGCCTTGACCACGTCGCCGGCGTCGG
CCGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGCCTCGGACTTAATCAGCGGCACGGCCTGACGCATCATGTTCGCGCCCATGAGGGCA
CGGTTGGCGTCGTCGTGCTCCAGGAAGGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCTTCGCC
GGTGACGTAGTCGACCTCGCCGCCCTTACGGCGGACCAGCACGCGGGCCTCGGCGAAGCGGAGCTCGTCGGTGAGCGGCGCGTTGGCCTGCGCG
ATGACGAAGCGGTCTTCCTCGTCGGCGGTCAGGTAGTCGACCTCGTCGGTGACCTGGCCCTCGAACACCTTGCGGTAGGGGGTCTCGACGAAACCG
AACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCAGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGA
GGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGACAGACCACCGGGACCCAGAGCCGAGAGACGACGCTTGTGCGTCAGACCCGAC
AGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTG
CGGCGTGATCGCCTCGACGTCCTGGGTCGTCATCCGCTCGCGGACGACACGCTCCATACGCGCCAGACCCGTACGGACCTGGTTCTGGATGAGCTC
GCCGACGCTGCGCAGACGGCGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGAGGTGCCGCTGTCGCCGACCGTCTCGGTCTCGC
CCGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGACGTCCTCGACGGTCAGGATGCCGGCGTCGAGCGGAGCGTCCGCGCCCAGCTTCTTGT
TGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGGGTCGGGGGC
TCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGACCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCG
CCGAACTCCTCGAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTGAGCAGGACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCGGACACCGAC
CATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGATCTTGGCGGAGAAGATGTCCTTGTCGGACGTCTTGTCGATGGA
GGAGTCGAAGTAGACACCGGGCGAACGGACCAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTCGTCATGAGCG
GGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTCGTGAACTCGGCGGTGACGAAGAGCGGGGCCGCGTACGTGAAG
TCGCGCTCCTTGCACTCGTCGATCGAGTTCTTCGGCGGCTCGAAGCGGTGGTCACGGAAGGTCAGCGACATCGACCCGCTGAAGTCCTCGATCGGG
GAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTGGTGGGGACGTCCTGACCGTTGTCCAGAGCCTCCTCGACCCGACTCTGCCAAGCGGTGTTGC
CGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGCAGGTTCGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAGGAAATGCGCAGCGGGGCG
GTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP050692.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCACGCATCTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTCAGGCAGAGAGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCTC
GGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAGGC
GGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATGGA
GTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGCAGCTTCAGGATGTACATGTAGCCGACCGAGATCGGGTCCGGGAAC
GGCTCACCGCTACGGCCGTCGAACAGCCGCGCCTTGCCGGACGGGAGCACCATGCGCTCGCCGTCGCGGTTCGGGATGGTGTGCTGGAGCAGACC
CGCCAGCTCGTCCTCACGGGCGCCGTCGAAGACCGGGGTGGCGACGTTGGTGCCCGGGGCGACCTGGTCGGCGCCGATGGCCTGGAGGCGCTGC
GCCCACTCGTCGGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGCACCTGTCCCGGGTTCATTCGGGACGG
CACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTCCCGTCCTCGAGGAACGGCATGTCCTCGATGGGCAGGATCTTGGAGATGACACCCTTGT
TGCCGTGCCGGCCGGCGAGCTTGTCGCCGTCCGTGATCTTGCGCTTCTGCGCAACATACACACGCACGAGCTGGTTGACGCCGGGCGGCAGTTCGT
CGCCCTCCTCGCGGTCGAAGACGCGCACGCCGATGACCTTGCCGGTCTCGCCGTGCGGCACCTTCAGCGAGGTGTCACGGACCTCACGGGCCTTCT
CACCGAAGATCGCGCGCAGCAGGCGCTCCTCCGGCGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACGAGGATGTCACCGGCGGTGACCT
CGGCGCCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCCAGC
TTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCTC
GTAGTTGTGACCCTCCCACGGCATGAACGCGACCAGCAGGTTCTTGCCGAGGGCCATCTCGCCGTTCTCGGTGGCCGGACCGTCGGCCAGGACCTG
GCCCTCGATGATCCGGTCGCCCTCGTTGACGATGACCTTCTGGTTGACCGAGGTGCCCTGGTTGGAGCGGGCGAACTTGGCCAGGCGGTACGTGAT
GTACGTGCCGTCGTCGTTGGCGGTGGTGATGTAGTCCGCGGAGACCTCCTGGACGACACCCGCCTTCTCGGCCTTGACCACGTCGCCGGCGTCGAC
GGCGGAGCGGTACTCCATGCCGGTGCCGACGAGCGGGGACTCGCTCTTGATGAGCGGCACGGCCTGGCGCATCATGTTCGCGCCCATGAGGGCAC
GGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCCACGTCGTCACCG
GGGACGTAGTCGACCTCGCCGCCACGGCGGCGGACCAGGACGCGGGACTCCTCGAACCGGAGATCGTCGTTGAGCGTGGCGTTGGCCTGCGCGA
TGACGAAGCGGTCCTCCTCGTCGGCGGTCAGGTAGTCCACCTCGTCGGTGACCTGGCCGTCGACGACCTTGCGGTACGGGGTCTCGACGAAACCGA
ACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCGGGCGTCTCGATCGGGCACATGCGGCCGTAGTGCGA
GGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCGAC
AGCGGGTTGTTCTGGTCCATGAACTGCGAGAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTCTG

CGGCGTGATCGCCTCGACGTCCTGGGTCGTCATCCGCTCGCGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAGCTC
GCCGACGCTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCGGTCTCGACGACGATCGTCTGGCCGTTGTCCCCGACCGTCTCGGTCTCGCC
GGCGTGCAGCTTGACCAGGTACTTGATGGTCGCGATGACGTCCTCGACGGTCAGGACACCGGCGTCGAGGGGCGCGTCCGCGGCCAGCTTCTTGT
TGACCTTGTAGCGGCCGACCTTGGCGAGGTCGTAGCGCTTGGGGTTGAAGTAGAGGTTCTCGAGCAGCGTCTGCGCGGCCTCACGCGTGGGGGGC
TCGCCCGGACGCAGCTTGCGGTAGATGTCCAAGAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGACTCGTACTCG
CCGAACTCCTCGAGGATCTGCTCGGTGGTCCAGCCGAGCGCCTTCAGGAGGACGGTGACCGACTGCTTGCGCTTGCGGTCGATGCGGACACCGAC
CATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGAGGGGATGATCTTGGCCGAGAAGATGTCCTTGTCGGAGGTCTTGTCGATCGA
GGAGTCGAAGTAGACACCCGGCGACCGGACGAGCTGCGACACGACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTGGTCATGAGC
GGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCCGTCACGAACAGCGGGGCCGCGTACGTGAA
GTCGCGGTCCTTGCACTCGTCGATGCTGTTCTTCGGCGGCTCGAAACGGTGGTCGCGGAAGGTCAGCGACATCGACCCGGAGAAGTCCTCGATCG
GGGAGATCTCCTCGAAGATCTCTTCGAGGCCGGACTTGGTGGGGACGTCCTGACCGTTCTCCAGAGCCTCCTCGACCCGACTCTGCCAGGCGGTGT
TGCCGAGCAGCCAGTCGAAGCTCTCGGTCTGCAGCGCGAGCAGGTTCGGAACCTCGAGAGGCTCCTTGATCTTTGCAAAAGAGATGCGCAGCGGG
GCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCGTTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP041650.2|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCACGCATTTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCAGAGAGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTATGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAGCGGGCGTGGAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCGGCGCCTTGCCGGACGGAAGCACCATGCGCTCGCCGTCGCGGTTCGGGATCGTGTGGTTCAGCAGAC
CGGCCAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTGGTGCCCGGGGCGACCACGTCGGCCTCGATGGCCTGCAGGCGCTG
CGCCCAGTCGTCGGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGAACCTGTCCCGGGTTCATTCGGGACG
GGACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACACCCTTG
TTGCCGTGGCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGAACCAGCTGGTTCACGCCCGGCGGCAGCTCG
TCGCCCTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTCGCGGCACCTTCAGCGAGGTGTCGCGCACCTCGCGCGCCTTCT
CACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTCGCCGGACCGTCGGCCAGAACCT
GACCGGCGACGACCCGGGCGCCCTCGTCCACGACAACCTTCTGGTTGACGGAGGTGCCCTGGTTCGAGCGGGAGAACTTGTGCAGGCGGTACGTG
GTGTACGTGCCGTCGTCGTTGGCGGTGGTGACGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCCGACTTCAGTACGTCACCGGCGTCG
ACCGCGCAGCGGTACTCCATGCCGGTGCCGACCAGCGGCGCCTCGGACTTGATGAGCGGAACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCCTCAC
CGGGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGGACGCGGGACTCCTCGAAGCGGAGCTCGTCCGTCAGCGGCGCGTTGGCCTGCGC
GATGACGAATCGGTCCTCTTCGTCGGCCGTGAGGTAGTCGACCTCGTCGGTGACGACACCGCCGGTGACCTTGCGGTACGGCGTCTCGACGAAACC
GAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCAGGGGTCTCGATCGGGCACATGCGGCCGTAGTGC
GAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAG
CTCGCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCCGTCTCGACGACGATCTGCGTGCCGTTGTCACCGACGGTCTCGGTCTC
ACCGGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCGGCGTCCAGCGGCTCG-CCGC-CC-
AGCTTCTTGTTGACCTTGTAACGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCAAGAAGCGTCTGCGCGGCCTCGCGG
GTCGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGA
CTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTCAGGAGAACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCG
GACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGACCTTCGCGGAGAAGATGTCCTTGTCGGACGTCTT
GTCGATGGAGGAGTCGAAGTAGACGCCCGGCGAGCGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTG
GTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGC
GTACGTGAAGTCGCGGTCCTTGCACTCGTCGATGGAATTCTTCGGAGGCTCGAAGCGGTGGTCCCGGAACGTCAGGGACATCGACCCGGAGAAGT
CCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTCGTGGGGACGTCCTGTCCGTTGTCCAGAGCCGCCTCGACACGAGCCTTCCA
CGCGTCGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCAAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGC
GCAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP041607.2|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCACGCATTTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCAGAGAGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTATGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG

AGTACGGACCGGTCGAGCGGGCGTGGAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCGGCGCCTTGCCGGACGGAAGCACCATGCGCTCGCCGTCGCGGTTCGGGATCGTGTGGTTCAGCAGAC
CGGCCAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGGTGGCGACGTTGGTGCCCGGGGCGACCACGTCGGCCTCGATGGCCTGCAGGCGCTG
CGCCCAGTCGTCGGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGAACCTGTCCCGGGTTCATTCGGGACG
GGACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACACCCTTG
TTGCCGTGGCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGAACCAGCTGGTTCACGCCCGGCGGCAGCTCG
TCGCCCTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGCACCTCGCGCGCCTTCT
CACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTCGCCGGACCGTCGGCCAGAACCT
GACCGGCGACGACCCGGGCGCCCTCGTCCACGACAACCTTCTGGTTGACGGAGGTGCCCTGGTTCGAGCGGGAGAACTTGTGCAGGCGGTACGTG
GTGTACGTGCCGTCGTCGTTGGCGGTGGTGACGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCCGACTTCAGTACGTCACCGGCGTCG
ACCGCGCAGCGGTACTCCATGCCGGTGCCGACCAGCGGCGCCTCGGACTTGATGAGCGGAACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCCTCAC
CGGGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGGACGCGGGACTCCTCGAAGCGGAGCTCGTCCGTCAGCGGCGCGTTGGCCTGCGC
GATGACGAATCGGTCCTCTTCGTCGGCCGTGAGGTAGTCGACCTCGTCGGTGACGACACCGCCGGTGACCTTGCGGTACGGCGTCTCGACGAAACC
GAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCAGGGGTCTCGATCGGGCACATGCGGCCGTAGTGC
GAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAG
CTCGCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCCGTCTCGACGACGATCTGCGTGCCGTTGTCACCGACGGTCTCGGTCTC
ACCGGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCGGCGTCCAGCGGCTCG-CCGC-CC-
AGCTTCTTGTTGACCTTGTAACGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCAAGAAGCGTCTGCGCGGCCTCGCGG
GTCGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGA
CTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTCAGGAGAACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCG
GACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGACCTTCGCGGAGAAGATGTCCTTGTCGGACGTCTT
GTCGATGGAGGAGTCGAAGTAGACGCCCGGCGAGCGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTG
GTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGC
GTACGTGAAGTCGCGGTCCTTGCACTCGTCGATGGAATTCTTCGGAGGCTCGAAGCGGTGGTCCCGGAACGTCAGGGACATCGACCCGGAGAAGT
CCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTCGTGGGGACGTCCTGTCCGTTGTCCAGAGCCGCCTCGACACGAGCCTTCCA
CGCGTCGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCAAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGC
GCAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP041609.2|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCACGCATTTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCAGAGAGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTATGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAGCGGGCGTGGAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCGGCGCCTTGCCGGACGGAAGCACCATGCGCTCGCCGTCGCGGTTCGGGATCGTGTGGTTCAGCAGAC
CGGCCAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGGTGGCGACGTTGGTGCCCGGGGCGACCACGTCGGCCTCGATGGCCTGCAGGCGCTG
CGCCCAGTCGTCGGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGAACCTGTCCCGGGTTCATTCGGGACG
GGACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACACCCTTG
TTGCCGTGGCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGAACCAGCTGGTTCACGCCCGGCGGCAGCTCG
TCGCCCTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGCACCTCGCGCGCCTTCT
CACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTCGCCGGACCGTCGGCCAGAACCT
GACCGGCGACGACCCGGGCGCCCTCGTCCACGACAACCTTCTGGTTGACGGAGGTGCCCTGGTTCGAGCGGGAGAACTTGTGCAGGCGGTACGTG
GTGTACGTGCCGTCGTCGTTGGCGGTGGTGACGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCCGACTTCAGTACGTCACCGGCGTCG
ACCGCGCAGCGGTACTCCATGCCGGTGCCGACCAGCGGCGCCTCGGACTTGATGAGCGGAACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCCTCAC
CGGGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGGACGCGGGACTCCTCGAAGCGGAGCTCGTCCGTCAGCGGCGCGTTGGCCTGCGC
GATGACGAATCGGTCCTCTTCGTCGGCCGTGAGGTAGTCGACCTCGTCGGTGACGACACCGCCGGTGACCTTGCGGTACGGCGTCTCGACGAAACC
GAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCAGGGGTCTCGATCGGGCACATGCGGCCGTAGTGC
GAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC

TGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAG
CTCGCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCCGTCTCGACGACGATCTGCGTGCCGTTGTCACCGACGGTCTCGGTCTC
ACCGGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCGGCGTCCAGCGGCTCG-CCGC-CC-
AGCTTCTTGTTGACCTTGTAACGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCAAGAAGCGTCTGCGCGGCCTCGCGG
GTCGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGA
CTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTCAGGAGAACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCG
GACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGACCTTCGCGGAGAAGATGTCCTTGTCGGACGTCTT
GTCGATGGAGGAGTCGAAGTAGACGCCCGGCGAGCGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTG
GTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGC
GTACGTGAAGTCGCGGTCCTTGCACTCGTCGATGGAATTCTTCGGAGGCTCGAAGCGGTGGTCCCGGAACGTCAGGGACATCGACCCGGAGAAGT
CCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTCGTGGGGACGTCCTGTCCGTTGTCCAGAGCCGCCTCGACACGAGCCTTCCA
CGCGTCGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCAAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGC
GCAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP041613.2|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCACGCATTTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCAGAGAGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCGGCGCCTTGCCGGACGGAAGCACCATGCGCTCGCCGTCGCGGTTCGGGATCGTGTGCTGCAGCAGAC
CCGCGAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTGGTGCCCGGGGCGACCACGTCGGCCTCGATGGCCTGCAGTCGCTGC
GCCCAGTCCTCGGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACACCCTTGT
TGCCGTGGCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGAACCAGCTGGTTCACGCCCGGCGGCAGCTCGT
CGCCCTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGCACCTCGCGCGCCTTCT
CACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTCGCCGGGCCGTCGGCCAGAACCT
GACCGGCGACGACCCGGGCGCCCTCGTCCACGACAACCTTCTGGTTGACGGAGGTGCCCTGGTTCGAGCGGGAGAACTTGTGCAGGCGGTACGTG
GTGTACGTGCCGTCGTCGTTGGCGGTGGTGACGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCCGACTTCAGTACGTCACCGGCGTCG
ACCGCGCAGCGGTACTCCATGCCGGTGCCGACCAGCGGCGCCTCGGACTTGATGAGCGGAACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCCTCAC
CGGGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGGACGCGGGACTCCTCGAAGCGGAGCTCGTCCGTCAGCGGCGCGTTGGCCTGCGC
GATGACGAATCGGTCCTCTTCGTCGGCCGTGAGGTAGTCGACCTCGTCGGTGACGACACCGCCGGTGACCTTGCGGTACGGCGTCTCGACGAAACC
GAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCAGGGGTCTCGATCGGGCACATGCGGCCGTAGTGC
GAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGAGAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAG
CTCGCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCCGTCTCGACGACGATCTGCGTGCCGTTGTCACCGACGGTCTCGGTCTC
ACCGGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCGGCGTCCAGCGGCTCG-CCGC-CC-
AGCTTCTTGTTGACCTTGTAACGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCAAGAAGCGTCTGCGCGGCCTCGCGG
GTCGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGA
CTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTCAGGAGAACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCG
GACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGACCTTCGCGGAGAAGATGTCCTTGTCGGACGTCTT
GTCGATGGAGGAGTCGAAGTAGACGCCCGGCGAGCGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTG
GTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGC
GTACGTGAAGTCGCGGTCCTTGCACTCGTCGATGGAATTCTTCGGAGGCTCGAAGCGGTGGTCCCGGAACGTCAGGGACATCGACCCGGAGAAGT
CCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTCGTGGGGACGTCCTGTCCGTTGTCCAGAGCCGCCTCGACACGAGCCTTCCA
CGCGTCGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCAAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGC
GCAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP041602.2|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCACGCATTTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCAGAGAGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTACGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG

AGTACGGACCGGTCGAGCGGGCGTGCAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCGGCGCCTTGCCGGACGGAAGCACCATGCGCTCGCCGTCGCGGTTCGGGATCGTGTGCTGCAGCAGAC
CCGCGAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGGTGGCGACGTTGGTGCCCGGGGCGACCACGTCGGCCTCGATGGCCTGCAGTCGCTGC
GCCCAGTCCTCGGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGGACCTGTCCCGGGTTCATTCGGGACGG
GACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACACCCTTGT
TGCCGTGGCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGAACCAGCTGGTTCACGCCCGGCGGCAGCTCGT
CGCCCTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGCACCTCGCGCGCCTTCT
CACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTCGCCGGGCCGTCGGCCAGAACCT
GACCGGCGACGACCCGGGCGCCCTCGTCCACGACAACCTTCTGGTTGACGGAGGTGCCCTGGTTCGAGCGGGAGAACTTGTGCAGGCGGTACGTG
GTGTACGTGCCGTCGTCGTTGGCGGTGGTGACGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCCGACTTCAGTACGTCACCGGCGTCG
ACCGCGCAGCGGTACTCCATGCCGGTGCCGACCAGCGGCGCCTCGGACTTGATGAGCGGAACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCCTCAC
CGGGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGGACGCGGGACTCCTCGAAGCGGAGCTCGTCCGTCAGCGGCGCGTTGGCCTGCGC
GATGACGAATCGGTCCTCTTCGTCGGCCGTGAGGTAGTCGACCTCGTCGGTGACGACACCGCCGGTGACCTTGCGGTACGGCGTCTCGACGAAACC
GAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCAGGGGTCTCGATCGGGCACATGCGGCCGTAGTGC
GAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGAGAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC
TGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAG
CTCGCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCCGTCTCGACGACGATCTGCGTGCCGTTGTCACCGACGGTCTCGGTCTC
ACCGGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCGGCGTCCAGCGGCTCG-CCGC-CC-
AGCTTCTTGTTGACCTTGTAACGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCAAGAAGCGTCTGCGCGGCCTCGCGG
GTCGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGA
CTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTCAGGAGAACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCG
GACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGACCTTCGCGGAGAAGATGTCCTTGTCGGACGTCTT
GTCGATGGAGGAGTCGAAGTAGACGCCCGGCGAGCGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTG
GTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGC
GTACGTGAAGTCGCGGTCCTTGCACTCGTCGATGGAATTCTTCGGAGGCTCGAAGCGGTGGTCCCGGAACGTCAGGGACATCGACCCGGAGAAGT
CCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTCGTGGGGACGTCCTGTCCGTTGTCCAGAGCCGCCTCGACACGAGCCTTCCA
CGCGTCGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCAAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGC
GCAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

>gb|CP041651.1|

TCAGACCTCTTCGACGCTGCTCGGCTCGCGCCGGGACAGGTCGATGCCGAGCTCCTCCGCCGCGCGGAAGACGTCCTCGTCGGTGTCACGCATTTC
GATGGACATACCGTCGCTGGACAGCACCTCCACGTTGAGGCAGAGAGACTGCATCTCCTTGATGAGCACCTTGAAGGACTCGGGGATGCCGGGCT
CGGGGATGTTCTCGCCCTTGACGATGGCCTCGTAGACCTTCACGCGGCCGGTCACGTCGTCGGACTTGATGGTCAGCAGCTCCTGGAGGGCGTAAG
CGGCGCCGTATGCCTCCAGCGCCCACACCTCCATCTCACCGAAGCGCTGGCCACCGAACTGGGCCTTACCACCCAGCGGCTGCTGGGTGATCATCG
AGTACGGACCGGTCGAGCGGGCGTGGAGCTTGTCGTCGACCAGGTGGTGGAGCTTGAGGATGTACATGTAGCCGACCGAGATCGGCTCCGGGAA
CGGCTCGCCGGAGCGGCCGTCGAACAGCGGCGCCTTGCCGGACGGAAGCACCATGCGCTCGCCGTCGCGGTTCGGGATCGTGTGGTTCAGCAGAC
CGGCCAGCTCGTCCTCACGCGCACCGTCGAAGACGGGGGTGGCGACGTTGGTGCCCGGGGCGACCACGTCGGCCTCGATGGCCTGCAGGCGCTG
CGCCCAGTCGTCGGCGAGACCGGAGACGTCCCAGCCGCGGCTGGCGAGCCAGCCGAGGTGGATCTCCAGAACCTGTCCCGGGTTCATTCGGGACG
GGACACCGAGCGGGTTGAGGATGATGTCGACCGGGGTGCCGTCCTCCAGGAACGGCATGTCCTCGATCGGCAGGATCTTCGAGATGACACCCTTG
TTGCCGTGGCGGCCGGCGAGCTTGTCACCGTCCGTGATCTTGCGCTTCTGCGCCACGTAGACACGAACCAGCTGGTTCACGCCCGGCGGCAGCTCG
TCGCCCTCTTCACGGTCGAAGACGCGGACGCCGATGACCTTGCCGATCTCGCCGTGCGGCACCTTCAGCGAGGTGTCGCGCACCTCGCGCGCCTTCT
CACCGAAGATCGCGCGGAGCAGGCGCTCCTCCGGGGTCAGCTCGGTCTCACCCTTGGGCGTGACCTTGCCGACCAGGATGTCGCCGGCGACGACC
TCGGCACCGATGCGGATGATGCCGCGCTCGTCGAGGTCGGCGAGGACCTCCTCGGAGACGTTCGGGATGTCCCGGGTGATCTCCTCGGGGCCGAG
CTTGGTGTCACGGGCGTCGACCTCGTGCTCCTCGATGTGGATCGAGGAGAGGACGTCGTCCTGCACGAGGCGCTGCGACAGGATGATCGCGTCCT
CGTAGTTGTGACCCTCCCACGGCATGAACGCCACGAGCAGGTTCTTGCCGAGCGCCATCTCGCCGTTCTCGGTCGCCGGACCGTCGGCCAGAACCT
GACCGGCGACGACCCGGGCGCCCTCGTCCACGACAACCTTCTGGTTGACGGAGGTGCCCTGGTTCGAGCGGGAGAACTTGTGCAGGCGGTACGTG
GTGTACGTGCCGTCGTCGTTGGCGGTGGTGACGTAGTCCGCGGAGACCTCCTGGACCACACCGTCCTTCTCCGACTTCAGTACGTCACCGGCGTCG
ACCGCGCAGCGGTACTCCATGCCGGTGCCGACCAGCGGCGCCTCGGACTTGATGAGCGGAACGGCCTGACGCATCATGTTCGCGCCCATGAGGGC
ACGGTTGGCGTCGTCGTGCTCGAGGAACGGGATCATGGCGGTCGCGACCGACACCATCTGGCGCGGCGAGACGTCCATGTAGTCGACGTCCTCAC
CGGGGACGTAGTCGACCTCTCCGCCACGACGGCGGACCAGGACGCGGGACTCCTCGAAGCGGAGCTCGTCCGTCAGCGGCGCGTTGGCCTGCGC
GATGACGAATCGGTCCTCTTCGTCGGCCGTGAGGTAGTCGACCTCGTCGGTGACGACACCGCCGGTGACCTTGCGGTACGGCGTCTCGACGAAACC
GAACGCGTTGACCCGGCCGTAGGAGGCGAGCGAGCCGATCAGACCGATGTTCGGGCCTTCAGGGGTCTCGATCGGGCACATGCGGCCGTAGTGC
GAGGGGTGCACGTCACGGACCTCGAAGCCGGCCCGCTCACGGGAGAGACCACCCGGGCCAAGAGCCGACAGACGGCGCTTGTGGGTGAGACCCG
ACAGCGGGTTGTTCTGGTCCATGAACTGCGACAGCTGGCTGGTGCCGAAGAACTCCTTGATGGAGGCGACGACCGGCCGGATGTTGATCAGGGTC

TGCGGCGTGATCGCCTCGACGTCCTGGGTCGTCATGCGCTCGCGGACGACGCGCTCCATACGAGCCAGACCCGTGCGGACCTGGTTCTGGATGAG
CTCGCCGACGTTGCGCAGACGACGGTTGCCGAAGTGGTCGATGTCGTCCGTCTCGACGACGATCTGCGTGCCGTTGTCACCGACGGTCTCGGTCTC
ACCGGCGTGCAGCTTCACCAGGTACTTGATCGTCGAGATGATGTCCTCGACGGTCAGGATGCCGGCGTCCAGCGGCTCG-CCGC-CC-
AGCTTCTTGTTGACCTTGTAACGGCCGACCTTCGCGAGGTCGTAGCGCTTCGGGTTGAAGTAGAGGTTCTCAAGAAGCGTCTGCGCGGCCTCGCGG
GTCGGGGGCTCGCCCGGACGCAGCTTGCGGTAGATGTCGAGCAGCGCGTCGTCCTGGCCCTGGGTGTGGTCCTTCTCCAGGGTGGCGCGCATGGA
CTCGTACTCGCCGAACTCCTCGAGGATCTGCTCGGTCGTCCAACCGAGAGCCTTCAGGAGAACGGTGACGGACTGCTTGCGCTTGCGGTCGATGCG
GACACCGACCATGTCGCGCTTGTCGATCTCCATCTCCAGCCAGGCACCCCGGGACGGGATGACCTTCGCGGAGAAGATGTCCTTGTCGGACGTCTT
GTCGATGGAGGAGTCGAAGTAGACGCCCGGCGAGCGGACGAGCTGCGACACCACGACACGCTCGGTGCCGTTGATGACGAAGGTGCCCTTGTTG
GTCATGAGCGGGAAGTCGCCCATGAAGACCGTCTGGGACTTGATCTCGCCGGTCTCGTTGTTGGTGAACTCGGCGGTGACGAAGAGCGGGGCCGC
GTACGTGAAGTCGCGGTCCTTGCACTCGTCGATGGAATTCTTCGGAGGCTCGAAGCGGTGGTCCCGGAACGTCAGGGACATCGACCCGGAGAAGT
CCTCGATCGGGGAGATCTCCTCGAAGATCTCCTCCAGACCGGACTTCGTGGGGACGTCCTGTCCGTTGTCCAGAGCCGCCTCGACACGAGCCTTCCA
CGCGTCGTTGCCGAGCAGCCAGTCGAAGCTCTCGGTTTGCAGCGCAAGAAGGTTCGGAACCTCGAGGGGCTCCTTGATCTTTGCAAAGGAGATGC
GCAGCGGGGCGGTGCTGGCGCCGTTGTTCGTATTCGCGGTCGAGGCAGTGCGCGAGGCGGCCAAGAGGGGGTCCTTCCGAGGGCTCGGACTCAC

A.2.3.3.1: NCBI reference *Streptomyces* and isolates *rpoB* sequences compiled

The *rpoB* sequences were used to create the phylogenetic tree comprising the *Streptomyces* reference sequences from NCBI (n= 211) and isolates sequences (n = 133). *M. tuberculosis* was used as an outgroup.

>NZ_CP023699.1_*S._kanamyceticus*

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCGTACGGCCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGACCTC
CACTACGCCGAGCCGCGCGTCCTGATCCGCCGTCGTGGCGGAGAGATCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATCACCGCCGAGGCCCCCCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACGTGATCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP045704.1_*S._rutgersensis*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGCCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTGGAGACGCCGTACCGCCGGGTCACCG---
AAGGTGTCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGGCTGACGGACGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTGTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGCGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCG
CCACCGACGCCGGTGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCAC
GTACACCACGTACCGCCTCGCC

>NZ_AP018365.1_S._sp._SN-593

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGGATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCCTACGGCCGGGTGAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCACCG---
AGGGCCAGGTCACCGACCACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCACCGACGAGCTG
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGCGGCGGCGAGATCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGCGCGCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGC
GCGGTCGACGCCGGCGACGTGATCAAGGCCGAGAAGGACGGCGTGGTCCAGGAGGTCTCCGCCGACTACGTCACGGTCACCAACGACGACGGCA
CCTACACCACCTACCGGGTCGCC

>NZ_CP048882.1_*S._bathyalis*

GCGGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTCGCCTCGTACGGGCGGGTGAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTGGTCG---
ACGGTGTCGTCGGCGATGAGGTGCACTACCTGACGGCCGATGAAGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGCCGCTCACCGAAGAGATG
CGCTTCGCCGAGGACCGGGTGCTCATCCGCCGCCGTGGCGGCGAGGTCGACCTGATC—
CCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACGGCCATGATCCCCTTCCTCGAGCACGACGACGCCAA
CCGCGCGCTCATGGGGTCGAACATGATGCGGCAGGCCGTGCCGCTCATCAAGGCCGAGTCGCCGCTGGTCGGCACCGGTATGGAGTACCGCTGCG
CGGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACGTCACGGTGGCCAACGACGACGGCAC
CTACACCACGTACCGCGTCGCG

>NZ_CP070326.1_*S._noursei*

GCCGGTCTGGACGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCAGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGCCGCGTCAACGTCTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGAGGAGGTGGACTACCTCACCGCAGATGAGGAGGACCGCTTCCTCATCGCCCAGGCCAACGCCAAGCTCACCGACGACCTG

CGCTTCGCCGAGCAGCGTGTGCTGGTCCGTCGCCGTGGCGGCGAGGTCGACCTCGTC—
CCCGCCGACGAGGTGCACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGCCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGACGCCAAC
CGCGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCGGAGTCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGAACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCGTCGCC

>NZ_CP071139.1_*S._nojiriensis*

GCCGGCTTCGAGGTCCGTGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGTCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGTCCGAGGACATG
CGCTTCACCGAGAACCGCGTGCTGGTGCGCCGTCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACTCGATCCGTGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACATCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCC

>NZ_CP047147.1_S._sp._GF20

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTCGTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>NZ_CP016825.1_*S._albus*

GCCGGCTTCGAGGTCCGCGACGTGCACCCCTCGCACTACGGCCGGATGTGTCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTGGAGACCCCGTACCGCCGCGTCACCG---
AGGGTGTCGTCACCGACGAGGTCGACTACCTCACCGCCGACGAAGAGGACCGCTTCGTCATCGCCCAGGCGAACGCCCCGCTCACCGAGGACCTCC
ACTTCGAGGAGTCCCGGGTGCTGGTGCGCCGCCGCGGCGGCGAGATCGACTACGTC—
CCCGGCGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGACGCCAAC
CGCGCGCTCATGGGCGCCAACATGATGCGCCAGGCCGTGCCGCTGATCACCGCCGAGGCCCCCCTCGTCGGTACGGGCATGGAGTACCGCTGCGC
CGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCGTCGCCAACGACGACGGCACCT
ACACCACGTACCGCATCGCC

>NZ_CP016824.1_*S._sampsonii*

GCCGGGTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTCGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>NZ_CP015098.1_*S._qaidamensis*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTGGAGACCCCGTACCGCAAGGTGTTCG---
AGGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCAGCTGACGGACGACCTC
CGCTTCGCCGAGGCCCGGGTGCTGGTCCGCCGTAAGGGCGGCGAGGTCGACTACGTC—
ACCGGCGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGTCCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGCG
GTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGCCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACGT
ACATCACGTACCGCCTGGCC

>NC_021985.1_*S._collinus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCAACG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGAACTG
CGGTTCGCCGAGGCGCGTGTCCTGGTCCGCCGCCGCGGCGGTGAGGTCGACTACGTC—
AGCCCCGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTCATCAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
GGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGGCCAACGACGACGGCACG
TACATCACGTACCGCCTGCAC

>NC_020990.1_*S._albidoflavus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>NZ_CP023701.1_*S._subrutilus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGTCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGTGTCGTCACGGACGACGTCGACTACCTGACCGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGAACGACGAGATG
CGCTTCACCGAGAACCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACTCGATCCGCGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCGGCCGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCGTCGCC

>NZ_CP011664.1_S._sp._Mg1

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGTCGCGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCATCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGTCCGAGGACATG
CGCTTCACCGAGAACCGCGTGCTGGTCCGCCGTCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACTCGATCCGTGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACGTCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCG

>NZ_CP033073.1_S._sp._Z022

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCCGCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCGGTGACGACATG
CGGTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GGCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCCGTGCCGCTGATCAAGTCCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGCC
GTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACGT
ACATCACGTACCGCCTGGCC

>NZ_CP023692.1_*S._vinaceus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGTCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGAACGACGAGATG
CGCTTCACCGAGAACCGCGTACTGGTCCGCCGTCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC

GGTCGACGCCGGCGACTCGATCAAGGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCGTCGCC

>NZ_CP010849.1_*S._cyaneogriseus*

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCACCGACGAGATG
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCAAGGGCGGCGAGGTCGACTACGTC—
GGCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATTAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
GGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP048641.1_*S._aureoverticillatus*

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCGTCCTACGGCCGTGTGAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAGGGTCGCCG---
ACGACCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGCGCTGACCGACGACATG
CGCTTCGCCGAGGCGCGTGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
CCGGGCTCCGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACGGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATCACCGCCGAGGCCCCCCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTGATCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCGTCACCAACGACGACGGCACG
TACACCACGTACCGAGTCGCC

>NZ_CP009802.1_S._sp._FR-008

GCCGGGTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>NZ_CP020555.1_S._sp._Sge12

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGTCCGAGGACATG
CGCTTCACCGAGAACCGCGTGCTGGTGCGCCGTCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACTCGATCCGTGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACATCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCC

>NZ_CP023702.1_*S._nitrosporeus*

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACATCACCGCCGACGAGGAAGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGAGGACATG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGTCGTGGCGGGGAGGTCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTCCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>NZ_CP029189.1_*S._venezuelae*

GCCGGCTTCGAGGTCCGAGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---

ACGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGTCCGAGGACATG
CGCTTCACCGAGAACCGCGTGCTGGTGCGCCGTCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACTCGATCCGTGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACATCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCC

>NZ_CP033581.1_S._sp._ADI95-16

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGTCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGTCCGAGGACATG
CGCTTCACCGAGAACCGCGTGCTGGTCCGCCGTCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACTCGATCCGCGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACGTCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCG

>NZ_CP016795.1_*S._olivaceus*

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGATCAACCCGTTCGGCTTCATCGAGACGCCCTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACCTG
CGGTTCGCCGAGGCCCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTCCG
CCGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGCCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>NZ_CP023693.1_*S._cinereoruber*

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACGTCACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGCGCTCGACGACGAGCTG
CGCTTCTCCGAGAACCGCGTCCTGGTCCGCAAGCGCGGCGGCGAGGTCGACTACGTCGAGCCGTCG—
GACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAACCGTGCCCTC
ATGGGCGCGAACATGATGCGTCAGGCGGTGCCCCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGCCACCGACGC
CGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGCTGTCGGCCGACTACATCACGGTCGCCAACGACGACGGCACGTACATCACGT
ACCGCCTGCAC

>NZ_CP034350.1_S._sp._W1SF4

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGTCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGCCGCGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCGCCGTCACCGACGAGGTCGACTACGTCACCGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCCCGCCTGGACGACGAGCTG
CGCTTCACCGAGAACCGCGTCCTGGTCCGCAAGCGCGGCGGCGAGGTCGACTACGTCGAGCCGTCG—
GACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTCGAGCACGACGACGCCAACCGCGCCCTC
ATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGCGGTCGACGC
CGGCGACTCGATCAAGGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACATCACCGTCGCCAACGACGACGGCACGTACACCACGT
ACCGCGTCGCC

>NZ_CP041610.2_S._sp._RLB3-17

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>NZ_CP019798.1_S._sp._fd1-xmd

GCCGGCTTCGAGGTCCGAGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ATGGTGTCGTCACCGACGAGGTCGACTACCTGACGGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCGGCCCTGTCCGAGGACATG
CGCTTCACGGAGAACCGCGTCCTGGTCCGCCGCCGCGGCGGCGAGATCGACTACATC—
GCCGGCGGCGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGCGACTCGATCCGCGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACATCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCG

>NZ_CP041612.2_S._sp._S1A1-8

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>NZ_CP034353.1_S._sp._KPB2

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGCATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGGTTCGCCGAGGCCCGCGTCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCCACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTCCGC
CGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGCCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGACGACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP029338.1_S._sp._SM17

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACCCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>NZ_CP029078.1_*S._griseoviridis*

GCCGGCTTCGAGGTCCGCGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGCCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGGTTCGAGGAGTCCCGGGTCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCCCTCATCAAGTCCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGCC
GTCGACGCCGGTGACGTGGTCAGGGCCGAGAAGGCGGGCGCCGTCCAGGAGGTCTCCGCCGACTACATCACCACCGCCAACGACGACGGCACGT
ACATCACGTACCGGCTGGCC

>NZ_CP011340.1_*S._pristinaespiralis*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCACGG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCCACCCTGAGCGACGAGCTG
CGCTTCACCGAGGCACGCGTGCTGGTCCGCCGTCGTGGCGGCGAGGTCGACTACGTC—
CCGCCGACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACGGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC

CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACGTCACCGTCGCGAACGACGACGGCACG
TACACCACGTACCGCATCGCG

>NZ_CP026304.1_*S._lunaelactis*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGGTGAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCCGGGTCGTCG---
ACGACCAGGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGGTTCGTCATCGCGCAGGCCAACGCCGCGCTGACCGACGACATG
CGGTTCGCCGAGAACCGCGTGCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCGGGCTCCGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATCACGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTCATCAAGGCGGAGAAGGAAGGTGTTGTCCAGGAGGTATCCGCGGACTACATCACCGTCGCCAATGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP023703.1_*S._galilaeus*

GCCGGCTTCGAGGTCCGCGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTGATCGGTTC
GCTCGCCTCGTACGGCCGGATCAACCCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGGTTCGTCATCGCGCAGGCCAACGCGCCGCTCAGTGACGAGTTC
CAGTTCGAGGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGCGACGTGGTCAGGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP032427.1_*S._griseorubiginosus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTGGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGGTTCGAGGAGGCCCGGGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACAGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP017248.1_*S._fodineus*

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGTATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCAGCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGAACGACGAGAT
GCGGTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GGCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATTAAGTCCGAGGCACCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
GGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP023407.1_S._*fungicidicus*

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAAACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCAACGACGAGATG
CGCTTCGTCGAGAACCGCATCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
GGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP023697.1_*S._prasinus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGAGTGACGACATG
CGGTTCGCCGAGGCGCGCGTGCTGGTCCGCCGCAAGGGCGGCGAGGTCGACTACGTC—
AGCCCCGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATCAGGTCCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGCC
GCCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCACGT
ACATCACGTACCGCCTGGCC

>NZ_CP029159.1_*S._tsukubensis*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCCATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCCACGCTCAACGACGAGCTG
CGGTTCGCCGAGGCCCGGGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
CCGCCGGCCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCACGGCCGAAGCGCCGCTGGTCGGTACCGGCATGGAGTACCGCTGCGC
CGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGGCAACGACGACGGCACCT
ACACCACCTACCGGGTGGCC

>NZ_CP015866.1_*S._parvulus*

GCCGGCTTCGAGGTCCGCGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGATCAACCCCTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGCTTCTCCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTCATCAAGAGCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGCCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP022685.1_*S._formicae*

GCAGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCCATTGAGACTCCCGAAGGCCCCAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGACCTC
CACTTCGCCGAGGCCCGCGTCCTGATCCGCCGTCGTGGCGGAGAGATCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATCACCGCCGAGGCCCCCCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTGATCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP047318.1_S._sp._HM190

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTGGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGCTTCGCCGAGAACCGCGTGCTCGTCCGCCGTCGTGGCGGCGAGGTCGACTACGTC—
GCCGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>NZ_CP031194.1_*S._paludis*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGTCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
ACGGTGTCGTCACCGACGAGGTCGACTACCTGACGGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCTCCGAGGACATG
CGCTTCACGGAGCCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
CCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCCCTGATCACGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCG

CGGTCGACGCCGGTGACGTCATCAAGGCGGAGAAGGAGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGTAC
GTACACCACGTACCGCATCGCC

>NZ_CP041650.2_S._sp._RLB1-8

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>NZ_CP019724.1_*S._pactum*

GCCGGCTTCGAGGTCCGTGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGACCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCCAGGTCACGGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGCTTCTCCGAGGCCCGCGTCCTGGTTCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTCCGC
CGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP022438.1_*S._peucetius*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCACCG---
ACGGTCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCGAACGCCACGCTGAACGACGAGCTC
CGCTTCACCGAGGCCCGCGTGCTGGTCCGCCGTCGTGGCGGCGAGGTCGACTACGTC—
CCGCCGGCCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGTGACGTGATCAAGGCCGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACGTCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>NZ_CP027306.1_*S._atratus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACATCACGGCCGACGAGGAAGACCGCTACGTCATCGCTCAGGCGAACGCGACCCTCTCCGAGGAACTG
CGTTTCGTCGAGCCCCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTGAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP028834.1_S._sp._M2

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTGACGGACGACCTG
CGCTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GCCGGTGACGACGTGGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTGATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCTG
CGGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>NZ_CP023695.1_*S._alboniger*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATTGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---

AGGGCCAGGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGACCTC
CACTTCGCCGAGGCCCGCGTCCTGATCCGCCGCCGTGGCGGAGAGATCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATCACCGCCGAGGCTCCCCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTGATCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCGTCACCAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP026490.1_S._sp._604F

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGCCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCCGGGTCACCG---
ACGGCATCGTCACGGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTTTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>NZ_CP027297.1_S._sp._SGAir0924

GCCGGCTTCGAGGTCCGTGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGATCAACCCGTTCGGCTTCATCGAGACGCCTTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGGTTCGCCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTCCGC
CGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP030862.1_*S._globosus*

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGCCCAGGTGACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCGCGCCTGGACGAGAACCTG
CGCTTCGCGGAGGGCCGCGTCCTCGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CGTCGACGCCGGCGACTCCATCAAGGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACATCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGGGTCGCC

>NZ_CP029043.1_*S._nigra*

GCCGGCTTCGAGGTCCGTGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGACCTC
CGCTTCGAGGAGGCCCGCGTCCTGGTCCGCCGGCGTGGCGGCGAGGTCGACTACGTC—
GGCGGCGAGGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CGGTCGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCAGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>NZ_CP010407.1_*S._vietnamensis*

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACGTCACCGCCGACGAGGAGGACCGGTTCGTCATCGCGCAGGCCAACGCCGCGCTCGACGACGAGCTC
CGCTTCTCCGAGAACCGCGTCCTGGTCCGCAAGCGTGGCGGCGAGGTCGACTACGTCGAGCCGTCG—
GACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTC
ATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGCCACCGACGC
CGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGCTGTCGGCCGACTACATCACGGTCGCCAACGACGACGGCACGTACATCACGT
ACCGCCTGCAC

>NZ_CP020039.1_S._sp._3211

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGTCCGAGGACATG
CGCTTCACCGAGAACCGCGTGCTGGTGCGCCGTCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACTCGATCCGTGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACATCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCC

>NZ_CP029617.1_S._sp._WAC_01529

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCCATTGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGTCAGGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGACCTC
CACTTCGCCGAGGCCCGCGTCCTGATCCGCCGCCGTGGCGGAGAGATCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATCACCGCCGAGGCTCCCCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTGATCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCGTCACCAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP022744.1_*S._lincolnensis*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTGGAGACGCCGTACCGCAAGGTCACCG---
CTGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGGTTCGCCGAGAACCGCGTGCTCGTCCGTCGTCGTGGTGGCGAGGTCGACTACGTC—
GCCGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
GGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGTACG
TACATCACGTACCGCCTGGCC

>NZ_CP012949.1_*S._ambofaciens*

GCCGGCTTCGAGGTCCGCGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGATCAACCCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGGTTCTCCGAGGCCCGTGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTCCGC
CGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP061072.1_S._sp._CB00271

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACGGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACGCTCTCCGACGAGCTG
CGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTCATCAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTCCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTATCCGCGGACTACATCACCGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP018870.1_S._sp._TN58

GCCGGCTTCGAGGTCCGTGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCGGCCCTGTCCGAGGACATG
CGCTTCACGGAGAACCGCGTCCTGGTCCGCCGCCGCGGCGGCGAGATCGACTACATC—
GCCGGCGGCGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTCCTGGAGCACGACGACGCCAAC

CGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTATCGCTGTGC
GGTCGACGCCGGTGACTCGATCCGCGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACATCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCC

>NZ_CP027858.1_*S._clavuligerus*

GCGGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGCTTCGTGGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGCGCTCAACGACGAGCTG
CAGTTCACCGAGGCCCGGGTGCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CTGCCGGCCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTGATCACGGCCGAGGCGCCGCTGGTCGGTACCGGTATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTGATCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACC
TTCACCACCTACCGGGTGGCC

>NZ_CP054926.1_*S._fulvissimus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGGCGCATGTGTCCGATCGAGACCCCTGAAGGCCCGAACATCGGCCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGTCAGGTCACCGACGACGTCGACTACATCACGGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGATCCTCTCCGACGAGCTG
CGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGTCCGAGGCCCCGCTCGTGGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP023688.1_*S._rimosus*

GCCGGCCTGGACGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGGGTCAACGTCTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCGGGTCACCGAGCAGGTCGACTACCTGACCGCCGACGAGGAAGACCGCTTCCTGATCGCCCAGGCGAACGCGCCCCTCACCGAGGACATG
CAGTTCGCCGAGGGCCGCGTGATGTGCCGCCGCCGCGGTGGTGAGGTCGACCTCGCC—
GCCCGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGCCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGCGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGGCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCG
CGGTCGACGCCGGTGACGTCATCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCAC
GTACACCACGTACCGCGTCGCC

>NZ_CP066774.1_*S._costaricanus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTC
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GGTGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCAAGTCCGAGGCGCCGCTCGTCGGTACCGGCATGGAGTACCGCTCCGC
CGTGGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGGCGAACGACGACGGCACG
TACATCACGTACCGCCTGCAC

>NZ_CP054918.1_S._sp._NA04227

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTGGAGACCCCGTACCGCAGGGTCACCG---
ACGGTGTCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCACCGAGGACCTG
CACTTCGAGGAGTCCCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
CCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACGTCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCC

>NZ_CP031742.1_*S._koyangensis*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>NZ_CP049838.1_*S._cacaoi*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTGGAGACGCCGTACCGCCGGGTCACCG---
ACGGTGTCGTCACCGACGAGGTCGACTACCTGACGGCCGACGAAGAGGACCGGTTCGTCATCGCGCAGGCCAACGCCACGCTCGACGACGACCTG
CGCTTCACCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGTGAGGTCGACTACGTC—
CCCGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP066831.1_S._sp._ZYC-3

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGTCCGATCGAGACCCCCGAAGGCCCGAACATCGGCCTGATCGGCTCG
CTCGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGTCAGGTCACCGACGACGTCGACTACCTCACCGCCGACGAGGAGGACCGGTTCGTCATCGCGCAGGCCAACGCCTCGCTCAATGACGACATG
CGCTTCGACGAGGCTCGTGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATCACCGCCGAGGCCCCCCTCGTCGGTACGGGCATGGAGTACCGCTGCGC
CGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGCGGGTGTCGTCCAGGAGGTGTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCC

>NZ_CP051486.1_*S._pratensis*

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGGCGCATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTCACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTGTCCGAGGACATG
CGCTTCACCGAGGCCCGCGTCCTCGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGTACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP060828.1_S._sp._CRXT-G-22

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGAACTG
CGCTTCACCGAGAACCGCGTCCTGGTCCGCCGTCGTGGTGGCGAGGTCGACTACGTC—
CCCGGTGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGTGACGTCGTCAAGGCCGAGAAGCCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCGAACGACGACGGCACGT
ACATCACGTACCGCCTGCAC

>NZ_CP043638.1_S._sp._INR7

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGTGTCGTCACCGACGAGGTCGACTACCTGACGGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCGGCCCTGTCCGAGGACATG
CGCTTCACGGAGAACCGCGTCCTGGTCCGCCGTCGTGGCGGCGAGATCGACTACATC—
GCCGGCGGCGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC

GGTCGACGCCGGTGACTCGATCCGTGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACATCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCC

>NZ_CP056773.1_*S._violascens*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACCCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>NZ_CP029188.1_*S._tirandamycinicus*

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGCCTGATCGGTTCG
CTCGCCACCTACGGCCGGGTGAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGAGG---
GCGGCCAGGTCACCGACCAGGTGGACTACCTCACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGCGCTCGGCGACGACATG
CGCTTCTCCGAGGCCCGCGTGCTCGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTG—
TCGCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTGGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGCGCCCTGATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCGGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCG
CCACCGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTGATCCAGGAGGTCTCGGCCGACTACGTCACCGTCGCCAACGACGACGGCAC
GTACACCACGTACCGCGTCGCG

>NZ_CP070242.1_*S._californicus*

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACCCTCAACGACGAGCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCCGAGGCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCG
CCACCGACGCCGGCGACGTCCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCAC
GTACACCACGTACCGCATCGCC

>NZ_CP051006.1_*S._griseofuscus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTG
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GGTGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCAAGTCCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTGGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACGGCGAACGACGACGGCACG
TACATCACGTACCGCCTGCAC

>NZ_CP061281.1_S._sp._CRXT-Y-14

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCAACGACGAGCTC
CGCTTCGTCGAGAACCGCGTGCTGGTCCGCCGTCGTGGTGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCATCACCAACGACGACGGCACGT
ACACCACGTACCGCCTGCAC

>NZ_CP029241.1_S._sp._NHF165

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGGATGTGCCCGATCGAGACGCCGGAAGGGCCGAACATCGGTCTGATCGGTTC
GCTCGCGTCCTACGGCCGGCTCAACCCCTTCGGTTTCGTCGAGACGCCCTACCGCAAGGTGCAGG---

ACGGCGTCGTCACCGACGAGGTGCACTACCTCACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCGAACGCGCCGCTCACCGAGGACATG
CGCTTCGCCGAGGACCGGGTGCTCATCCGCCGCCGTGGCGGCGAGGTCGACCTGATC—
CCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGCGCGCTCATGGGGTCGAACATGATGCGGCAGGCCGTGCCGCTCATCAAGGCCGAGGCGCCGCTGGTCGGCACCGGTATGGAGTACCGCTGCG
CGGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTCACGGTGGCCAACGACGACGGCAC
GTACACCACCTACCGCGTCGCC

>AP018036.1_*M._tuberculosis*

GCCGGGCTGGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGGATGTGCCCGATCGAAACCCCTGAGGGGCCCAACATCGGTCTGATCGGCTC
GCTGTCGGTGTACGCGCGGGTCAACCCGTTCGGGTTCATCGAAACGCCGTACCGCAAGGTGGTCG---
ACGGCGTGGTTAGCGACGAGATCGTGTACCTGACCGCCGACGAGGAGGACCGCCACGTGGTGGCACAGGCCAATTCGCCGATCGATGCGGACGG
TCGCTTCGTCGAGCCGCGCGTGCTGGTCCGCCGCAAGGCGGGCGAGGTGGAGTACGTG—
CCCTCGTCTGAGGTGGACTACATGGACGTCTCGCCCCGCCAGATGGTGTCGGTGGCCACCGCGATGATTCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGGGCAAACATGCAGCGCCAGGCGGTGCCGCTGGTCCGTAGCGAGGCCCCGCTGGTGGGCACCGGGATGGAGCTGCGCGCGG
CGATCGACGCCGGCGACGTCGTCGTCGCCGAAGAAAGCGGCGTCATCGAGGAGGTGTCGGCCGACTACATCACTGTGATGCACGACAACGGCACC
CGGCGTACCTACCGGATGCGC

>NZ_CP021121.1_S._sp._SCSIO_3032

GCCGGCCTCGAAGTCCGCGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTCGCCGCCTTCGGCCGGGTGAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCACCG---
ACGGCGTCGTCACCGACGAGGTGCACTACCTGACCGCCGACGAGGAGGACAGGTACGTCATCGCGCAGGCCAACGCGCGGCTGACGGACGACTA
CCACTTCGCCGAGAACCGCGTCCTCGTCCGCCGCCGCGGCGGCGAGATCGACTACGTG—
ACCGGCTCCGAGGTTCACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTTGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTCATCCAGGCCGAGTCGCCGCTGGTCGGCACCGGCATGGAGTACCGGTCCGC
GGTCGACGCGGGCGACGTGATCAAGGCCGCGCAGGACGGCGTGATCACCGACGTCTCCGCCGACTACATCACCATCGCCAACGACGACGGCACGA
ACGTCACCCACCGGCTGGCG

>NZ_CP009754.1_S._sp._CCM_MD2014

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGGTTCGCCGAGGCCCGCGTCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCGGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTCCGC
CGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGCCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP021080.1_*S._pluripotens*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCCTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGATGACCTG
CGTTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
GGCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTTGAGCACGACGACGCCAA
CCGCGCCCTCATGGGTGCCAACATGATGCGTCAGGCCGTGCCGCTGATTAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
GGTCGACGCCGGCGATGTCGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP045643.1_*S._fagopyri*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACGG---
GCGGCATCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGGTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTC
CGCTTCGAGGAGTCCCGCGTTCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGTCCGAGGCACCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CGTCGACGCCGGCGACGTGCTCAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>NZ_CP017316.1_*S._rubrolavendulae*

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGTGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCATCG---
ACGGCCGGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACATG
CGCTTCGCCGAGAACCGCGTCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GCCGGGGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAA
CCGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATCCGCTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGCTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP029378.1_*S._bacillaris*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGGCGCATGTGTCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACCCTCTCCGACGAGCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGCCGCGGGGGAGAGGTCGACTACGTC—
CCCGCCTCCGAGGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NC_021055.1_S._sp._PAMC_26508

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGTACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP009438.1_*S._glaucescens*

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCCTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCGGCGACGACATG
CGCTTCTCCGAGGCTCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
AGCCCCGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCCAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGCC
GTCGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACGT
ACATCACGTACCGCCTGGCC

>NZ_CP029618.1_S._sp._WAC_06738

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGTCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTCATCGGCTC
GCTCGCCTCGTACGGCCGGGTGAACGCCTTCGGCTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
GCGGCGTCGTCACCGACGACGTGGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCACGAACGACCTG
CACTTCGCCGAGAACCGGGTCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCCAACATGATGCGCCAGGCCGTGCCGCTGATCCGGGCCGAGGCCCCCCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGCGACGTCATCAAGGCCGACAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCGTCAAC

>NC_015953.1_S._sp._SirexAA-E

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAGGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGTCAGGTCACCGACGAGGTCGACTACATCACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACGCTCTCCGACGAGCTG
CGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC

CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGTCGGAGGCCCCGCTCGTGGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>NZ_CP029541.1_S._sp._NEAU-S7GS2

GCCGGCCTGGACGTCCGTGACGTGCACCCCTCGCACTACGGCCGTATGTGCCCGATTGAGACCCCTGAAGGTCCCAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGGTCAACGTCTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGAGGAGGTGGACTACCTCACCGCTGATGAGGAGGACCGCTTCCTGATCGCCCAGGCGAACGCGAAGCTCTCCGACGACATG
CGCTTCTCCGAGCAGCGTGTCCTGGTCCGCCGTCGTGGCGGCGAGGTCGACCTCGTC—
CCCGCCGACGAGGTGGACTTCATGGACGTCTCGCCGCGCCAGATGGTGTCGGCCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCGGAGTCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCGTCGCC

>NZ_CP009922.3_*S._xiamenensis*

GCCGGCCTGGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCCAACATCGGTCTGATCGGCTCG
CTGGCCGCGTTCGGCCGGATCAACGCCTTCGGTTTCGTGGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCAAGCTGACCGACGAGCTG
ACCTTCGCCGAGGCCCGCGTCCTGGTGCGCCGTCGTGGCGGGGAGATCGACTACATC—
CCCGGCACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCGGTGCCCCTGATCAACGCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTCATCAAGGCCGAGCAGAACGGTGTCGTCTCCGACGTCTCCGCCGACTACGTCTCGATCGCCAACGACGACGGCACCCA
CGTCACCCATCGGCTGGCC

>NZ_CP045740.1_S._sp._SUK_48

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTG
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GGCGGCGCCGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACGGCCATGATCCCCTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGTACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGGCGAACGACGACGGCACG
TACATCACGTACCGCCTGCAC

>NZ_CP007699.2_*S._lydicus*

GCCGGCCTGGACGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGTCCCAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGGTCAACGTCTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGAGGAGGTGGACTACCTCACCGCCGATGAGGAGGACCGCTTCCTGATCGCGCAGGCCAACGCCAAGCTCAGCGACGACAT
GCGCTTCTCCGAGCAGCGTGTGCTGGTCCGCCGTCGTGGCGGCGAGGTCGACCTCGTC—
CCCGCCGACGAGGTGGACTTCATGGACGTCTCGCCGCGCCAGATGGTGTCGGCCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCGGAGTCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCGTCGCC

>NZ_CP034279.1_*S._ficellus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCTTACCGCAAGGTCGTCG---
AGGGTCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTC
CGGTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GCGCCCGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCGGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACGTCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>NZ_CP028369.1_S._sp._P3

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGTTTCGTGGAGACTCCGTACCGCAGGGTCGACG---
GCGACGTCGTCACCGACGAGGTCGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCGACGACGACATG
CGCTTCACCGAGAACCGCGTGCTGGTCCGCCGTCGTGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGACAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_AP018517.1_*S._rochei*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTG
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GGTGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCAAGTCCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTGGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACGGCGAACGACGACGGCACG
TACATCACGTACCGCCTGCAC

>NZ_CP026652.1_*S._dengpaensis*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCGTCGTCACCGACGAGGTCGACTACCTCACGGCGGACGAGGAGGACCGGTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGAACGAGCTC
CGCTTCGAGGAGTCCCGTGTTCTGGTTCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATTAAGTCCGAGGCGCCGCTCGTCGGTACCGGCATGGAGTACCGCTGCGC
CGTCGACGCCGGTGACGTACTCAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACGT
ACACCACGTACCGCCTCGCG

>NZ_CP015362.1_S._sp._S8

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGGCGCATGTGTCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACCCTCTCCGACGAGCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGCCGCGGGGGAGAGGTCGACTACGTC—
CCCGCCTCCGAGGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCACGT
TACACCACGTACCGCATCGCC

>NZ_CP013738.1_*S._globisporus*

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACCCTCAACGACGAGCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGCGACGTCCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>NZ_CP029601.1_S._sp._WAC_01438

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCCTACGGGCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCAGCGACGACATG
CGCTTCGCCGAGAACCGCATCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGAGCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC

CGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGCCTCGCC

>NZ_CP042324.1_*S._coelicolor*

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGCATCAACCCCTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGCGCTGGGCGACGACAT
GCGGTTCGCCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTCCGC
CGCCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP022310.1_*S._asterosporus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGAGATG
CGCTTCGTCGAGAACCGCATCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATCCGCTCCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGCC
GTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACGT
ACATCACGTACCGCCTGGCC

>NZ_CP023694.1_*S._coeruleorubidus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGTCCGATCGAGACCCCTGAAGGCCCGAACATCGGCCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTTGTCGAGACCCCGTACCGCAAGGTGTTCG---
AGGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGACCTC
CGCTTCGCCGAGGCCCGCGTCCTGGTCCGCCGCAAGGGCGGCGAGGTCGACTACGTC—
GGCGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCCAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CGGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>NZ_CP048397.1_S._sp._S4.7

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGTCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCGTACGGCCGGGTCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCGACGCTCTCCGAGGACATG
CGGTTCACCGAGCCGCGCGTCCTGGTCCGCCGTCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTTGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTACCGCTGATCACGGCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCGGGTGACGTCATCAAGGCGGAGAAGGAAGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP018047.1_*S._niveus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGTCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGGTCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCGACGCTCTCCGAGGACATG
CGGTTCACCGAGCCGCGCGTCCTGGTCCGCCGTCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTCCCGCTGATCACGGCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCG
CGGTCGACGCGGGTGACGTCATCAAGGCGGAGAAGGAAGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCAC
GTACACCACGTACCGCATCGCC

>NZ_CP018627.1_*S._hygroscopicus*

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---

ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTG
CGCTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCACC
TACACCACCTACCGGGTGGCC

>NZ_CP019779.1_S._sp._MOE7

GCCGGCCTGGACGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATTGAGACCCCTGAAGGTCCCAACATCGGTCTCATCGGCTCG
CTGGCCTCCTACGGCCGGGTCAACGTCTTCGGCTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTCAGGTCACCGAGGAGGTGGACTACCTCACCGCCGATGAGGAGGACCGCTTCCTGATCGCGCAGGCCAACGCGAAGCTCTCCGAGGACATG
CGGTTCTCCGAGCAGCGTGTGCTGGTCCGCCGTCGTGGCGGCGAGGTCGACCTCGTC—
CCCGCCGACGAGGTGGACTTCATGGACGTCTCGCCGCGCCAGATGGTGTCGGCCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCGGAGTCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTCATCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCGTCGCC

>NZ_CP016279.1_*S._griseochromogenes*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGTATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTGGAGACCCCGTACCGCAAGGTCGACGAGCACGGTGTGGTCACCGACGAGGTGGACT
ACCTGACCGCCGACGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGAACGACGAGATGCGGTTCGCCGAGGCCCGCGTCCTGGTC
CGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GCCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATTAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
GGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACTTACCGCCTGCAC

>NZ_CP031969.1_S._sp._CC0208

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTGGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGGTTCGAGGAGTCCCGGGTCCTGGTCCGCCGTCGTGGCGGCGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGCC
GTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACGT
ACATCACGTACCGCCTGGCC

>NZ_CP041654.1_S._sp._RLB1-9

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>NC_017586.1_*S._cattleya*

GCCGGTCTGGACGTGCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACCCCGTACCGCCGGGTCGTCG---
AGGGCCGGGTCACCGAGGACGTGCACTACCTGACCGCGGACGAGGAGGACCGCTTCCTCATCGCGCAGGCCAACGCGCCGCTCACCGAGGAGGG
CACCTTCGCCGAGGCCCGTGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACCTCGTC—
CCCTCGCAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTGGCCACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCCAACATGATGCGCCAGGCGGTGCCGCTGCTGAAGGCGGAGTCGCCGCTGGTCGGCACCGGCATCGAGTACCGCGCCGC
GGTCGACGCCGGTGACGTCATCACCGCCGAGAAGGACGGCGTGGTCCAGGAGGTCTCCGCGGACTACGTCACGGTGGCCAACGACGACGGCACC
TACACCACCTACCGGGTCGCC

>NZ_CP023747.1_*S._nodosus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCGACG---
AGGGCCAGGTCACCGACGAGGTCGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACCCTGACGGACGACCTG
CGCTTCGCCGAGAACCGTGTCCTGGTCCGCCGCCGCGGCGGTGAGGTCGACTACGTC—
AGCCCCTCGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTGATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGTCGGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGCC
GTCGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACGT
ACATCACCTACCGCCTGGCC

>NZ_CP050975.1_S._sp._RPA4-2

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCACGG---
GCGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGGTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CGTCGACGCCGGCGACGTGCTCAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>NZ_CP025018.1_S._sp._M56

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGTCCGATCGAGACACCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGAGGTCCACTACCTCACGGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCCCGCTGACGAACGACCTGC
ACTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGTGAGGTCGACTACATC—
CCCGGCGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGCGCCCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGTATGGAGTACCGCTGCG
CGGTCGACGCCGGCGATGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCAC
CTACACCACCTACCGGGTGGCC

>NZ_CP029254.1_*S._spongiicola*

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCCCACTACGGACGGATGTGCCCGATCGAGACCCCCGAGGGCCCGAACATCGGTCTGATCGGTTC
GCTCGCCACCTACGGCCGGGTGAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCGAGG---
GCGGCCAGGTCACCGACCAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGCGCTCGGCGACGACCTG
CGCTTCGCCGAGGCCCGCGTGCTCGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTG—
TCGCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTGGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGCGCCCTGATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCGGAGGCGCCGCTCGTCGGCACCGGTATGGAGTACCGCTGCG
CCACCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTGATCCAGGAGGTCTCGGCCGACTACGTCACCGTCGCCAACGACGACGGCAC
GTACACCACCTACCGCGTCGCG

>NZ_CP009124.1_*S._lividans*

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGCATCAACCCCTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGCGCTGGGCGACGACAT
GCGGTTCGCCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTCCGC
CGCCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP034463.1_*S._aquilus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGTATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTGGAGACGCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGAGGTCGACTACCTGACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCCACCCTCGACGAGGGCATG
CGCTTCACCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC

CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATTAAGTCCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGCC
GTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGGCCAACGACGACGGCACGT
ACATCACGTACCGCCTGGCC

>NZ_CP011492.1_S._sp._CNQ-509

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGTCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTCATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCCTTCGGCTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
GCGGCGTCGTCACCGACGACGTGGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCACGAACGACCTG
CACTTCGCCGAGAACCGGGTCCTGGTCCGCCGCCGCGGCGGTGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCCAACATGATGCGCCAGGCCGTGCCGCTGATCCGGGCCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGCGACGTCATCAAGGCCGACAAGGACGGCGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCGTCAAC

>NZ_CP023691.1_*S._platensis*

GCCGGCCTGGACGTCCGTGACGTGCACCCCTCGCACTACGGCCGTATGTGCCCGATTGAGACCCCTGAAGGTCCCAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGGTCAACGTCTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGAGGAGGTGGACTACCTCACCGCTGATGAGGAGGACCGCTTCCTGATCGCCCAGGCCAACGCGAAGCTCTCCGAGGACATG
CGGTTCTCCGAGCAGCGTGTCCTGGTCCGCCGTCGTGGCGGCGAGGTCGACCTCGTC—
CCCGCCGACGAGGTGGACTTCATGGACGTCTCGCCGCGCCAGATGGTGTCGGCCGCGACCGCCATGATTCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCGTCGCC

>NZ_CP011522.1_S._sp._CFMR_7

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGGCGCATGTGTCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGTCAGGTCACCGACGACGTCGACTACATCACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACCCTCTCCGACGAGCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGCCGCGGGGGGAGAGGTCGACTACGTC—
CCCGCCTCCGAGGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP045547.1_S._sp._SYP-A7193

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACCTG
CGGTTCGCCGAGGCCCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGCCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP047020.1_*S._broussonetiae*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGTATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCGACGAGCACGGTGTGGTCACCGACGAGGTGGACTA
CCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGAACGACGAGATGCGGTTCGCCGAGGCCCGTGTCCTGGTCC
GCCGCCGTGGTGGCGAGGTCGACTACGTC—
GCCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATCAAGTCCGAGGCACCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
GGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGTCTGGCC

>NZ_CP023202.1_*S._xinghaiensis*

GCCGGCTTCGAGGTCCGTGACGTCCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCCGGGTCACCG---
AGGGCATCGTCACCGACAGCGTGGACTACCTCACGGCCGACGAGGAAGACCGCTTTGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGACTTC
CACTTCGCCGAGGGCCGGGTCCTGGTCCGCCGCCGCGGCGGTGAGGTCGACTACGTC—
CCCGGGACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGCGCGCTCATGGGCTCCAACATGATGCGCCAGGCGGTGCCGCTGCTGAAGGCGGAGGCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCG
CGGTCGACGCCGGTGACGTGATCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCAC
CTACACCACCTACCGGGTGGCG

>NZ_CP034539.1_*S._cyaneochromogenes*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGTATGTGTCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACCTG
CGGTTCGCCGAGGCCCGCGTCCTGGTTCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GGCGGCGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CAGTCGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>NZ_CP019458.1_*S._autolyticus*

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGTCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGAGGTCCACTACCTCACGGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCCCGCTGACGAACGACCTGC
ACTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGTGAGGTCGACTACATC—
CCCGGCGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGCGCCCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGTATGGAGTACCGCTGCG
CGGTCGACGCCGGCGATGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCAC
CTACACCACCTACCGGGTGGCC

>NZ_CP023696.1_*S._fradiae*

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCATCG---
ACGGCCGGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACATG
CGCTTCGCCGAGAACCGCGTCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GCCGGGGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAA
CCGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATCCGCTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGCTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP027022.1_S._sp._WAC00288

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACGTCACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGCGCTCGACGACGAGCTG
CGCTTCTCCGAGAACCGCGTCCTGGTCCGCAAGCGCGGCGGCGAGGTCGACTACGTCGAGCCGTCG—
GACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAACCGTGCCCTC
ATGGGCGCGAACATGATGCGTCAGGCGGTGCCCCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGCCACCGACGC
CGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGCTGTCGGCCGACTACATCACGGTCGCCAACGACGACGGCACGTACATCACGT
ACCGCCTGCAC

>NZ_CP007574.1_*S._albulus*

GCCGGTCTGGACGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCGTCCTACGGCCGGGTCAACGTCTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGAGGAGGTGGACTACCTCACCGCGGACGAGGAGGACCGCTTCCTCATCGCCCAGGCCAACGCCAAGCTCAGCGAGGACCT
GCGCTTCTCCGAGCAGCGTGTGCTGGTCCGTCGTCGTGGCGGCGAGGTCGACCTCGTC—
CCCGCCGACGAGGTGCACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGCCGCGACCGCCATGATCCCGTTCCTTGAGCACGACGACGCCAAC
CGCGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCGGAGTCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC

GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGAACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCGTCGCC

>NZ_CP023689.1_*S._chartreusis*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGTATGTGCCCGATCGAGACGCCCGAAGGCCCCAACATCGGTCTGATCGGTTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGCTTCGTGGAGACGCCGTACCGCAAGGTCGTCG---
AGGGTCAGGTCACCGACGAGGTCGACTACCTGACGGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGCTTCACCGAGAACCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GCCGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCCCTCATCAAGTCCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP036534.1_S._sp._VN1

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGATCAACCCGTTCGGCTTCATCGAGACGCCCTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACCTG
CGGTTCGCCGAGGCCCGCGTCCTGGTCCGCCGCCGTGGCGGGGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTCCG
CCGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGCCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>NZ_CP038147.1_S._sp._S501

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP045096.1_S._sp._GY16

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGTATGTGTCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCCTACGGGCGCGTCAACGCGTTCGGTTTCGTGGAGACGCCCTACCGCCGGGTCACCG---
ACGGTGTCGTCACCGACGAGGTCGACTACCTGACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCCACGCTCGACGACGACATG
CGCTTCACCGAGAACCGCGTCCTGGTCCGCCGTCGTGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP047140.1_S._sp._Tu_2975

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCACCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACCCTGAGCGACGAACTG
CGCTTCACCGAGGCGCGCGTGCTGGTCCGCCGTCGTGGCGGCGAGGTCGACTACGTC—
CCGCCGACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTGGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGGATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACGTCACCGTCGCGAACGACGACGGTACG
TACACCACGTACCGCATCGCG

>NZ_CP020570.1_*S._violaceoruber*

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---

ACGGCCAGGTCACCGACGACGTCGACTACATCACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACCCTCAACGACGAGCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCTCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCCGAGGCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCG
CCACCGACGCCGGCGACGTCCTCAAGGCAGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCAC
GTACACCACGTACCGCATCGCC

>NZ_CP032229.1_*S._seoulensis*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGGTTCGAGGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTGTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGCAC

>NZ_CP043959.1_*S._tendae*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCCTACGGGCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTGG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCGCCGCTGAGCGACGACATG
CGCTTCGTCGAGAACCGCATCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCCGGCGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCCGTGCCGCTGATCAAGTCCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGTGACGTGGTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP050177.1_S._sp._QMT-12

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGGCGTATGTGTCCGATCGAGACCCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGGTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCGACGTTGTCCGAGGAGATG
CGGTTCACCGAGCCGCGCGTCCTCGTCCGCCGTCGTGGCGGCGAGGTCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCACGTCGGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCG
CGGTGGACGCCGGCGACGTCGTCAAGGCGGAGAAGGAGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCAC
GTACACCACGTACCGCATCGCC

>NZ_CP032698.1_*S._hundungensis*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACGTCACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGCGCTCAGCGACGACATG
CGGTTCACCGAGCCGCGCGTCGTGGTCCGCCGCCGTGGTGGCGAGGTCGACTACGTG—
CCCGGCACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATCACGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTCATCAAGGCGGAGAAGGCGGGTGTCGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGGATCGCC

>NZ_CP020569.1_*S._gilvosporeus*

GCCGGCCTGGACGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACTCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCGTCCTACGGCCGCGTCAACGTCTTCGGCTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGAGGAGGTGGACTACCTCACCGCCGATGAGGAGGACCGCTTCCTGATCGCGCAGGCCAACGCGAAGCTCTCCGAGGACATG
CGGTTCTCCGAGCAGCGTGTGCTGGTCCGCCGTCGTGGCGGCGAGGTCGACCTCGTC—
CCCGCCGACGAGGTGGACTTCATGGACGTCTCGCCGCGCCAGATGGTGTCGGCCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCGGAGTCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCGTCGCC

>NZ_CP025407.1_S._sp._CMB-StM0423

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGTCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTCATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCCTTCGGCTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
GCGGCGTCGTCACCGACGACGTGGACTACCTGACGGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGAACGACCTG
CACTTCGCCGAGAACCGGGTCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCCAACATGATGCGCCAGGCCGTGCCGCTGATCCGGGCCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCGGGCGACGTCATCAAGGCCGAGAAGGACGGCGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCGTCAAC

>NZ_CP030771.1_S._sp._YIM_121038

GCGGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGTCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGTCAGGTCACCGACGACGTCGACTACCTGACCGCTGACGAAGAGGACCGATTCGTCATCGCCCAGGCCAACGCGCCGCTGACCGACGAGCTC
CGGTTCGCCGAGGCCCGCGTCCTGGTCCGCCGCCGTGGCGGCGAAGTCGACTACGTC—
GCGCCGACGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGCCAA
CCGCGCGCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATCACCGCCGAGGCCCCCCTCGTCGGTACGGGCATGGAGTACCGCTGCG
CCGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCAC
GTACACCACGTACCGCGTCGCG

>NZ_CP013743.1_S._sp._CdTB01

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCTCCGACGACATG
CGGTTCGCCGAGAACCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GCCGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP021978.1_*S._hawaiiensis*

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCTTCG---
AGGGTCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGACCTC
CGCTTCGCCGAGGCACGAGTCCTGGTCCGCCGTAAGGGCGGCGAGGTCGACTACGTC—
ACCGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATTAAGTCCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
GGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGCCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP028719.1_S._sp._endophyte_N2

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTC
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GGTGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTGGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGGCGAACGACGACGGCAC
GTACATCACGTACCGCCTGCAC

>NZ_CP023698.1_*S._viridifaciens*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGTCCGATCGAGACCCCGGAAGGCCCGAACATCGGTCTGATCGGGTCG
CTGGCCTCGTACGGCCGGGTCAACGCCTTCGGCTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGAGAAGGTGGACTACCTCACCGCCGACGAGGAGGACCGCTACGTCATCGCCCAGGCGAACGCCCCGCTGACCGCGGACCTG
CACTTCGCCGAGCCGCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGATCGACTACATC—
CCCGGCACCGAGATCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC

CGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCGGTGCCGCTGCTGAAGAGCGAGGCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCG
CCGTCGACGCCGCGGACGTCATCACCGCCGAGAAGGCCGGTGTGGTGCAGGAGGTCTCGGCCGACTACGTCACCGTGGCCAACGACGACGGCAC
GTACACCACGTACCGCGCCGCC

>NZ_CP015849.1_S._sp._SAT1

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGGCCTCAACGACGACATG
CGCTTCTCCGAGAACCGCGTCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGTGACGTCGTCAAGGCCGAGAAGGCCGGTGTGGTGCAGGAGGTCTCCGCCGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGCAC

>NZ_CP034587.1_*S._luteoverticillatus*

GCCGGCTTCGAGGTCCGAGACGTCCACCCGTCGCACTACGGCCGTATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGTCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGAACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCCGCCCTGGGTGAGGACAT
GCGCTTCGCCGAGGCCCGCGTGCTGGTCCGTCGCCGTGGCGGCGAGATCGACTACATC—
GCGCCGGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGCGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCG
CGGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTCACGGTGGCCAACGACGACGGCAC
GTACAACACCTACCGCGTCGCC

>NZ_CP047146.1_S._sp._GS7

GCCGGTCTGGACGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCGTCGTACGGCCGGGTCAACGTCTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGAGGAGGTGGACTACCTCACCGCTGACGAGGAGGACCGCTTCCTGATCGCCCAGGCGAACGCGAAGCTCAGCGACGACAT
GCGCTTCGCCGAGCAGCGTGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACCTCGTC—
CCCGCCGACGAGGTGGACTTCATGGACGTCTCGCCGCGCCAGATGGTGTCGGCCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGCGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCGGAGTCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCG
CGGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCAC
GTACACCACGTACCGCGTCGCC

>NZ_CP029823.1_*S._malaysiensis*

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGTCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGAGGTCCACTACCTCACGGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCCCGCTGACGAACGACCTGC
ACTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGTGAGGTCGACTACATC—
CCCGGCGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGCGCCCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGTATGGAGTACCGCTGCG
CGGTCGACGCCGGCGATGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCAC
CTACACCACCTACCGGGTGGCC

>NZ_CP021118.1_S._sp._CLI2509

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTCGCCTCCTACGGTCGCGTCAACGCCTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGCCGCTCACGGACGACCTC
CACTTCGAGGAGGCCCGCGTCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTGATCAAGTCCGAGGCGCCCATCGTCGGCACCGGCATGGAGTACCGCTCCGC
GACCGACGCCGGCGACGTGCTCAAGGCCGACAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGACCGCGAACGACGACGGCACG
TACACCACGTACCGCCTCGCG

>NZ_CP023700.1_*S._viridosporus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCACCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCAACG---
ACGGCCAGGTCACCGACACCGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGACCTC
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCAAGGGCGGCGAGGTCGACTACGTC—
GGCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATTAAGTCCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGCC
GTCGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACGT
ACATCACGTACCGCCTGCAC

>NZ_CP053109.1_S._sp._Z423-1

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTTGCCACCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTGAACG---
ACGGCCAGGTCACCGACGAGGTGAACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAACTC
CGGTTCGCCGAGGCCCGCGTCCTGGTCCGCCGCCGTGGCGGTGAGGTCGACTACGTC—
GCCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATCCAGGCGGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CGGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACATCACGTACCGGCTGGCC

>NZ_LN831790.1_*S._leeuwenhoekii*

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGTATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTGGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCACCGACGAGATG
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCAAGGGCGGCGAGGTCGACTACGTC—
GGCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACGGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATTAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CGGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>NZ_CP030073.1_*S._cadmiisoli*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCACCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGAACTACCTGACCGCCGACGAGGAGGACCGGTTCGTCATCGCGCAGGCCAACGCGCCGCTCACCGACGAGCTC
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GCGCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCCAGGCGGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
GGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCCGGCGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP047144.1_S._sp._HF10

GCCGGCTTCGAGGTCCGCGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTC
CGTTTTGCCGAGGCCCGCGTGCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GGCGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGTACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGGCGAACGACGACGGCACG
TACATCACGTACCGCCTGCAC

>NZ_CP010833.1_S._sp._Tue_6075

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACCCTCAACGACGAGCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACCGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC

CACCGACGCCGGCGACGTCCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>NZ_CP023690.1_*S._spectabilis*

GCGGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCATCAACCCGTTCGGCTTCATCGAGACGCCCTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGACGTCGACTACCTGACCGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGAGGAGCT
CCGCTTCGCCGAGGCCCGCGTCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GCGCCGACGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGCGCGCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATCACCGCCGAGGCCCCCCTCGTCGGTACGGGCATGGAGTACCGCTGCG
CCGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCAC
GTACACCACGTACCGCGTCGCC

>NZ_CP024957.1_*S._cavourensis*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGGCGCATGTGTCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACCCTCTCCGACGAGCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGCCGCGGGGGAGAGGTCGACTACGTC—
CCCGCCTCCGAGGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP026730.1_S._sp._CB09001

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGCATCAACCCCTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGCGCTGGGCGACGACAT
GCGGTTCGCCGAGGCCCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTCCG
CCGCCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>NZ_CP032543.1_*S._griseus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGGCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGCCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACGGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGATCCTCTCCGACGAGCTG
CGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTG—
CCTGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTGGTCGGCACCGGTATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTGAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP042266.1_*S._qinzhouensis*

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACTCCGTACCGCAAGGTCACCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCCACGCTCACCGACGAGCTG
CGTTTCGCCGAGGCCCGGGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
CCGCCGACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCACGGCCGAGGCGCCGCTGGTCGGTACCGGCATGGAGTACCGCTGCGC
CGTCGACGCCGGTGACGTGATCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGGCAACGACGACGGCACC
TACACCACCTACCGGGTGGCC

>NZ_CP040244.1_*S._exfoliatus*

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---

ACGGTCAGGTCACCGACGAGGTCGACTACATCACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACCCTGAGCGACGAGCTC
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGCCGTGGTGGCGAGGTCGACTACGTC—
CCGCCGACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACGGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCCGAGGCCCCGCTCGTCGGTACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGTGACGTGCCGAAGGCCGACAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACATCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGGATGGCC

>NZ_CP050504.1_S._sp._DSM_40868

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCAGCGACCTC
CGCTTCGAGGAGAGCCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GGCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGTACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACGT
ACATCACGTACCGCCTGGCC

>NZ_CP015588.1_*S._alfalfae*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCCATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACCTCACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGACCTC
CACTTCGCCGAGGCCCGCGTCCTGATCCGCCGCCGCGGCGGCGAGATCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATCACCGCCGAGGCCCCCCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCGTCACCAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP029788.1_*S._actuosus*

GCCGGCTTCGAGGTCCGTGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCACCTACGGCCGCGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGAACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCACCGCCGACATG
CGGTTCGAGGAGGCCCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GCCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTGATTAAGTCCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGCC
GTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACGT
ACATCACGTACCGCCTGGCC

>NZ_AP019620.1_*S._antimycoticus*

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTCGCGGACGACCTG
CGCTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCACC
TACACCACCTACCGGGTGGCC

>NZ_CP046905.1_S._sp._QHH-9511

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACATCACGGCCGACGAGGAAGACCGTTACGTCATCGCCCAGGCGAACGCGACCCTCAGCGACGAGCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGCCGTGGTGGCGAGGTCGACTACGTC—
CCGCCGACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACGGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCACTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACGGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTGATCCAGGAGGTCTCGGCCGACTACATCACCGTCACCAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP015726.1_S._sp._RTd22

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCCACG---
AGGGTGTCGTCACCGACAAGGTGGACTTCCTCACCGCCGATGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCGCCGCTGGCCGAGGACAAC
ACCTACGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCTCCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGTCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGCGACGTCATCAAGGCCGACAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCACC
TACACCACGTACCGGATCGCC

>NZ_CP026121.1_S._sp._Go-475

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCACCTACGGCCGCGTGAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTGTTCG---
AGGGCCAGGTCACCGACCAGGTCGACTACCTCACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGACCTC
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCAAGGGCGGCGAGGTCGACTACGTC—
GGCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATTAAGTCCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
GGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGCCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP063373.1_*S._ferrugineus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGTATGTGCCCCATTGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTGAACG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGAGCGACCTC
CGCTTCGAGGAGGCCCGCGTACTCGTCCGCCGGCGTGGCGGCGAGGTCGACTACGTC—
ACCGGCGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>NZ_CP040752.1_*S._rectiverticillatus*

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACGGACGAGGTCGACTACCTGACTGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCCGCCCTCACCGAGGACATG
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTCACGGTGGCCAACGACGACGGCACG
TACAACACCTACCGGGTCGCG

>NZ_CP021707.1_S._sp._S063

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACCCTCAACGACGAGCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGCCGTGGGGGAGAGGTCGACTACGTC—
CCCGCGTCCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGCGACGTCCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>NZ_CP054939.1_S._sp._NA02536

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCAACGACGAGATG
CGCTTCGTCGAGAACCGCATCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC

CGCGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
GGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP041607.2_S._sp._S1D4-14

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>NZ_CP021748.1_*S._alboflavus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCCTCCTACGGCCGTGTGAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGGTTCGTCATCGCCCAGGCCAACGCGCCGCTGACCGATGACCTC
CGGTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGTGGCGAGGTCGACTACGTC—
GCCGGTTCCGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACGGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATCACCGCCGAGGCCCCCCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTGATCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCGTCACCAACGACGACGGCACG
TACACCACGTACCGCGTCGCG

>NZ_CP039123.1_S._sp._SS52

GCCGGCTTCGAGGTCCGCGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGATCAACCCGTTCGGCTTCATCGAGACGCCTTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGGTTCGCCGAGGCCCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTCCGC
CGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP054920.1_S._sp._NA03103

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGCATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGACATG
CGGTTCGCCGAGGCCCGCGTCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTCCG
CCGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGCCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>NZ_CP022545.1_S._sp._11-1-2

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGTCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTG
CGCTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCACC
TACACCACCTACCGGGTGGCC

>NZ_CP021744.1_*S._albireticuli*

GCCGGCTTCGAGGTCCGAGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCCTACGGCCGTGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
AGGGTGTCGTCACGGACGACGTCGACTACCTGACTGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCCGCCCTCACCGAGGACATG
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCGGAGGCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACGGTCGCCAACGACGACGGCACG
TACAACACGTACCGGGTCGCC

>NZ_CP029342.1_S._sp._SM18

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCGACGCTCTCCGACGAGCTG
CGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGCCGTGGCGGGGAGGTCGACTACGTG—
CCCGGCACGGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGTCCGAGGCCCCGCTCGTGGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTCCTGAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>NZ_CP063374.1_*S._chromofuscus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCACCTACGGCCGGGTCAATGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTGAACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGAT
GCGGTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GCGCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATTAAGTCCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGCG
GTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACGT
ACATCACCTACCGGCTGGCC

>NZ_CP003987.1_S._sp._769

GCCGGTCTGGACGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGCCGCGTCAACGTCTTCGGCTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGAGGAGGTGGACTACCTCACCGCAGATGAGGAGGACCGCTTCCTCATCGCCCAGGCCAACGCCAAGCTCTCCGACGACCTG
CGCTTCGCCGAGCAGCGTGTGCTGGTCCGTCGCCGTGGCGGCGAGGTCGACCTCGTC—
CCCGCCGACGAGGTGCACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGCCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCGGAGTCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGAACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCGTCGCC

>NZ_CP054929.1_*S._buecherae*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGGATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTGGAGACCCCGTACCGCAAGGTCACCG---
GCGGTCAGGTCACGGACGACGTGGACTACCTGACCGCGGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCACCGACGACCTG
CGGTTCGCCGAGTCCCGCGTGCTGGTGCGCCGCCGTGGTGGCGAGGTCGACTACATC—
CCGGGCGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACGGCCATGATCCCCTTCCTCGAGCACGACGACGCCAA
CCGCGCGCTCATGGGCTCCAACATGATGCGCCAGGCCGTGCCGCTGATTAAGGCGGAGGCGCCGCTGGTGGGCACCGGCATGGAGTACCGCTGCG
CGGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGACGGCGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCAC
GTACACCACGTACCGGGTCGCC

>NZ_CP048835.1_S._sp._So13.3

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGTCCGATCGAGACCCCGGAAGGCCCGAACATCGGCCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTGGAGACCCCGTACCGCAAGGTCGTCG---
AGGGCATCGTCACCGATGACGTCGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCAAACGCGCCGCTGACGGACGACAAC
CACTTCGAAGAGGCCCGAGTCCTGGTCCGCCGCCGTGGCGGCGAGATCGACTACATC—
CCCGGCGCCGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGATCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGAGCGAGGCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC

GGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACGGTCGCCAACGACGACGGCACC
TACACCACCTACCGGGTGGCG

>NZ_LR881953.1_S._sp._KY75

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGGCGCATGTGTCCGATCGAGACCCCTGAAGGCCCGAACATCGGCCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACGGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGATCCTCTCCGAGGAGCTG
CGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTCATCAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP066801.1_S._sp._HSG2

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGGATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCACCTACGGTCGGGTCAACGCGTTCGGCTTCGTGGAGACGCCCTACCGCAAGGTCATCG---
ACGGCCAGGTGACCGACCAGGTCAACTACCTCACCGCGGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGAGGACTTC
CGGTTCGCCGAGCCGCGGGTCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GGCGGTCAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACGGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGCGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTCCCGCTGATCCAGGCGGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGCCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCTCCAACGACGACGGCACC
TACATCACGTACCGGCTGGCC

>NZ_CP071839.1_*S._cyanogenus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGTATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCAGCGACCTC
CGTTTCGAGGAGAGCCGCGTGCTCGTCCGCCGTCGTGGCGGCGAGGTCGACTACGTC—
GGCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGGCCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NZ_CP050692.1_*S._antibioticus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCAACGACGATCTC
CGGTTCGAGGAGTCCCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCGCTCATCAAGAGCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGCC
GTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTCGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACGT
ACATCACGTACCGCCTGGCC

>NZ_CP058693.1_S._sp._Rer75

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCCACG---
AGGGTGTCGTCACCGACAAGGTGGACTTCCTCACCGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCGCCGCTGGCCGAGGACAAC
ACCTACGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGTGAGGTCGACTACGTG—
CCCGGCTCCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGTCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCACC
TACACCACCTACCGGATCGCC

>NZ_CP050974.1_S._sp._RLB1-33

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---

GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGTGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>NZ_CP065050.1_*S._solisilvae*

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGTCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGAGGTCCACTACCTCACGGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCCCGCTGACGAACGACCTGC
ACTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGTGAGGTCGACTACATC—
CCCGGCGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGCGCCCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGTATGGAGTACCGCTGCG
CGGTCGACGCCGGCGATGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCAC
CTACACCACCTACCGGGTGGCC

>NZ_CP070373.1_S._sp._MBT27

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACGTCACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGCGCTCAGTGACGAGATG
CGGTTCACCGAGCCGCGCGTCCTGGTTCGCCGCCGTGGTGGCGAGGTCGACTACGTG—
CCCGGCACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATCACGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTCATCAAGGCGGAGAAGGCGGGTGTCGTCCAGGAGGTTTCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGGATCGCC

>JAC70

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>NC_003155.5_*S._avermitilis*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCTTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCCGGGTCACCG---
ACGGTGTCGTCACCGACGAGGTCGACTACCTGACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTCGGCGACGACTTC
CGGTTCGAGGAGACCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATCAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CGTCGACGCCGGCGACGTGCTCAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>NC_013929.1_*S._scabiei*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTGGAGACGCCGTACCGCCGGGTCACCG---
ACGGTGTCGTCACCGACGAGGTCGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCGACGACGACATG
CGCTTCACCGAGAACCGCGTCCTGGTCCGCCGTCGTGGCGGCGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCCGACTACATCACCACCACCAACGACGACGGCACG
TACATCACGTACCGCCTGGCC

>NC_015957.1_*S._violaceusniger*

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTG
CGCTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGTCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCACC
TACACCACCTACCGGGTGGCC

>NZ_AP023439.1_*S._tuirus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTGGAGACCCCGTACCGCAAGGTGTTCG---
AGGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCAGCTGACGGACGACCTC
CGCTTCGCCGAGGCCCGGGTGCTCGTCCGCCGTAAGGGCGGCGAGGTCGACTACGTC—
ACCGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTGATTAAGTCCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGCGACGTGGTCAAGGCCGAGAAGCCCGGTGTGGTCCAGGAGGTCTCTGCGGACTACATCACCACCGCCAACGACGACGGCACGT
ACATCACGTACCGGCTGGCC

>NZ_AP023440.1_*S._glomeroaurantiacus*

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGCCTGATCGGCTCG
CTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCCTACCGCCGGGTCACCG---
ACGGTTTCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCGCCGCTGAGCGAGGAGTTC
CGGTTCGAGGAGGCGCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCTGTGCCGCTGATCACCGCCGAGGCTCCCCTCGTCGGTACGGGCATGGAGTACCGCTGCGCC
GTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTATCCGCGGACTACGTCACCATCGCCAACGACGACGGCACCTA
CAACACCTACCGGGTGGCC

>NZ_CP031455.1_*S._olivoreticuli*

GCCGGCTTCGAGGTCCGAGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCTTCGTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCGTGG---
GCGGCGTCGTCACCGACGACGTGGACTACCTGACTGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCGAACGCCCCCCTCACCGAGGACATG
CGTTACGCCGAGTCGCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCGGAGGCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACGTCATCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCGTCGCCAACGACGACGGCACG
TACAACACGTACCGGGTCGCC

>NZ_CP045702.1_*S._finlayi*

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGACGCGTCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGTCAGGTCACGGACGACGTCGACTACATCACGGCCGACGAAGAAGACCGTTACGTCATCGCCCAGGCGAACGCGCTCCTCTCCGAGGAGCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTG—
CCTGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGTCCGAGGCGCCGCTCGTCGGCACGGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_CP046907.1_S._sp._SCUT-3

GCCGGCTTCGAGGTCCGCGACGTCCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAGGGTCCGAACATCGGTCTGATCGGCTCC
CTGGCCTCCTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGCGTCGTCACCGACCAGGTGGACTTCCTCACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGAGC
ACCTTCGAGGAGGACCGCGTCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCCGGCAGCGAGATCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCCACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC

CGCGCGCTCATGGGGTCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGGCGGAGGCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGCGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCGCCAACGACGACGCCACCT
ACACCACGTACCGCGTCGCC

>NZ_CP053189.1_S._sp._jing01

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGTATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCCTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGAACGACCTC
CGGTTCGAGGAGAGCCGCGTGCTCGTCCGCCGTCGTGGCGGCGAGGTCGACTACGTC—
GGCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGCC
GTCGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCACGT
ACATCACGTACCGCCTGGCC

>NZ_CP054916.1_S._sp._NA02950

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGGATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTGGCCTCCTACGGCCGGGTCAATGCCTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCACCG---
ACGGTGTGGTCACCGACGGCGTGGACTTCCTCACCGCCGATGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCGCCGCTGGCCGACGACAAC
CGCTTCGCCGAGGCCCGGGTCCTGGTCCGCCGCCGTGGCGGTGAGGTCGACTACGTC—
CCCGGCTCCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGGTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGTATGGAGTACCGCTGCGC
GGTCGACGCCGGTGATGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTCACGGTGGCCAACGACGACGGCACCT
ACACCACCTACCGGGTGGCC

>NZ_CP054938.1_*S._harbinensis*

GCCGGTCTGGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCCATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCGGCGTTCGCGCGGATCAACGCCTTCGGTTTCGTGGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTGGTCACCGACGAGGTGAACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCAAGCTGACCGACGACTAC
CGGTTCGCCGAGGCCCGCGTCCTGGTGCGCCGCCGTGGTGGGGAGATCGACTACATC—
CCCGGCAGCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCGGTGCCGCTGATCAACGCCGAGGCTCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCG
CCACCGACGCGGGCGATGTCATCCGGGCCGAGCAGAACGGCGTCATCTCCGACGTCTCCGCCGACTACGTGACGATCGCCAACGACGACGGCACC
CACGTCACGCACCGGCTGTCG

>NZ_CP058556.1_S._sp._CB04723

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGCCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACCCTCAACGACGAGCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCCGAGGCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGCGACGTCCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>NZ_CP058977.1_S._sp._NEAU-sy36

GCCGGCTTCGAGGTCCGCGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTT
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GGCCCCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCGC
CGTCGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGGCCAACGACGACGGCACG
TACATCACGTACCGCCTGCAC

>NZ_CP059898.1_S._sp._M54

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGACCCTCAACGACGAGCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCCGAGGCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCG
CCACCGACGCCGGCGACGTCCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCAC
GTACACCACGTACCGCATCGCC

>NZ_CP059991.1_S._gardneri

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACGTCACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGCGCTCAACGACGACATG
CGGCTCACCGAGAACCGCGTCCTGGTGCGTAAGCGTGGCGGCGAGGTCGACTACGTCCTCCCGGCC—
GAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTC
ATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGCCACCGACGC
CGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGCTGTCGGCCGACTACGTCACGGTCGCCAACGACGACGGCACGTACATCACGT
ACCGCCTGCAC

>NZ_CP060825.1_S._sp._CRPJ-33

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGTCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTGGTCG---
ACGGTGTCGTCGGCGACGAGGTCGACTACCTCACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGTTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCCCTGATGGGCGCGAACATGATGAGGCAGGCCGTGCCGCTCATCAAGTCCGAGGCGCCGCTCGTGGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCATCGCCAACGACGACGGCACGT
ACACCACGTACCGCCTGCAC

>NZ_CP064057.1_S._sp._XC_2026

GCCGGTCTGGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCCAACATCGGTCTGATCGGCTCG
CTGGCCGCGTTCGGCCGGATCAACGCCTTCGGTTTCGTGGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCAAGCTGACCGACGAGCTG
ACCTTCGCCGAGGCCCGCGTCCTGGTGCGCCGTCGTGGCGGGGAGATCGACTACATC—
CCCGGCACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCGGTGCCCCTGATCAACGCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTCATCAAGGCCGAGCAGAACGGTGTCGTCTCCGACGTCTCCGCCGACTACGTCTCGATCGCCAACGACGACGGCACCCA
CGTCACCCACCGGCTGGCC

>NZ_CP071044.1_S._sp._CA-256286

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGGCGCATGTGTCCGATCGAGACCCCTGAAGGCCCGAACATCGGCCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGTCAGGTCACCGACGACGTCGACTACATCACGGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGATCCTCTCCGACGAGCTG
CGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGTCCGAGGCCCCGCTCGTGGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NZ_LR881952.1_S._sp._KY70

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGGCGCATGTGTCCGATCGAGACCCCTGAAGGCCCGAACATCGGCCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACGGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCGAACGCGATCCTCTCCGAGGAGCTG
CGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTCATCAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC

CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC1

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC17

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>NC_016582.1_*S._bingchenggensis*

GCGGGCTTCGAGGTCCGAGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGGAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGATCTG
CACTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
CCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGCGCGCTCATGGGATCGAACATGATGCGCCAGGCCGTGCCGCTGCTCAAGGCGGAGTCCCCGCTGGTCGGCACCGGTATGGAGTACCGCTGCG
CGGTCGACGCCGGTGACGTCATCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCAC
CTACAACACCTACCGGGTGGCC

>JAC35

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTG
CGCTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCACC
TACACCACCTACCGGGTGGCC

>JAC63

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTG
CGGTTCGCCGAGACCCGCGTCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCATGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>JAC53

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---

ACGGCCAGGTCACGGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTG
CGGTTCGCCGAGACCCGCGTCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>JAC65

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTG
CGGTTCGCCGAGACCCGCGTCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>JAC24

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC97

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATG
CGGTTCACCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGATCGACTACATC—
CCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCGGCACGGGCATGGAGTACCGCTGCG
CCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACACCACGTACCGCCTCGCG

>JAC124

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCACCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACAGGTTCGTCATCGCCCAGGCGAACGCGCTCCTCTCCGACGACCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGATGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGCGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCCCTGATCACCTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCGGGTGACGTCATCAAGGCGGAGAAGGAGGGTGTGGTCCAGGAGGTATCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC3

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC69

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCACCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACAGGTTCGTCATCGCCCAGGCGAACGCGCTCCTCTCCGACGACCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGATGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGCGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCCCTGATCACCTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCGGGTGACGTCATCAAGGCGGAGAAGGAGGGTGTGGTCCAGGAGGTATCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC80

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCGGGTCAGCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGTGCTGACCGACGACATG
CTGTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTGATCAAGGCGGAGGCCCCGCTCGTCGGTACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACGTCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCC

>JAC18

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC119

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC48

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTC
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GGTGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTGGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGGCGAACGACGACGGCAC
GTACATCACGTACCGCCTGCAC

>JAC64

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTG
CGGTTCGCCGAGACCCGCGTCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA

CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>JAC76

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>JAC86

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCACCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACAGGTTCGTCATCGCCCAGGCGAACGCGCTCCTCTCCGACGACCTG
CGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGATGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGCGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCCCTGATCACCTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCGGGTGACGTCATCAAGGCGGAGAAGGAGGGTGTGGTCCAGGAGGTATCCGCGGACTACATCACCGTCGCCAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC60

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTC
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GGTGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTGGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGGCGAACGACGACGGCAC
GTACATCACGTACCGCCTGCAC

>JAC78

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCGGGTCAGCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGTGCTGACCGACGACATG
CTGTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTGATCAAGGCGGAGGCCCCGCTCGTCGGTACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACGTCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCC

>JAC85L

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>JAC57

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTCGCCTCCTACGGGCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCGGCGACGACATG
CGCTTCGTCGAGAACCGCATCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACATCACGTACCGCCTCGCC

>JAC43

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC74C

GCCGGCCTGGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCCAACATCGGTCTGATCGGCTCG
CTGGCCGCGTTCGGCCGGATCAACGCCTTCGGTTTCGTGGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCAAGCTGACCGACGAGCTG
ACCTTCGCCGAGGCCCGCGTCCTGGTGCGCCGTCGTGGCGGGGAGATCGACTACATC—
CCCGGCACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGTCAGGCGGTGCCCCTGATCAACGCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTCATCAAGGCCGAGCAGAACGGTGTCGTCTCCGACGTCTCCGCCGACTACGTCTCGATCGCCAACGACGACGGCACCCA
CGTCACCCACCGGCTGGCC

>JAC107

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCGGGTCAGCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGTGCTGACCGACGACATG
CTGTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTGATCAAGGCGGAGGCCCCGCTCGTCGGTACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACGTCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCC

>JAC113

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATG
CGGTTCACCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGATCGACTACATC—
CCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCGGCACGGGCATGGAGTACCGCTGCG
CCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACACCACGTACCGCCTCGCG

>JAC72

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATG
CGGTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGATCGACTACATC—
CCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCGGCACGGGCATGGAGTACCGCTGCG

CCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACACCACGTACCGCCTCGCG

>JAC99

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATG
CGGTTCACCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGATCGACTACATC—
CCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCGGCACGGGCATGGAGTACCGCTGCG
CCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACACCACGTACCGCCTCGCG

>JAC100

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCgAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGGTCACGAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC13

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC19

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC20

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC22

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---

AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC26

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC36

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC40

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC98

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACCGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTGATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCG
CCGTCGACGCCGGCGACGTGCTCAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCAC
GTACACCACGTACCGCCTGCAC

>JAC103

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC104

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>JAC105

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACGTCACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGGGCTCGACGACGACCTG
CGCTTCACCGAGAACCGCGTCCTGGTCCGCAAGCGCGGCGGCGAGGTCGACTACGTCGAGCCGTCG—
GACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTC
ATGGGCGCGAACATGATGCGTCAGGCGGTGCCCCTCATCAAGTCGGAGGCCCCGCTCGTCGGTACCGGCATGGAGTACCGCTGCGCCACCGACGC
CGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGCTGTCGGCCGACTACATCACGGTCGCCAACGACGACGGCACGTACATCACGT
ACCGCCTGCAC

>JAC108

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC11

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGTACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC111

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC115

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC

CGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC116

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC117

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCGGGTCAGCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGTGCTGACCGACGACATG
CTGTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTGATCAAGGCGGAGGCCCCGCTCGTCGGTACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACGTCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCC

>JAC118

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCGGGTCAGCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGTGCTGACCGACGACATG
CTGTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTGATCAAGGCGGAGGCCCCGCTCGTCGGTACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACGTCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCC

>JAC120

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCGACGCTCTCCGACGAGCTG
CGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGCCGTGGCGGGGAGGTCGACTACGTG—
CCCGGCACGGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGTCCGAGGCCCCGCTCGTGGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTCCTGAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC121

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACGTCACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGGGCTCGACGACGACCTG
CGCTTCACCGAGAACCGCGTCCTGGTCCGCAAGCGCGGCGGCGAGGTCGACTACGTCGAGCCGTCG—
GACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTC
ATGGGCGCGAACATGATGCGTCAGGCGGTGCCCCTCATCAAGTCGGAGGCCCCGCTCGTCGGTACCGGCATGGAGTACCGCTGCGCCACCGACGC
CGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGCTGTCGGCCGACTACATCACGGTCGCCAACGACGACGGCACGTACATCACGT
ACCGCCTGCAC

>JAC126

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC127

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCGACGCTCTCCGACGAGCTG
CGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGCCGTGGCGGGGAGGTCGACTACGTG—
CCCGGCACGGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGTCCGAGGCCCCGCTCGTGGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTCCTGAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC128

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGTACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCTGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC15

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>JAC16

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>JAC21

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC

GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>JAC33

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>JAC34

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC37

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGTACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC38

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC39

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC41

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---

GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>JAC42

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC4B

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC51

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTG
CGGTTCGCCGAGACCCGCGTCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGTGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>JAC58

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTCGCCTCCTACGGGCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCGGCGACGACATG
CGCTTCGTCGAGAACCGCATCCTGGTCCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACATCACGTACCGCCTCGCC

>JAC6

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC61A

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTC
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GGTGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTGGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGGCGAACGACGACGGCAC
GTACATCACGTACCGCCTGCAC

>JAC61B

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTC
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GGTGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTGGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGGCGAACGACGACGGCAC
GTACATCACGTACCGCCTGCAC

>JAC7

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>JAC71

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>JAC73

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATG
CGGTTCACCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGATCGACTACATC—
CCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCGGCACGGGCATGGAGTACCGCTGCG
CCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACACCACGTACCGCCTCGCG

>JAC75

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC

CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>JAC79

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCGACGCTCTCCGACGAGCTG
CGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGCCGTGGCGGGGAGGTCGACTACGTG—
CCCGGCACGGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGTCCGAGGCCCCGCTCGTGGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTCCTGAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC8

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>JAC81

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>JAC82

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC85C

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>JAC87

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC9

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC90

GCCGGCTTCGAGGTCCCGGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGTACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCTGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC92

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC94

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC96

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTGGCCTCCTACGGCCGGGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AAGGGCAGGTCACCGACGATGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCATCGAGCCGCGTGTCCTCGTCCGCCGTCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAGGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC

CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGGTGACCAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC10

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC101

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACCGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTGATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCG
CCGTCGACGCCGGCGACGTGCTCAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCAC
GTACACCACGTACCGCCTGCAC

>JAC102

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGTACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCTGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC109

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATG
CGGTTCACCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGATCGACTACATC—
CCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCGGCACGGGCATGGAGTACCGCTGCG
CCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACACCACGTACCGCCTCGCG

>JAC112

GCCGGCCTGGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCCAACATCGGTCTGATCGGCTCG
CTGGCCGCGTTCGGCCGGATCAACGCCTTCGGTTTCGTGGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCAAGCTGACCGACGAGCTG
ACCTTCGCCGAGGCCCGCGTCCTGGTGCGCCGTCGTGGCGGGGAGATCGACTACATC—
CCCGGCACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGTCAGGCGGTGCCCCTGATCAACGCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTCATCAAGGCCGAGCAGAACGGTGTCGTCTCCGACGTCTCCGCCGACTACGTCTCGATCGCCAACGACGACGGCACCCA
CGTCACCCACCGGCTGGCC

>JAC114

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---

GCGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACCGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGACGCCAA
CCGTGCCCTGATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCG
CCGTCGACGCCGGCGACGTGCTCAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCAC
GTACACCACGTACCGCCTGCAC

>JAC23

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>JAC28

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTG
CGCTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCACC
TACACCACCTACCGGGTGGCC

>JAC30L

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGATGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>JAC45

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTG
CGGTTCGCCGAGACCCGCGTCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCATGACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>JAC46

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTG
CGGTTCGCCGAGACCCGCGTCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>JAC47

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTG
CGGTTCGCCGAGACCCGCGTCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGTGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>JAC54

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGTCGCGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCATCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGTCCGAGGACATG
CGCTTCACCGAGAACCGCGTGCTGGTCCGCCGTCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACTCGATCCGTGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACGTCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCG

>JAC55

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGTCGCGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCATCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGTCCGAGGACATG
CGCTTCACCGAGAACCGCGTGCTGGTCCGCCGTCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACTCGATCCGTGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACGTCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCG

>JAC59

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTG
CGGTTCGCCGAGACCCGCGTCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>JAC66

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACGGACGACGTGGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTG
CGGTTCGCCGAGACCCGCGTCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
GCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGACCAACGACGACGGCAC
GTACATCACGTACCGCCTGGCC

>JAC68

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTCGCCTCCTACGGGCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCCAGGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCGGCGACGACATG
CGCTTCGTCGAGAACCGCATCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGCCAA

CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACATCACGTACCGCCTCGCC

>JAC77

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC88

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACGTCACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGGGCTCGACGACGACCTG
CGCTTCACCGAGAACCGCGTCCTGGTCCGCAAGCGCGGCGGCGAGGTCGACTACGTCGAGCCGTCG—
GACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTC
ATGGGCGCGAACATGATGCGTCAGGCGGTGCCCCTCATCAAGTCGGAGGCCCCGCTCGTCGGTACCGGCATGGAGTACCGCTGCGCCACCGACGC
CGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGCTGTCGGCCGACTACATCACGGTCGCCAACGACGACGGCACGTACATCACGT
ACCGCCTGCAC

>JAC93

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATG
CGGTTCACCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGATCGACTACATC—
CCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCGGCACGGGCATGGAGTACCGCTGCG
CCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACACCACGTACCGCCTCGCG

>JAC125

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>JAC32

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTG
CGCTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCACC
TACACCACCTACCGGGTGGCC

>JAC49

GCAGGCATGGAGGTCCGTGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTCATCGGTTC
GCTCGCGTCGTTCGGTCGCATCAACCCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACTTGTCGGCGGACGAGGAGGACACGTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGCCGGACAAC
CACTTCGCCGAGGACCGCGTCATGGCGCGCACCAAGGGCGGAGAGGCGAGCGACGTC—
GCGATCGAGGACGTCGACTACATGGACGTCTCGGCGCGTCAGATGGTGTCGGCCGCGACCGCACTCATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCCCTCATGGGTGCCAACATGCAGCGTCAGGCGGTGCCGCTCGTTCGTTCGGAGGCTCCGCTCGTCGGTACCGGCATCGAGTACCGAGCGGC
ACTCGACTCGGGTGACGTCGTGCGAGCCACGAAGGCGGGTGTCGTGCAGGAGGTCTCCGCCGACCTCGTCACGGTCGCCAACGACGACGCGACGT
ACTCGACGTACCGCATCGCG

>JAC4A

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTG
CGCTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCACC
TACACCACCTACCGGGTGGCC

>JAC62

GCAGGCATGGAGGTCCGTGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTCATCGGTTC
GCTCGCGTCGTTCGGTCGCATCAACCCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACTTGTCGGCGGACGAGGAGGACACGTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGCCGGACAAC
CACTTCGCCGAGGACCGCGTCATGGCGCGCACCAAGGGCGGAGAGGCGAGCGACGTC—
GCGATCGAGGACGTCGACTACATGGACGTCTCGGCGCGTCAGATGGTGTCGGCCGCGACCGCACTCATCCCGTTCCTCGAGCACGACGACGCCAAC
CGCGCCCTCATGGGTGCCAACATGCAGCGTCAGGCGGTGCCGCTCGTTCGTTCGGAGGCTCCGCTCGTCGGTACCGGCATCGAGTACCGAGCGGC
ACTCGACTCGGGTGACGTCGTGCGAGCCACGAAGGCGGGTGTCGTGCAGGAGGTCTCCGCCGACCTCGTCACGGTCGCCAACGACGACGCGACGT
ACTCGACGTACCGCATCGCG

>JAC91

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATG
CGGTTCACCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGATCGACTACATC—
CCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCGGCACGGGCATGGAGTACCGCTGCG
CCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACACCACGTACCGCCTCGCG

>JAC106

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC12

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC

CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC50

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCG---
ACGGCCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTC
CGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTC—
GGTGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGTGGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACGGCGAACGACGACGGCAC
GTACATCACGTACCGCCTGCAC

>JAC84

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCG---
ACGGCATCGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTC
CACTTCGCCGAGGACCGCGTCCTGGTCCGCCGCCGTGGCGGAGAGGTCGACTACGTC—
CCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCACCAACGACGACGGCACG
TACACCACGTACCGCCTCGCC

>JAC95

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCGACGCTCTCCGACGAGCTG
CGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGCCGTGGCGGGGAGGTCGACTACGTG—
CCCGGCACGGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGTCCGAGGCCCCGCTCGTGGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTCCTGAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC122

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACCGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGCCAA
CCGTGCCCTGATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCG
CCGTCGACGCCGGCGACGTGCTCAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCAC
GTACACCACGTACCGCCTGCAC

>JAC56

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGTCGCGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCATCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGTCCGAGGACATG
CGCTTCACCGAGAACCGCGTGCTGGTCCGCCGTCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACTCGATCCGTGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACGTCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCG

>JAC67

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTCGCCTCCTACGGGCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---

AGGGCCAGGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCGGCGACGACATG
CGCTTCGTCGAGAACCGCATCCTGGTGCGCCGCCGCGGCGGCGAGGTCGACTACGTC—
CCCGGTGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTGGAGCACGACGACGACGCCAA
CCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTCCG
CCGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCACCGCCAACGACGACGGCAC
GTACATCACGTACCGCCTCGCC

>JAC89

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCGTTCCTGGAGCACGACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC25

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGTACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC14

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC2

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTG
CGCTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCACC
TACACCACCTACCGGGTGGCC

>JAC27

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTG
CGCTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCACC
TACACCACCTACCGGGTGGCC

>JAC30C

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGACCAGGTCACCGACGAGGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTC
CGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACGTG—
CCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACGGTCACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC31

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCGACGCTCTCCGAGGAGATG
CGGTTCACCGAGCCGCGTGTCCTGGTCCGCCGTCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCCACGGCCATGATCCCGTTCCTGGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATCCAGTCCGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACGTCCTCAAGGCGGACAAGGGCGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCATCGCCAACGACGACGGCACGT
ACTCGACGTACCGCATCGCC

>JAC5

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTG
CGCTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCACC
TACACCACCTACCGGGTGGCC

>JAC110

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCG---
GCGGTGTCGTCACCGACGAGGTCGACTACCTCACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTC
CGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACGTC—
CCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
GGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGTCACCACCGCCAACGACGACGGCACG
TACACCACGTACCGCCTGCAC

>JAC29

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
AGGGCGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAAGAGGACCGCTTCGTCGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATG
CGCTTCACCGAGGCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGATCGACTACATC—
CCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCGTTCCTCGAGCACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
CACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTGACGAACGACGACGGCACG
TACACCACGTACCGCATCGCC

>JAC52

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGTCGCGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCATCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGTCCGAGGACATG
CGCTTCACCGAGAACCGCGTGCTGGTCCGCCGTCGTGGCGGCGAGATCGACTACATC—
GCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGCCAAC

CGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCGCCGCTCGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGTGACTCGATCCGTGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACGTCACCGTCGCCAACGACGACGGCACGT
ACACCACGTACCGCGTCGCG

>JAC74L

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGACGTGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGAGGTCGACTACATCACCGCCGACGAGGAAGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGAGGACATG
CGGTTCACCGAGCCCCGCGTCCTGGTCCGCCGTCGTGGCGGAGAGGTCGACTACATC—
CCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCCCTTCCTGGAGCACGACGACGACGCCAAC
CGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTCGGTACGGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTCCTGAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC123

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCG---
ACGGCCAGGTCACCGACGACGTCGACTACATCACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCGACGCTCTCCGACGAGCTG
CGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGCCGTGGCGGGGAGGTCGACTACGTG—
CCCGGCACGGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCCGTTCCTGGAGCACGACGACGACGCCAAC
CGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGTCCGAGGCCCCGCTCGTGGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTCCTGAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATCACCGTCACCAACGACGACGGCACGT
ACACCACGTACCGCATCGCC

>JAC44

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTG
CGCTTCGCCGAGAACCGCGTCCTGGTCCGCCGCCGTGGCGGCGAGGTCGACTACATC—
CCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCCCTTCCTCGAGCACGACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTCGGCACCGGCATGGAGTACCGCTGTGC
GGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACGTGACGGTGGCCAACGACGACGGCACC
TACACCACCTACCGGGTGGCC

>JAC83

GCCGGCCTGGAAGTCCGTGAT-
TGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCCAACATCGGTCTGATCGGCTCGCTGGCCGCGTTCGGCCGGATCA
ACGCCTTCGGTTTCGTGGAGACCCCGTACCGCAAGGTCGTCG---
ACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCAAGCTGACCGACGAGCTG
ACCTTCGCCGAGGCCCGCGTCCTGGTGCGCCGTCGTGGCGGGGAGATCGACTACATC--
CCCGGCACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGACGCCAAC
CGCGCGCTCATGGGCTCGAACATGATGCGTCAGGCGGTGCCCCTGATCAACGCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTGCGC
CACCGACGCCGGTGACGTCATCAAGGCCGAGCAGAACGGTGTCGTCTCCGACGTCTCCGCCGACTACGTCTCGATCGCCAACGACGACGGCACCCA
CGTCACCCACCGGCTGGCC

A.2.3.3.2: *Streptomyces* isolates *rpoB* sequences compiled

The *rpoB* sequences were used to create the phylogenetic tree comprising the *Streptomyces* isolates sequences (n = 133). *M. tuberculosis* was used as an outgroup.

>*M_tuberculosis*_AP018036.1

GCCGGGCTGGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGGATGTGCCCGATCGAAACCCCTGAGGGGCCCAACATCGGTCTGATCGGCTC
GCTGTCGGTGTACGCGCGGGTCAACCCGTTCGGGTTCATCGAAACGCCGTACCGCAAGGTGGTCGACGGCGTGGTTAGCGACGAGATCGTGTACC
TGACCGCCGACGAGGAGGACCGCCACGTGGTGGCACAGGCCAATTCGCCGATCGATGCGGACGGTCGCTTCGTCGAGCCGCGCGTGCTGGTCCGC
CGCAAGGCGGGCGAGGTGGAGTACGTGCCCTCGTCTGAGGTGGACTACATGGACGTCTCGCCCCGCCAGATGGTGTCGGTGGCCACCGCGATGAT
TCCCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGGGCAAACATGCAGCGCCAGGCGGTGCCGCTGGTCCGTAGCGAGGCCCCGCTGG
TGGGCACCGGGATGGAGCTGCGCGCGGCGATCGACGCCGGCGACGTCGTCGTCGCCGAAGAAAGCGGCGTCATCGAGGAGGTGTCGGCCGACTA
CATCACTGTGATGCACGACAACGGCACCCGGCGTACCTACCGGATGCGC

>JAC70

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCGACGGCATCGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTCCACTTCGCCGAGGACCGCGTCCTGGTCCGCCG
CCGTGGCGGAGAGGTCGACTACGTCCCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CCTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCG
GCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCACCACCAACGACGACGGCACGTACACCACGTACCGCCTCGCC

>JAC63

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACGGACGACGTGGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTGCGGTTCGCCGAGACCCGCGTCCTGGTGCGCCG
CCGCGGCGGCGAGGTCGACTACGTCGCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTGGAGCATGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTCCGCCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCACGACCAACGACGACGGCACGTACATCACGTACCGCCTGGCC

>JAC1

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC17

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC53

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACGGACGACGTGGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTGCGGTTCGCCGAGACCCGCGTCCTGGTGCGCCG
CCGCGGCGGCGAGGTCGACTACGTCGCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTCCGCCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCACGACCAACGACGACGGCACGTACATCACGTACCGCCTGGCC

>JAC35

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCGACGGTGTCGTCACCGACGACGTCGACTACCT
GACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTGCGCTTCGCCGAGAACCGCGTCCTGGTCCGCC
GCCGTGGCGGCGAGGTCGACTACATCCCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCCTTCCTCGAGCACGACGACGCCAACCGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCGGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACG
TGACGGTGGCCAACGACGACGGCACCTACACCACCTACCGGGTGGCC

>JAC65

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACGGACGACGTGGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTGCGGTTCGCCGAGACCCGCGTCCTGGTGCGCCG
CCGCGGCGGCGAGGTCGACTACGTCGCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTCCGCCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCCGGACTACA
TCACCACGACCAACGACGACGGCACGTACATCACGTACCGCCTGGCC

>JAC24

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACGGTCACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC97

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCTG
ACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATGCGGTTCACCGAGAACCGCGTCCTGGTCCGCCG
CCGTGGCGGCGAGATCGACTACATCCCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCG
GCACGGGCATGGAGTACCGCTGCGCCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACAT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTCGCG

>JAC124

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCACCGACGGCCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACAGGTTCGTCATCGCCCAGGCGAACGCGCTCCTCTCCGACGACCTGCGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGT
CGTGGCGGAGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGATGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCC
GTTCCTGGAGCACGACGACGCCAACCGCGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCCCTGATCACCTCCGAGGCGCCGCTCGTCGG
CACCGGCATGGAGTACCGCTGCGCGGTCGACGCGGGTGACGTCATCAAGGCGGAGAAGGAGGGTGTGGTCCAGGAGGTATCCGCGGACTACATC
ACCGTCGCCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC3

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG

ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACGGTCACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC69

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCACCGACGGCCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACAGGTTCGTCATCGCCCAGGCGAACGCGCTCCTCTCCGACGACCTGCGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGT
CGTGGCGGAGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGATGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCC
GTTCCTGGAGCACGACGACGCCAACCGCGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCCCTGATCACCTCCGAGGCGCCGCTCGTCGG
CACCGGCATGGAGTACCGCTGCGCGGTCGACGCGGGTGACGTCATCAAGGCGGAGAAGGAGGGTGTGGTCCAGGAGGTATCCGCGGACTACATC
ACCGTCGCCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC80

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCGACGGCCGGGTCAGCGACGACGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGTGCTGACCGACGACATGCTGTTCGCCGAGAACCGCGTCCTGGTCCGCCG
CCGTGGCGGAGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTGATCAAGGCGGAGGCCCCGCTCGTC
GGTACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACG
TCACCGTCACCAACGACGACGGCACGTACACCACGTACCGCGTCGCC

>JAC18

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC119

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACGGTCACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC48

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCGACGGCCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTCCGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTCGGTGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCAAGTCCGAGGCGCCGCTCGTCG
GCACCGGCATGGAGTACCGCTCCGCCGTGGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCACGGCGAACGACGACGGCACGTACATCACGTACCGCCTGCAC

>JAC64

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACGGACGACGTGGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTGCGGTTCGCCGAGACCCGCGTCCTGGTGCGCCG
CCGCGGCGGCGAGGTCGACTACGTCGCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC

CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTCCGCCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCACGACCAACGACGACGGCACGTACATCACGTACCGCCTGGCC

>JAC76

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCGACGGCATCGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTCCACTTCGCCGAGGACCGCGTCCTGGTCCGCCG
CCGTGGCGGAGAGGTCGACTACGTCCCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CCTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCG
GCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCACCACCAACGACGACGGCACGTACACCACGTACCGCCTCGCC

>JAC86

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCGTACGGCCGGGTCAACGCGTTCGGCTTCGTCGAGACCCCGTACCGCAAGGTCACCGACGGCCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACAGGTTCGTCATCGCCCAGGCGAACGCGCTCCTCTCCGACGACCTGCGCTTCACCGAGCCGCGCGTCCTGGTCCGCCGT
CGTGGCGGAGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGATGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCCC
GTTCCTGGAGCACGACGACGCCAACCGCGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCCCTGATCACCTCCGAGGCGCCGCTCGTCGG
CACCGGCATGGAGTACCGCTGCGCGGTCGACGCGGGTGACGTCATCAAGGCGGAGAAGGAGGGTGTGGTCCAGGAGGTATCCGCGGACTACATC
ACCGTCGCCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC60

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCGACGGCCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTCCGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTCGGTGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCAAGTCCGAGGCGCCGCTCGTCG
GCACCGGCATGGAGTACCGCTCCGCCGTGGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCACGGCGAACGACGACGGCACGTACATCACGTACCGCCTGCAC

>JAC78

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCGACGGCCGGGTCAGCGACGACGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGTGCTGACCGACGACATGCTGTTCGCCGAGAACCGCGTCCTGGTCCGCCG
CCGTGGCGGAGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTGATCAAGGCGGAGGCCCCGCTCGTC
GGTACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACG
TCACCGTCACCAACGACGACGGCACGTACACCACGTACCGCGTCGCC

>JAC85L

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCGACGGCATCGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTCCACTTCGCCGAGGACCGCGTCCTGGTCCGCCG
CCGTGGCGGAGAGGTCGACTACGTCCCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CCTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCG
GCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCACCACCAACGACGACGGCACGTACACCACGTACCGCCTCGCC

>JAC57

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTCGCCTCCTACGGGCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCCAGGTCACCGACGACGTCGACTACCT
GACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCGGCGACGACATGCGCTTCGTCGAGAACCGCATCCTGGTCCGCC
GCCGCGGCGGCGAGGTCGACTACGTCCCCGGTGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATC
CCGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTCGTC

GGCACCGGCATGGAGTACCGCTCCGCCGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCACCGCCAACGACGACGGCACGTACATCACGTACCGCCTCGCC

>JAC43

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC74C

GCCGGCCTGGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCCAACATCGGTCTGATCGGCTCG
CTGGCCGCGTTCGGCCGGATCAACGCCTTCGGTTTCGTGGAGACCCCGTACCGCAAGGTCGTCGACGGTGTCGTCACCGACGACGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCAAGCTGACCGACGAGCTGACCTTCGCCGAGGCCCGCGTCCTGGTGCGCCG
TCGTGGCGGGGAGATCGACTACATCCCCGGCACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGCGCGCTCATGGGCTCGAACATGATGCGTCAGGCGGTGCCCCTGATCAACGCCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTCATCAAGGCCGAGCAGAACGGTGTCGTCTCCGACGTCTCCGCCGACTACGT
CTCGATCGCCAACGACGACGGCACCCACGTCACCCACCGGCTGGCC

>JAC107

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCGACGGCCGGGTCAGCGACGACGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGTGCTGACCGACGACATGCTGTTCGCCGAGAACCGCGTCCTGGTCCGCCG
CCGTGGCGGAGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTGATCAAGGCGGAGGCCCCGCTCGTC
GGTACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACG
TCACCGTCACCAACGACGACGGCACGTACACCACGTACCGCGTCGCC

>JAC113

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCTG
ACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATGCGGTTCACCGAGAACCGCGTCCTGGTCCGCCG
CCGTGGCGGCGAGATCGACTACATCCCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCG
GCACGGGCATGGAGTACCGCTGCGCCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACAT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTCGCG

>JAC72

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCTG
ACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATGCGGTTCGCCGAGAACCGCGTCCTGGTCCGCCG
CCGTGGCGGCGAGATCGACTACATCCCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCG
GCACGGGCATGGAGTACCGCTGCGCCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACAT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTCGCG

>JAC99

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCTG
ACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATGCGGTTCACCGAGAACCGCGTCCTGGTCCGCCG
CCGTGGCGGCGAGATCGACTACATCCCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCG
GCACGGGCATGGAGTACCGCTGCGCCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACAT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTCGCG

>JAC100

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCgAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACGGTCACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC13

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC19

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC20

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC22

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC26

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC36

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC40

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC98

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGACGTCGACTACCTG
ACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACCGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTGATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGTCCGAGGCGCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCGTCGACGCCGGCGACGTGCTCAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACG
TCACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC103

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTCACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC104

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGAGGTCGACTACCTC
ACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCG
GCACCGGCATGGAGTACCGCTGCGCGGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC105

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACCGACGAGGTCGACTACGTC
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGGGCTCGACGACGACCTGCGCTTCACCGAGAACCGCGTCCTGGTCCGCAA
GCGCGGCGGCGAGGTCGACTACGTCGAGCCGTCGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCCCTCATCAAGTCGGAGGCCCCGCTCGTCG
GTACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGCTGTCGGCCGACTACATC
ACGGTCGCCAACGACGACGGCACGTACATCACGTACCGCCTGCAC

>JAC108

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG

ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACGGTCACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC11

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGTACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC111

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACGGTCACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC115

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTCACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC116

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACGGTCACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC117

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCGACGGCCGGGTCAGCGACGACGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGTGCTGACCGACGACATGCTGTTCGCCGAGAACCGCGTCCTGGTCCGCCG
CCGTGGCGGAGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTGATCAAGGCGGAGGCCCCGCTCGTC
GGTACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACG
TCACCGTCACCAACGACGACGGCACGTACACCACGTACCGCGTCGCC

>JAC118

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCGACGGCCGGGTCAGCGACGACGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGTGCTGACCGACGACATGCTGTTCGCCGAGAACCGCGTCCTGGTCCGCCG
CCGTGGCGGAGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC

CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTGATCAAGGCGGAGGCCCCGCTCGTC
GGTACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTCATCAAGGCGGAGAAGGACGGTGTCATCCAGGAGGTCTCGGCCGACTACG
TCACCGTCACCAACGACGACGGCACGTACACCACGTACCGCGTCGCC

>JAC120

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACCGACGACGTCGACTACATC
ACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCGACGCTCTCCGACGAGCTGCGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGC
CGTGGCGGGGAGGTCGACTACGTGCCCGGCACGGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCC
GTTCCTGGAGCACGACGACGCCAACCGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGTCCGAGGCCCCGCTCGTGG
GCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTCCTGAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCGTCACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC121

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACCGACGAGGTCGACTACGTC
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGGGCTCGACGACGACCTGCGCTTCACCGAGAACCGCGTCCTGGTCCGCAA
GCGCGGCGGCGAGGTCGACTACGTCGAGCCGTCGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCCCTCATCAAGTCGGAGGCCCCGCTCGTCG
GTACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGCTGTCGGCCGACTACATC
ACGGTCGCCAACGACGACGGCACGTACATCACGTACCGCCTGCAC

>JAC126

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTCACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC127

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACCGACGACGTCGACTACATC
ACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCGACGCTCTCCGACGAGCTGCGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGC
CGTGGCGGGGAGGTCGACTACGTGCCCGGCACGGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCC
GTTCCTGGAGCACGACGACGCCAACCGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGTCCGAGGCCCCGCTCGTGG
GCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTCCTGAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCGTCACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC128

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGTACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCTGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC15

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGAGGTCGACTACCTC
ACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCG

GCACCGGCATGGAGTACCGCTGCGCGGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC16

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGAGGTCGACTACCTC
ACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCG
GCACCGGCATGGAGTACCGCTGCGCGGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC21

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGAGGTCGACTACCTC
ACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCG
GCACCGGCATGGAGTACCGCTGCGCGGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC33

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGAGGTCGACTACCTC
ACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCG
GCACCGGCATGGAGTACCGCTGCGCGGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC34

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC37

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGTACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC38

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC39

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC41

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGAGGTCGACTACCTC
ACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCG
GCACCGGCATGGAGTACCGCTGCGCGGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC42

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC4B

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACGGTCACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC51

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACGGACGACGTGGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTGCGGTTCGCCGAGACCCGCGTCCTGGTGCGCCG
CCGCGGCGGCGAGGTCGACTACGTCGCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGTGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTCCGCCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCACGACCAACGACGACGGCACGTACATCACGTACCGCCTGGCC

>JAC58

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTCGCCTCCTACGGGCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCCAGGTCACCGACGACGTCGACTACCT
GACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCGGCGACGACATGCGCTTCGTCGAGAACCGCATCCTGGTCCGCC
GCCGCGGCGGCGAGGTCGACTACGTCCCCGGTGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATC
CCGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTCCGCCGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCACCGCCAACGACGACGGCACGTACATCACGTACCGCCTCGCC

>JAC6

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACGGTCACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC61A

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCGACGGCCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTCCGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTCGGTGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCAAGTCCGAGGCGCCGCTCGTCG
GCACCGGCATGGAGTACCGCTCCGCCGTGGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCACGGCGAACGACGACGGCACGTACATCACGTACCGCCTGCAC

>JAC61B

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCGACGGCCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTCCGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTCGGTGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCAAGTCCGAGGCGCCGCTCGTCG
GCACCGGCATGGAGTACCGCTCCGCCGTGGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCACGGCGAACGACGACGGCACGTACATCACGTACCGCCTGCAC

>JAC7

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGAGGTCGACTACCTC
ACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCG
GCACCGGCATGGAGTACCGCTGCGCGGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC71

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGAGGTCGACTACCTC
ACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCG
GCACCGGCATGGAGTACCGCTGCGCGGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC73

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCTG
ACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATGCGGTTCACCGAGAACCGCGTCCTGGTCCGCCG
CCGTGGCGGCGAGATCGACTACATCCCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCG
GCACGGGCATGGAGTACCGCTGCGCCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACAT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTCGCG

>JAC75

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCGACGGCATCGTCACCGACGAGGTCGACTACCTG

ACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTCCACTTCGCCGAGGACCGCGTCCTGGTCCGCCG
CCGTGGCGGAGAGGTCGACTACGTCCCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CCTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCG
GCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCACCACCAACGACGACGGCACGTACACCACGTACCGCCTCGCC

>JAC79

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACCGACGACGTCGACTACATC
ACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCGACGCTCTCCGACGAGCTGCGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGC
CGTGGCGGGGAGGTCGACTACGTGCCCGGCACGGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCC
GTTCCTGGAGCACGACGACGCCAACCGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGTCCGAGGCCCCGCTCGTGG
GCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTCCTGAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCGTCACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC8

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGAGGTCGACTACCTC
ACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCG
GCACCGGCATGGAGTACCGCTGCGCGGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC81

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCGACGGCATCGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTCCACTTCGCCGAGGACCGCGTCCTGGTCCGCCG
CCGTGGCGGAGAGGTCGACTACGTCCCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CCTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCG
GCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCACCACCAACGACGACGGCACGTACACCACGTACCGCCTCGCC

JAC82

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACGGTCACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC85C

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCGACGGCATCGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTCCACTTCGCCGAGGACCGCGTCCTGGTCCGCCG
CCGTGGCGGAGAGGTCGACTACGTCCCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CCTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCG
GCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCACCACCAACGACGACGGCACGTACACCACGTACCGCCTCGCC

>JAC87

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC

CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTCACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC9

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC90

GCCGGCTTCGAGGTCCCGGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGTACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCTGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC92

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTCACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC94

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTCACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC96

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTGGCCTCCTACGGCCGGGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGAAGGGCAGGTCACCGACGATGTCGACTACCT
GACCGCCGACGAGGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCATCGAGCCGCGTGTCCTCGTCCGCCG
TCGTGGCGGCGAGGTCGACTACATCCCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCC
GTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAGGTCCGAGGCGCCGCTGGTCG
GCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACATC
ACGGTGACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC10

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTC

GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACGGTCACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC101

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGACGTCGACTACCTG
ACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACCGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTGATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGTCCGAGGCGCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCGTCGACGCCGGCGACGTGCTCAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACG
TCACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC102

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGTACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCTGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC109

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCTG
ACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATGCGGTTCACCGAGAACCGCGTCCTGGTCCGCCG
CCGTGGCGGCGAGATCGACTACATCCCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCG
GCACGGGCATGGAGTACCGCTGCGCCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACAT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTCGCG

>JAC112

GCCGGCCTGGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCCAACATCGGTCTGATCGGCTCG
CTGGCCGCGTTCGGCCGGATCAACGCCTTCGGTTTCGTGGAGACCCCGTACCGCAAGGTCGTCGACGGTGTCGTCACCGACGACGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCAAGCTGACCGACGAGCTGACCTTCGCCGAGGCCCGCGTCCTGGTGCGCCG
TCGTGGCGGGGAGATCGACTACATCCCCGGCACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGCGCGCTCATGGGCTCGAACATGATGCGTCAGGCGGTGCCCCTGATCAACGCCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTCATCAAGGCCGAGCAGAACGGTGTCGTCTCCGACGTCTCCGCCGACTACGT
CTCGATCGCCAACGACGACGGCACCCACGTCACCCACCGGCTGGCC

>JAC114

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGACGTCGACTACCTG
ACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACCGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTGATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGTCCGAGGCGCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCGTCGACGCCGGCGACGTGCTCAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACG
TCACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC23

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGAGGTCGACTACCTC
ACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCG
GCACCGGCATGGAGTACCGCTGCGCGGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC28

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCGACGGTGTCGTCACCGACGACGTCGACTACCT
GACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTGCGCTTCGCCGAGAACCGCGTCCTGGTCCGCC
GCCGTGGCGGCGAGGTCGACTACATCCCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCCTTCCTCGAGCACGACGACGCCAACCGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCGGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACG
TGACGGTGGCCAACGACGACGGCACCTACACCACCTACCGGGTGGCC

>JAC30L

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGAGGTCGACTACCTC
ACGGCCGATGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCG
GCACCGGCATGGAGTACCGCTGCGCGGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC45

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACGGACGACGTGGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTGCGGTTCGCCGAGACCCGCGTCCTGGTGCGCCG
CCGCGGCGGCGAGGTCGACTACGTCGCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTGGAGCATGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTCCGCCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCACGACCAACGACGACGGCACGTACATCACGTACCGCCTGGCC

>JAC46

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACGGACGACGTGGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTGCGGTTCGCCGAGACCCGCGTCCTGGTGCGCCG
CCGCGGCGGCGAGGTCGACTACGTCGCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTCCGCCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCACGACCAACGACGACGGCACGTACATCACGTACCGCCTGGCC

>JAC47

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACGGACGACGTGGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTGCGGTTCGCCGAGACCCGCGTCCTGGTGCGCCG
CCGCGGCGGCGAGGTCGACTACGTCGCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGTGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTCCGCCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCACGACCAACGACGACGGCACGTACATCACGTACCGCCTGGCC

>JAC54

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGTCGCGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCATCGACGGTGTCGTCACCGACGACGTCGACTACCTG
ACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGTCCGAGGACATGCGCTTCACCGAGAACCGCGTGCTGGTCCGCCG
TCGTGGCGGCGAGATCGACTACATCGCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCC
CTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCGCCGCTCGTCG
GCACCGGCATGGAGTACCGCTGTGCGGTCGACGCCGGTGACTCGATCCGTGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACGTC
ACCGTCGCCAACGACGACGGCACGTACACCACGTACCGCGTCGCG

>JAC55

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGTCGCGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCATCGACGGTGTCGTCACCGACGACGTCGACTACCTG
ACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGTCCGAGGACATGCGCTTCACCGAGAACCGCGTGCTGGTCCGCCG
TCGTGGCGGCGAGATCGACTACATCGCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCC
CTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCGCCGCTCGTCG
GCACCGGCATGGAGTACCGCTGTGCGGTCGACGCCGGTGACTCGATCCGTGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACGTC
ACCGTCGCCAACGACGACGGCACGTACACCACGTACCGCGTCGCG

>JAC59

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACGGACGACGTGGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTGCGGTTCGCCGAGACCCGCGTCCTGGTGCGCCG
CCGCGGCGGCGAGGTCGACTACGTCGCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTCCGCCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCACGACCAACGACGACGGCACGTACATCACGTACCGCCTGGCC

>JAC66

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACGGACGACGTGGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCACGCTCACCGACGACCTGCGGTTCGCCGAGACCCGCGTCCTGGTGCGCCG
CCGCGGCGGCGAGGTCGACTACGTCGCGGGTGACGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTGCCGCTCATCAAGAGCGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTCCGCCGTCGACGCCGGTGACGTGGTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCACGACCAACGACGACGGCACGTACATCACGTACCGCCTGGCC

>JAC68

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTCGCCTCCTACGGGCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCCAGGTCACCGACGACGTCGACTACCT
GACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCGGCGACGACATGCGCTTCGTCGAGAACCGCATCCTGGTGCGCC
GCCGCGGCGGCGAGGTCGACTACGTCCCCGGTGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATC
CCGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTCCGCCGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCACCGCCAACGACGACGGCACGTACATCACGTACCGCCTCGCC

>JAC77

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACGGTCACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC88

GCCGGCTTCGAGGTCCGAGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACCGACGAGGTCGACTACGTC
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCGGGCTCGACGACGACCTGCGCTTCACCGAGAACCGCGTCCTGGTCCGCAA
GCGCGGCGGCGAGGTCGACTACGTCGAGCCGTCGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCCCTCATCAAGTCGGAGGCCCCGCTCGTCG
GTACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGCTGTCGGCCGACTACATC
ACGGTCGCCAACGACGACGGCACGTACATCACGTACCGCCTGCAC

>JAC93

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCTG

ACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATGCGGTTCACCGAGAACCGCGTCCTGGTCCGCCG
CCGTGGCGGCGAGATCGACTACATCCCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCG
GCACGGGCATGGAGTACCGCTGCGCCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACAT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTCGCG

>JAC125

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCGACGGCATCGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTCCACTTCGCCGAGGACCGCGTCCTGGTCCGCCG
CCGTGGCGGAGAGGTCGACTACGTCCCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CCTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCG
GCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCACCACCAACGACGACGGCACGTACACCACGTACCGCCTCGCC

>JAC32

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCGACGGTGTCGTCACCGACGACGTCGACTACCT
GACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTGCGCTTCGCCGAGAACCGCGTCCTGGTCCGCC
GCCGTGGCGGCGAGGTCGACTACATCCCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCCTTCCTCGAGCACGACGACGCCAACCGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCGGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACG
TGACGGTGGCCAACGACGACGGCACCTACACCACCTACCGGGTGGCC

>JAC49

GCAGGCATGGAGGTCCGTGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTCATCGGTTC
GCTCGCGTCGTTCGGTCGCATCAACCCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACCGACGACGTCGACTACTT
GTCGGCGGACGAGGAGGACACGTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGCCGGACAACCACTTCGCCGAGGACCGCGTCATGGCGCGC
ACCAAGGGCGGAGAGGCGAGCGACGTCGCGATCGAGGACGTCGACTACATGGACGTCTCGGCGCGTCAGATGGTGTCGGCCGCGACCGCACTCA
TCCCGTTCCTCGAGCACGACGACGCCAACCGCGCCCTCATGGGTGCCAACATGCAGCGTCAGGCGGTGCCGCTCGTTCGTTCGGAGGCTCCGCTCG
TCGGTACCGGCATCGAGTACCGAGCGGCACTCGACTCGGGTGACGTCGTGCGAGCCACGAAGGCGGGTGTCGTGCAGGAGGTCTCCGCCGACCTC
GTCACGGTCGCCAACGACGACGCGACGTACTCGACGTACCGCATCGCG

>JAC4A

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCGACGGTGTCGTCACCGACGACGTCGACTACCT
GACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTGCGCTTCGCCGAGAACCGCGTCCTGGTCCGCC
GCCGTGGCGGCGAGGTCGACTACATCCCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCCTTCCTCGAGCACGACGACGCCAACCGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCGGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACG
TGACGGTGGCCAACGACGACGGCACCTACACCACCTACCGGGTGGCC

>JAC62

GCAGGCATGGAGGTCCGTGACGTCCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGTCTCATCGGTTC
GCTCGCGTCGTTCGGTCGCATCAACCCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACCGACGACGTCGACTACTT
GTCGGCGGACGAGGAGGACACGTTCGTCATCGCGCAGGCCAACGCGCCGCTGACGCCGGACAACCACTTCGCCGAGGACCGCGTCATGGCGCGC
ACCAAGGGCGGAGAGGCGAGCGACGTCGCGATCGAGGACGTCGACTACATGGACGTCTCGGCGCGTCAGATGGTGTCGGCCGCGACCGCACTCA
TCCCGTTCCTCGAGCACGACGACGCCAACCGCGCCCTCATGGGTGCCAACATGCAGCGTCAGGCGGTGCCGCTCGTTCGTTCGGAGGCTCCGCTCG
TCGGTACCGGCATCGAGTACCGAGCGGCACTCGACTCGGGTGACGTCGTGCGAGCCACGAAGGCGGGTGTCGTGCAGGAGGTCTCCGCCGACCTC
GTCACGGTCGCCAACGACGACGCGACGTACTCGACGTACCGCATCGCG

>JAC91

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCTCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTGGCTTCCTACGGTCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCTG
ACCGCGGACGAGGAAGACCGCTTCGTGATCGCCCAGGCCAACGCCGGCCTCGGCGAGGACATGCGGTTCACCGAGAACCGCGTCCTGGTCCGCCG
CCGTGGCGGCGAGATCGACTACATCCCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATCC

CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTCGTCG
GCACGGGCATGGAGTACCGCTGCGCCACTGACGCGGGCGACGTGCTCAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTTTCCGCGGACTACAT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTCGCG

>JAC106

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTCACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC12

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC50

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACCCCGTACCGCAAGGTCATCGACGGCCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAAGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACGGACGAGCTCCGCTTCGCCGAGGCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTCGGTGGCGAGGACGTGGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTGCCCCTGATCAAGTCCGAGGCGCCGCTCGTCG
GCACCGGCATGGAGTACCGCTCCGCCGTGGACGCCGGCGACGTCGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCACGGCGAACGACGACGGCACGTACATCACGTACCGCCTGCAC

>JAC84

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCCTTCGGCTTCGTCGAGACGCCCTACCGCCGGGTCACCGACGGCATCGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGACCGAGGACCTCCACTTCGCCGAGGACCGCGTCCTGGTCCGCCG
CCGTGGCGGAGAGGTCGACTACGTCCCCGGGGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CCTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTCCCGCTCATCAAGAGCGAGGCCCCGCTCGTCG
GCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTGAAGGCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCACCACCAACGACGACGGCACGTACACCACGTACCGCCTCGCC

>JAC95

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACCGACGACGTCGACTACATC
ACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCGACGCTCTCCGACGAGCTGCGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGC
CGTGGCGGGGAGGTCGACTACGTGCCCGGCACGGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCC
GTTCCTGGAGCACGACGACGCCAACCGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGTCCGAGGCCCCGCTCGTGG
GCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTCCTGAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCGTCACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC122

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGACGTCGACTACCTG
ACGGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGACCGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTGATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTGATCAAGTCCGAGGCGCCGCTCGTC

GGCACCGGCATGGAGTACCGCTGCGCCGTCGACGCCGGCGACGTGCTCAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACG
TCACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC56

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGTCGCGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCATCGACGGTGTCGTCACCGACGACGTCGACTACCTG
ACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGTCCGAGGACATGCGCTTCACCGAGAACCGCGTGCTGGTCCGCCG
TCGTGGCGGCGAGATCGACTACATCGCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCC
CTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCGCCGCTCGTCG
GCACCGGCATGGAGTACCGCTGTGCGGTCGACGCCGGTGACTCGATCCGTGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACGTC
ACCGTCGCCAACGACGACGGCACGTACACCACGTACCGCGTCGCG

>JAC67

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCGGAAGGCCCGAACATCGGCCTGATCGGCTC
GCTCGCCTCCTACGGGCGGATCAACCCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCCAGGTCACCGACGACGTCGACTACCT
GACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTCGGCGACGACATGCGCTTCGTCGAGAACCGCATCCTGGTGCGCC
GCCGCGGCGGCGAGGTCGACTACGTCCCCGGTGACGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATC
CCGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTCATCAAGAGCGAGTCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTCCGCCGCCGACGCCGGCGACGTGGTCAAGGCCGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCACCGCCAACGACGACGGCACGTACATCACGTACCGCCTCGCC

>JAC89

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGAGCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CGTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGGCAGGCGGTGCCGCTGATCAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGCGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTCACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC25

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGTACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC14

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC2

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCGACGGTGTCGTCACCGACGACGTCGACTACCT
GACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTGCGCTTCGCCGAGAACCGCGTCCTGGTCCGCC
GCCGTGGCGGCGAGGTCGACTACATCCCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCCTTCCTCGAGCACGACGACGCCAACCGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCGGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACG
TGACGGTGGCCAACGACGACGGCACCTACACCACCTACCGGGTGGCC

>JAC27

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCGACGGTGTCGTCACCGACGACGTCGACTACCT
GACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTGCGCTTCGCCGAGAACCGCGTCCTGGTCCGCC
GCCGTGGCGGCGAGGTCGACTACATCCCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCCTTCCTCGAGCACGACGACGCCAACCGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCGGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACG
TGACGGTGGCCAACGACGACGGCACCTACACCACCTACCGGGTGGCC

>JAC30C

GCCGGCTTCGAAGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTGGCCTCCTACGGCCGGATCAACGCGTTCGGTTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGACCAGGTCACCGACGAGGTCGACTACCTG
ACCGCCGACGAGGAGGACCGCTTCGTCATCGCGCAGGCCAACGCGCCGCTGACCGACGATCTCCGGTTCGCCGAGTCCCGCGTGCTGGTCCGCCG
CCGTGGCGGCGAGGTCGACTACGTGCCCGGCACGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTC
GGCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTGCTCAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACGGTCACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC31

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCCTACGGCCGCGTCAACGCGTTCGGCTTCGTCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCTG
ACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCGACGCTCTCCGAGGAGATGCGGTTCACCGAGCCGCGTGTCCTGGTCCGCCG
TCGTGGCGGCGAGGTCGACTACATCCCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCCACGGCCATGATCCC
GTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATCCAGTCCGAGGCGCCGCTCGTCG
GCACCGGCATGGAGTACCGCTGTGCGGTCGACGCCGGTGACGTCCTCAAGGCGGACAAGGGCGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCATCGCCAACGACGACGGCACGTACTCGACGTACCGCATCGCC

>JAC5

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCGACGGTGTCGTCACCGACGACGTCGACTACCT
GACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTGCGCTTCGCCGAGAACCGCGTCCTGGTCCGCC
GCCGTGGCGGCGAGGTCGACTACATCCCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCCTTCCTCGAGCACGACGACGCCAACCGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCGGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACG
TGACGGTGGCCAACGACGACGGCACCTACACCACCTACCGGGTGGCC

>JAC110

GCCGGCTTCGAGGTCCGTGACGTGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCCTCCTACGGCCGGGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCACCGGCGGTGTCGTCACCGACGAGGTCGACTACCTC
ACGGCCGACGAAGAGGACCGATTCGTCATCGCGCAGGCCAACGCGCCGCTGACGGACGAGCTCCGCTTCGAGGAGTCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACGTCCCCGGTGAGGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCC
CGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTTCCGCTCATCAAGTCCGAGGCGCCGCTGGTCG
GCACCGGCATGGAGTACCGCTGCGCGGTCGACGCCGGTGACGTACTGAAGTCGGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACGT
CACCACCGCCAACGACGACGGCACGTACACCACGTACCGCCTGCAC

>JAC29

GCCGGCTTCGAGGTCCGCGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGTTC
GCTGGCCTCCTACGGGCGCATCAACCCGTTCGGGTTCATCGAGACGCCGTACCGCAAGGTCGTCGAGGGCGTCGTCACCGACGACGTCGACTACCT
GACGGCCGACGAAGAGGACCGCTTCGTCGATCGCCCAGGCGAACGCGACGCTCTCCGACGACATGCGCTTCACCGAGGCCCGCGTCCTGGTCCGCC
GTCGTGGCGGAGAGATCGACTACATCCCCGGTGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCGTTCCTCGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCCGTCCCGCTCATCAAGAGCGAGGCGCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCCACCGACGCCGGCGACGTGCTCAAGGCGGAGAAGGCGGGTGTGGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTGACGAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC52

GCCGGCTTCGAGGTCCGTGACGTTCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGCTCG
CTCGCGTCCTACGGTCGCGTCAACGCGTTCGGTTTCGTCGAGACGCCGTACCGCAAGGTCATCGACGGTGTCGTCACCGACGACGTCGACTACCTG
ACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCCAACGCCGGCCTGTCCGAGGACATGCGCTTCACCGAGAACCGCGTGCTGGTCCGCCG
TCGTGGCGGCGAGATCGACTACATCGCCGGCGACGACGTCGACTACATGGACGTCTCCCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCCC
CTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGCCAGGCCGTTCCGCTCATCAAGGCGGAGGCGCCGCTCGTCG
GCACCGGCATGGAGTACCGCTGTGCGGTCGACGCCGGTGACTCGATCCGTGCGGAGAAGGACGGTGTCGTCCAGGAGGTCTCGGCCGACTACGTC
ACCGTCGCCAACGACGACGGCACGTACACCACGTACCGCGTCGCG

>JAC74L

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCTGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGACGTGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACCGACGAGGTCGACTACATC
ACCGCCGACGAGGAAGACCGCTTCGTGATCGCCCAGGCGAACGCGACGCTCTCCGAGGACATGCGGTTCACCGAGCCCCGCGTCCTGGTCCGCCG
TCGTGGCGGAGAGGTCGACTACATCCCGGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCGACCGCGATGATCC
CCTTCCTGGAGCACGACGACGCCAACCGTGCCCTCATGGGCGCGAACATGATGCGTCAGGCGGTGCCGCTGATTAAGTCGGAGGCCCCGCTCGTC
GGTACGGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTCCTGAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCGGACTACA
TCACCGTCACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC123

GCCGGCTTCGAGGTCCGTGACGTGCACCCGTCCCACTACGGACGCATGTGCCCGATCGAGACCCCCGAAGGCCCGAACATCGGTCTGATCGGTTCG
CTCGCCTCGTACGGCCGCGTCAACGCGTTCGGCTTCATCGAGACGCCGTACCGCAAGGTCGTCGACGGCCAGGTCACCGACGACGTCGACTACATC
ACGGCCGACGAGGAAGACCGCTTCGTCATCGCCCAGGCGAACGCGACGCTCTCCGACGAGCTGCGCTTCACCGAGCCCCGCGTCCTGGTCCGCCGC
CGTGGCGGGGAGGTCGACTACGTGCCCGGCACGGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCCGTCGCCACCGCGATGATCCC
GTTCCTGGAGCACGACGACGCCAACCGCGCCCTCATGGGCGCGAACATGATGCGCCAGGCGGTGCCGCTCATCAAGTCCGAGGCCCCGCTCGTGG
GCACCGGCATGGAGTACCGCTGCGCCACCGACGCCGGTGACGTCCTGAAGGCCGAGAAGGACGGTGTGGTCCAGGAGGTCTCCGCGGACTACAT
CACCGTCACCAACGACGACGGCACGTACACCACGTACCGCATCGCC

>JAC44

GCGGGCTTCGAGGTCCGTGACGTGCACCCGTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCGAACATCGGTCTGATCGGCTC
GCTCGCCTCGTACGGCCGGGTCAACGCGTTCGGTTTCATCGAGACCCCGTACCGCAAGGTCGTCGACGGTGTCGTCACCGACGACGTCGACTACCT
GACCGCCGATGAAGAGGACCGCTTCGTCATCGCGCAGGCCAACGCCCCGCTGGCGGACGACCTGCGCTTCGCCGAGAACCGCGTCCTGGTCCGCC
GCCGTGGCGGCGAGGTCGACTACATCCCCGGCGACGACGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCGATGATC
CCCTTCCTCGAGCACGACGACGCCAACCGCGCGCTCATGGGCTCGAACATGATGCGCCAGGCCGTGCCGCTGATCAAGGCGGAGTCCCCGCTGGTC
GGCACCGGCATGGAGTACCGCTGTGCGGTCGACGCCGGCGACGTCATCAAGGCCGAGAAGGACGGTGTCGTCCAGGAGGTCTCCGCCGACTACG
TGACGGTGGCCAACGACGACGGCACCTACACCACCTACCGGGTGGCC

>JAC83

GCCGGCCTGGAAGTCCGTGAT-
TGCACCCCTCGCACTACGGCCGCATGTGCCCGATCGAGACGCCCGAAGGCCCCAACATCGGTCTGATCGGCTCGCTGGCCGCGTTCGGCCGGATCA
ACGCCTTCGGTTTCGTGGAGACCCCGTACCGCAAGGTCGTCGACGGTGTCGTCACCGACGACGTCGACTACCTGACCGCCGACGAGGAGGACCGCT
TCGTCATCGCGCAGGCCAACGCCAAGCTGACCGACGAGCTGACCTTCGCCGAGGCCCGCGTCCTGGTGCGCCGTCGTGGCGGGGAGATCGACTAC
ATCCCCGGCACCGAGGTCGACTACATGGACGTCTCGCCGCGCCAGATGGTGTCGGTCGCGACCGCCATGATCCCGTTCCTCGAGCACGACGACGCC
AACCGCGCGCTCATGGGCTCGAACATGATGCGTCAGGCGGTGCCCCTGATCAACGCCGAGGCGCCGCTGGTCGGCACCGGCATGGAGTACCGCTG
CGCCACCGACGCCGGTGACGTCATCAAGGCCGAGCAGAACGGTGTCGTCTCCGACGTCTCCGCCGACTACGTCTCGATCGCCAACGACGACGGCAC
CCACGTCACCCACCGGCTGGCC

A.2.5.3.1a: Integration R-code for GC-MS before GC column cut

```
library(tidyverse)

library(stringr)

#write.csv(data2, "med-3-4-kovats.csv")

#getting the table of integrations into shape. Export as a text (tab delimited), put it in the working directory

intDF = read.table("med-3-4-int.txt", sep = "",skip = 2)

#here we read in the table of integrations

str(intDF) #check the structure

colnames(intDF)=c("RT","area","percent","SN","desc")#renaming the columns

intDF <- intDF %>% mutate("num" = 1:n())

head(intDF)

##########################################################################################################

#Now, we must read in another dataframe, that containing the chemical ID's, which we copied out of the

#Bruker method builder program

chemID=read.csv("med-3-4-ids.csv",head=F) #Read in the chem ID CSV corresponding to the integration dataframe from above

head(chemID)

str(chemID)

str(intDF)#check the structure, be sure column 1 contains RT's and col 3 contains ID's

chemID=data.frame(chemID$V1,chemID$V3,chemID$V8) #combine the only three parameters we want

colnames(chemID)=c("RT","ID","Ions") #rename the columns of the ID dataframe

str(chemID)

chemID$ID=as.character(chemID$ID) #make sure the ID column is character

chemID$RT=as.double(chemID$RT)

chemID$Ions=as.character(chemID$Ions)

chemID <- chemID %>% mutate("num" = 1:n())
```

```
##########################################################################
##############################################

#Now we will join the two dataframes

merged.df<-merge(x=intDF,y=chemID,by="num",all=TRUE)

str(merged.df)

tail(merged.df)

merged.df <- subset(merged.df, select = c("num", "SN","RT.y","area","ID","percent","Ions"))

names(merged.df)[names(merged.df) == "RT.y"] <- "RT"

##########################################################################
##############################################

#And now we will write this merged DF into a CSV

#write.csv(merged.df,"f18 hex_12-16-2019_mike standard apis db-5_liquid-final.csv")

#Be sure to give the final CSV an informative name!!! IE

#make sure the culture is stated, and that you indicate it is combined

#I preface this with a 1, so they will collect at the top of the folder

#1) load dplyr

library(dplyr)

#2) Load the txt file and set it to the container data

data=merged.df

#data<-read.csv("f18 hex_12-16-2019_mike standard apis db-5_liquid-final.csv")

head(data)

tail(data)

str(data)

#carbons=read.csv("carbons.csv")

#3) set the carbon standards

C08<-4.829

C09<-7.017

C10<-9.518
```

C11<-12.174

C12<-14.618

C13<-16.502

C14<-18.078

C15<-19.467

C16<-20.729

C17<-21.899

C18<-22.996

C19<-24.034

C20<-25.022

#C21<-25.902

#C22<-26.801

#C23<-27.712

#C24<-28.790

#C25<-30.201

#C26<-31.896

#C27<-34.067

#C28<-36.875

#C29<-40.524

#C30<-45.297

#C31<-51.49

#C32<-44.038

#C33<-47.437

#C34<-51.459

#C35<-56.635

#C36<-63.144

#C37<-71.263

#C38<-NA

```
#4) calculate the Kovats
# first separate by carbons
#(you will have to tailor this list to reflect how many hydrocarbons in the ladder you have)
part01<-data %>% filter(RT < C08)
part02<-data %>% filter(((RT > C08)&(RT < C09))|RT==C08)
part03<-data %>% filter(((RT > C09)&(RT < C10))|RT==C09)
part04<-data %>% filter(((RT > C10)&(RT < C11))|RT==C10)
part05<-data %>% filter(((RT > C11)&(RT < C12))|RT==C11)
part06<-data %>% filter(((RT > C12)&(RT < C13))|RT==C12)
part07<-data %>% filter(((RT > C13)&(RT < C14))|RT==C13)
part08<-data %>% filter(((RT > C14)&(RT < C15))|RT==C14)
part09<-data %>% filter(((RT > C15)&(RT < C16))|RT==C15)
part10<-data %>% filter(((RT > C16)&(RT < C17))|RT==C16)
part11<-data %>% filter(((RT > C17)&(RT < C18))|RT==C17)
part12<-data %>% filter(((RT > C18)&(RT < C19))|RT==C18)
part13<-data %>% filter(((RT > C19)&(RT < C20))|RT==C19)
part14<-data %>% filter(RT>=C20)
#part14<-data %>% filter(((RT > C20)&(RT < C21))|RT==C20)
#part15<-data %>% filter(((RT > C21)&(RT < C22))|RT==C21)
#part16<-data %>% filter(((RT > C22)&(RT < C23))|RT==C22)
#part17<-data %>% filter(((RT > C23)&(RT < C24))|RT==C23)
#part18<-data %>% filter(RT>=C24)
#part18<-data %>% filter(((RT > C24)&(RT < C25))|RT==C24)
#part19<-data %>% filter(((RT > C25)&(RT < C26))|RT==C25)
#part20<-data %>% filter(((RT > C26)&(RT < C27))|RT==C26)
#part21<-data %>% filter(((RT > C27)&(RT < C28))|RT==C27)
#part22<-data %>% filter(((RT > C28)&(RT < C29))|RT==C28)
#part23<-data %>% filter(((RT > C29)&(RT < C30))|RT==C29)
```

```
#part24<-data %>% filter(((RT > C30)&(RT < C31))|RT==C30)
#part25<-data %>% filter(((RT > C31)&(RT < C32))|RT==C31)
#part20<-data %>% filter(RT >= C31) #trimming off everything greater than C37
#part26<-data %>% filter(((RT > C32)&(RT < C33))|RT==C32)
#part27<-data %>% filter(((RT > C33)&(RT < C34))|RT==C33)
#part28<-data %>% filter(((RT > C34)&(RT < C35))|RT==C34)
#part29<-data %>% filter(((RT > C35)&(RT < C36))|RT==C35)
#part30<-data %>% filter(((RT > C36)&(RT < C37))|RT==C36)
#part31<-data %>% filter(((RT > C37)&(RT < C38))|RT==C36)
#part31<-data %>% filter(RT >= C37) #trimming off everything greater than C37
#part15<-data %>% filter(((RT > C21)&(RT < C22))|RT==C21)
#part16<-data %>% filter(((RT > C22)&(RT < C23))|RT==C22)
#part17<-data %>% filter(((RT > C23)&(RT < C24))|RT==C23)
#part18<-data %>% filter(((RT > C24)&(RT < C25))|RT==C24)
#part19<-data %>% filter(((RT > C25)&(RT < C26))|RT==C25)
#part20<-data %>% filter(((RT > C26)&(RT < C27))|RT==C26)
#part21<-data %>% filter(((RT > C27)&(RT < C28))|RT==C27)
#part22<-data %>% filter(((RT > C28)&(RT < C29))|RT==C28)
#part23<-data %>% filter(((RT > C29)&(RT < C30))|RT==C29)
#part24<-data %>% filter(((RT > C30)&(RT < C31))|RT==C30)
#part25<-data %>% filter(((RT > C31)&(RT < C32))|RT==C31)
#part26<-data %>% filter(((RT > C32)&(RT < C33))|RT==C32)
#part27<-data %>% filter(((RT > C33)&(RT < C34))|RT==C33)
#part23<-data %>% filter(RT > C29)
# second pull out just the retention times
head(data)
str(part01)
part01 <-part01$RT
```

```
part02 <-part02$RT
part03 <-part03$RT
part04 <-part04$RT
part05 <-part05$RT
part06 <-part06$RT
part07 <-part07$RT
part08 <-part08$RT
part09 <-part09$RT
part10 <-part10$RT
part11 <-part11$RT
part12 <-part12$RT
part13 <-part13$RT
part14 <-part14$RT
#part15 <-part15$RT
#part16 <-part16$RT
#part17 <-part17$RT
#part18 <-part18$RT
#part19 <-part19$RT
#part20 <-part20$RT
#part21 <-part21$RT
#part22 <-part22$RT
#part23 <-part23$RT
#part24 <-part24$RT
#part25 <-part25$RT
#part26 <-part26$RT
#part27 <-part27$RT
#part28 <-part28$RT
#part29<-part29$RT
```

```
#part30<-part30$RT
#part31<-part31$RT
# next calculate your Kovats
kov07<-700+(part01*0)
kov08 <-800+((part02-C08)/(C09-C08)*100)
kov09 <-900+((part03-C09)/(C10-C09)*100)
kov10<-1000+((part04-C10)/(C11-C10)*100)
kov11<-1100+((part05-C11)/(C12-C11)*100)
kov12<-1200+((part06-C12)/(C13-C12)*100)
kov13<-1300+((part07-C13)/(C14-C13)*100)
kov14<-1400+((part08-C14)/(C15-C14)*100)
kov15<-1500+((part09-C15)/(C16-C15)*100)
kov16<-1600+((part10-C16)/(C17-C16)*100)
kov17<-1700+((part11-C17)/(C18-C17)*100)
kov18<-1800+((part12-C18)/(C19-C18)*100)
kov19<-1900+((part13-C19)/(C20-C19)*100)
#kov20<-2000+((part14-C20)/(C21-C20)*100)
#kov21<-2100+((part15-C21)/(C22-C21)*100)
#kov22<-2200+((part16-C22)/(C23-C22)*100)
#kov23<-2300+((part17-C23)/(C24-C23)*100)
kov20<-NA
#kov25<-2500+((part19-C25)/(C26-C25)*100)
#kov26<-2600+((part20-C26)/(C27-C26)*100)
#kov27<-2700+((part21-C27)/(C28-C27)*100)
#kov28<-2800+((part22-C28)/(C29-C28)*100)
#kov29<-2900+((part23-C29)/(C30-C29)*100)
#kov30<-3000+((part24-C30)/(C31-C30)*100)
#kov31<-3100+((part25-C31)/(C32-C31)*100)
```

```
#kov32<-3200+((part26-C32)/(C33-C32)*100)

#kov33<-3300+((part27-C33/(C34-C33)*100))

#kov34<-3400+((part28-C34/(C35-C34)*100))

#kov35<-3500+((part29-C35/(C36-C35)*100))

#kov36<-3600+((part30-C36/(C37-C36)*100))

#kov37<-3700+((part31-C37/(10000000-C37)*100))

# Almagmate the data and put it back into the orginal

Kovats <-
c(kov07,kov08,kov09,kov10,kov11,kov12,kov13,kov14,kov15,kov16,kov17,kov18,kov1
9,kov20)#,

#kov21,kov22,kov23,kov24)#,kov24,kov25,kov26,kov27,kov28,kov29,kov30)#,kov31,k
ov31,kov33,kov34,#kov35,

#kov36,#kov37)#,kov33,kov34)we need to delete any parts we did not calculate

str(Kovats)

str(data)

Kovats=as.data.frame(Kovats)

Kovats <- Kovats %>% mutate("num" = 1:n())

data2 = full_join(data, Kovats,by="num")

data2$Kovats<-round(data2$Kovats,1)

head(data2)

tail(data2)

# save data2 (change the filename to reflect which files you processed)

write.csv(data2, "med-3-4-kovats.csv")
```

A.2.5.3.1b: Integration R-Code for GC-MS after GC column cut

```
library(tidyverse)

library(stringr)

#write.csv(data2, "med-3-4-kovats.csv")

#getting the table of integrations into shape. Export as a text (tab delimited), put it in the working directory

intDF = read.table("med-3-4-int.txt", sep = "",skip = 2)

#here we read in the table of integrations

str(intDF) #check the structure

colnames(intDF)=c("RT","area","percent","SN","desc")#renaming the columns

intDF <- intDF %>% mutate("num" = 1:n())

head(intDF)

##########################################################################################################

#Now, we must read in another dataframe, that containing the chemical ID's, which we copied out of the

#Bruker method builder program

chemID=read.csv("med-3-4-ids.csv",head=F) #Read in the chem ID CSV corresponding to the integration dataframe from above

head(chemID)

str(chemID)

str(intDF)#check the structure, be sure column 1 contains RT's and col 3 contains ID's

chemID=data.frame(chemID$V1,chemID$V3,chemID$V8) #combine the only three parameters we want

colnames(chemID)=c("RT","ID","Ions") #rename the columns of the ID dataframe

str(chemID)

chemID$ID=as.character(chemID$ID) #make sure the ID column is character

chemID$RT=as.double(chemID$RT)

chemID$Ions=as.character(chemID$Ions)

chemID <- chemID %>% mutate("num" = 1:n())
```

```
##########################################################################
##############################################

#Now we will join the two dataframes

merged.df<-merge(x=intDF,y=chemID,by="num",all=TRUE)

str(merged.df)

tail(merged.df)

merged.df <- subset(merged.df, select = c("num", "SN","RT.y","area","ID","percent","Ions"))

names(merged.df)[names(merged.df) == "RT.y"] <- "RT"

##########################################################################
##############################################

#And now we will write this merged DF into a CSV

#write.csv(merged.df,"f18 hex_12-16-2019_mike standard apis db-5_liquid-final.csv")

#Be sure to give the final CSV an informative name!!! IE

#make sure the culture is stated, and that you indicate it is combined

#I preface this with a 1, so they will collect at the top of the folder

#1) load dplyr

library(dplyr)

#2) Load the txt file and set it to the container data

data=merged.df

#data<-read.csv("f18 hex_12-16-2019_mike standard apis db-5_liquid-final.csv")

head(data)

tail(data)

str(data)

#carbons=read.csv("carbons.csv")

#3) set the carbon standards

C08<-4.752

C09<-6.936

C10<-9.435
```

C11<-12.082

C12<-14.546

C13<-16.436

C14<-18.016

C15<-19.405

C16<-20.666

C17<-21.834

C18<-22.931

C19<-23.969

C20<-24.955

#C21<-25.902

#C22<-26.801

#C23<-27.712

#C24<-28.790

#C25<-30.201

#C26<-31.896

#C27<-34.067

#C28<-36.875

#C29<-40.524

#C30<-45.297

#C31<-51.49

#C32<-44.038

#C33<-47.437

#C34<-51.459

#C35<-56.635

#C36<-63.144

#C37<-71.263

#C38<-NA

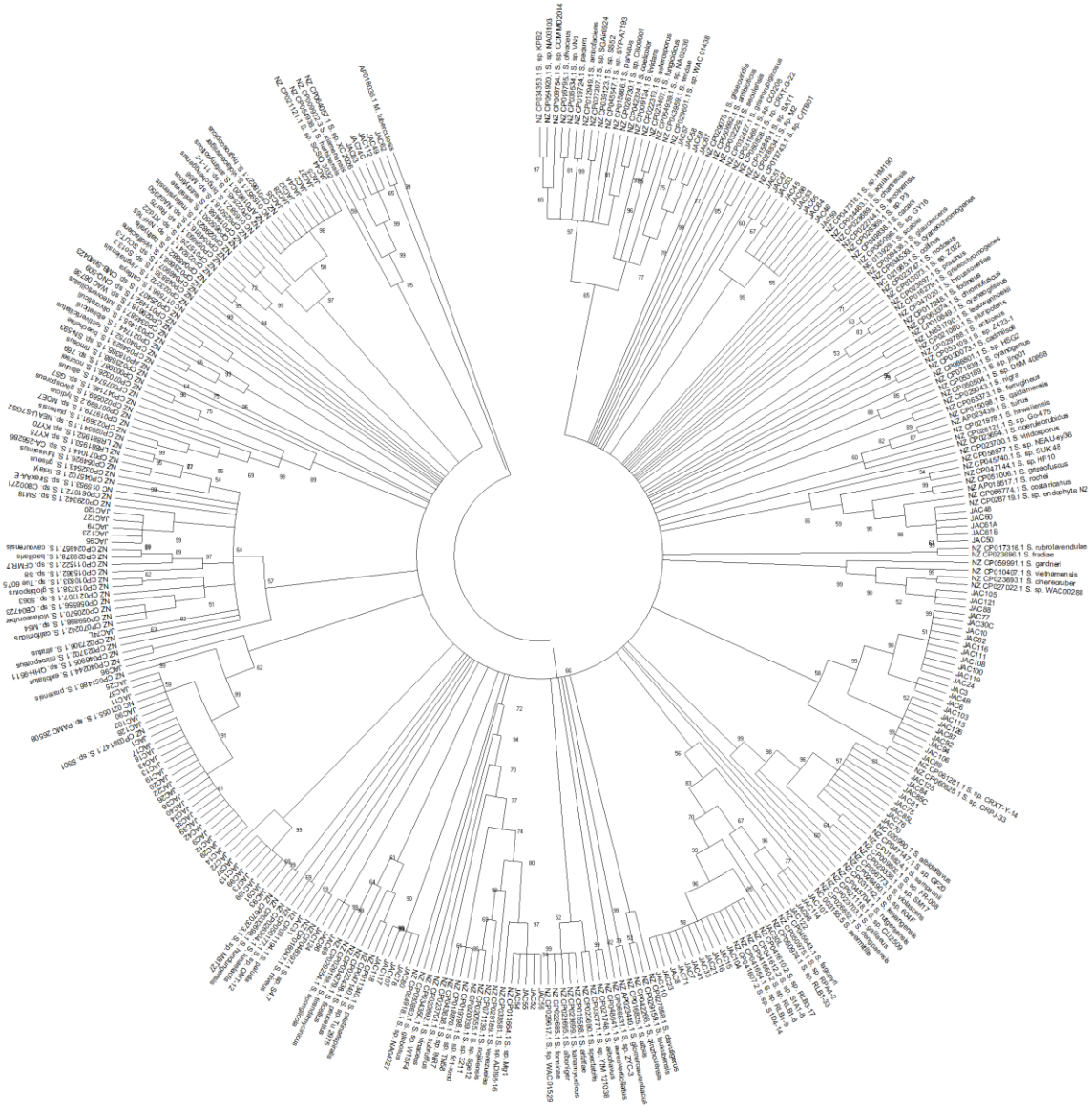
```
#4) calculate the Kovats
# first separate by carbons
#(you will have to tailor this list to reflect how many hydrocarbons in the ladder you have)
part01<-data %>% filter(RT < C08)
part02<-data %>% filter(((RT > C08)&(RT < C09))|RT==C08)
part03<-data %>% filter(((RT > C09)&(RT < C10))|RT==C09)
part04<-data %>% filter(((RT > C10)&(RT < C11))|RT==C10)
part05<-data %>% filter(((RT > C11)&(RT < C12))|RT==C11)
part06<-data %>% filter(((RT > C12)&(RT < C13))|RT==C12)
part07<-data %>% filter(((RT > C13)&(RT < C14))|RT==C13)
part08<-data %>% filter(((RT > C14)&(RT < C15))|RT==C14)
part09<-data %>% filter(((RT > C15)&(RT < C16))|RT==C15)
part10<-data %>% filter(((RT > C16)&(RT < C17))|RT==C16)
part11<-data %>% filter(((RT > C17)&(RT < C18))|RT==C17)
part12<-data %>% filter(((RT > C18)&(RT < C19))|RT==C18)
part13<-data %>% filter(((RT > C19)&(RT < C20))|RT==C19)
part14<-data %>% filter(RT>=C20)
#part14<-data %>% filter(((RT > C20)&(RT < C21))|RT==C20)
#part15<-data %>% filter(((RT > C21)&(RT < C22))|RT==C21)
#part16<-data %>% filter(((RT > C22)&(RT < C23))|RT==C22)
#part17<-data %>% filter(((RT > C23)&(RT < C24))|RT==C23)
#part18<-data %>% filter(RT>=C24)
#part18<-data %>% filter(((RT > C24)&(RT < C25))|RT==C24)
#part19<-data %>% filter(((RT > C25)&(RT < C26))|RT==C25)
#part20<-data %>% filter(((RT > C26)&(RT < C27))|RT==C26)
#part21<-data %>% filter(((RT > C27)&(RT < C28))|RT==C27)
```

```
#part22<-data %>% filter(((RT > C28)&(RT < C29))|RT==C28)
#part23<-data %>% filter(((RT > C29)&(RT < C30))|RT==C29)
#part24<-data %>% filter(((RT > C30)&(RT < C31))|RT==C30)
#part25<-data %>% filter(((RT > C31)&(RT < C32))|RT==C31)
#part20<-data %>% filter(RT >= C31) #trimming off everything greater than C37
#part26<-data %>% filter(((RT > C32)&(RT < C33))|RT==C32)
#part27<-data %>% filter(((RT > C33)&(RT < C34))|RT==C33)
#part28<-data %>% filter(((RT > C34)&(RT < C35))|RT==C34)
#part29<-data %>% filter(((RT > C35)&(RT < C36))|RT==C35)
#part30<-data %>% filter(((RT > C36)&(RT < C37))|RT==C36)
#part31<-data %>% filter(((RT > C37)&(RT < C38))|RT==C36)
#part31<-data %>% filter(RT >= C37) #trimming off everything greater than C37
#part15<-data %>% filter(((RT > C21)&(RT < C22))|RT==C21)
#part16<-data %>% filter(((RT > C22)&(RT < C23))|RT==C22)
#part17<-data %>% filter(((RT > C23)&(RT < C24))|RT==C23)
#part18<-data %>% filter(((RT > C24)&(RT < C25))|RT==C24)
#part19<-data %>% filter(((RT > C25)&(RT < C26))|RT==C25)
#part20<-data %>% filter(((RT > C26)&(RT < C27))|RT==C26)
#part21<-data %>% filter(((RT > C27)&(RT < C28))|RT==C27)
#part22<-data %>% filter(((RT > C28)&(RT < C29))|RT==C28)
#part23<-data %>% filter(((RT > C29)&(RT < C30))|RT==C29)
#part24<-data %>% filter(((RT > C30)&(RT < C31))|RT==C30)
#part25<-data %>% filter(((RT > C31)&(RT < C32))|RT==C31)
#part26<-data %>% filter(((RT > C32)&(RT < C33))|RT==C32)
#part27<-data %>% filter(((RT > C33)&(RT < C34))|RT==C33)
#part23<-data %>% filter(RT > C29)
# second pull out just the retention times
head(data)
```

```
str(part01)

part01 <-part01$RT
part02 <-part02$RT
part03 <-part03$RT
part04 <-part04$RT
part05 <-part05$RT
part06 <-part06$RT
part07 <-part07$RT
part08 <-part08$RT
part09 <-part09$RT
part10 <-part10$RT
part11 <-part11$RT
part12 <-part12$RT
part13 <-part13$RT
part14 <-part14$RT
#part15 <-part15$RT
#part16 <-part16$RT
#part17 <-part17$RT
#part18 <-part18$RT
#part19 <-part19$RT
#part20 <-part20$RT
#part21 <-part21$RT
#part22 <-part22$RT
#part23 <-part23$RT
#part24 <-part24$RT
#part25 <-part25$RT
#part26 <-part26$RT
```

```
#part27 <-part27$RT
#part28 <-part28$RT
#part29<-part29$RT
#part30<-part30$RT
#part31<-part31$RT
# next calculate your Kovats
kov07<-700+(part01*0)
kov08 <-800+((part02-C08)/(C09-C08)*100)
kov09 <-900+((part03-C09)/(C10-C09)*100)
kov10<-1000+((part04-C10)/(C11-C10)*100)
kov11<-1100+((part05-C11)/(C12-C11)*100)
kov12<-1200+((part06-C12)/(C13-C12)*100)
kov13<-1300+((part07-C13)/(C14-C13)*100)
kov14<-1400+((part08-C14)/(C15-C14)*100)
kov15<-1500+((part09-C15)/(C16-C15)*100)
kov16<-1600+((part10-C16)/(C17-C16)*100)
kov17<-1700+((part11-C17)/(C18-C17)*100)
kov18<-1800+((part12-C18)/(C19-C18)*100)
kov19<-1900+((part13-C19)/(C20-C19)*100)
#kov20<-2000+((part14-C20)/(C21-C20)*100)
#kov21<-2100+((part15-C21)/(C22-C21)*100)
#kov22<-2200+((part16-C22)/(C23-C22)*100)
#kov23<-2300+((part17-C23)/(C24-C23)*100)
kov20<-NA
#kov25<-2500+((part19-C25)/(C26-C25)*100)
#kov26<-2600+((part20-C26)/(C27-C26)*100)
#kov27<-2700+((part21-C27)/(C28-C27)*100)
#kov28<-2800+((part22-C28)/(C29-C28)*100)
```

```
#kov29<-2900+((part23-C29)/(C30-C29)*100)

#kov30<-3000+((part24-C30)/(C31-C30)*100)

#kov31<-3100+((part25-C31)/(C32-C31)*100)

#kov32<-3200+((part26-C32)/(C33-C32)*100)

#kov33<-3300+((part27-C33/(C34-C33)*100))

#kov34<-3400+((part28-C34/(C35-C34)*100))

#kov35<-3500+((part29-C35/(C36-C35)*100))

#kov36<-3600+((part30-C36/(C37-C36)*100))

#kov37<-3700+((part31-C37/(10000000-C37)*100))

# Almagmate the data and put it back into the orginal

Kovats <-
c(kov07,kov08,kov09,kov10,kov11,kov12,kov13,kov14,kov15,kov16,kov17,kov18,kov1
9,kov20)#,
#kov21,kov22,kov23,kov24)#,kov24,kov25,kov26,kov27,kov28,kov29,kov30)#,kov31,k
ov31,kov33,kov34,#kov35,

#kov36,#kov37)#,kov33,kov34)we need to delete any parts we did not calculate

str(Kovats)

str(data)

Kovats=as.data.frame(Kovats)

Kovats <- Kovats %>% mutate("num" = 1:n())

data2 = full_join(data, Kovats,by="num")

data2$Kovats<-round(data2$Kovats,1)

head(data2)

tail(data2)

# save data2 (change the filename to reflect which files you processed)

write.csv(data2, "med-3-4-kovats.csv")
```



A.3.1.1: Phylogenetic comparison of *Streptomyces* isolates to *Streptomyces* reference sequences. The phylogenetic tree consists of 345 sequences, 211 Streptomyces and 133 isolates rpoB sequences. M. tuberculosis served as an outgroup

A.3.5.1: AntiSMASH predicted BGCs in the genome of *S. griseofuscus*.

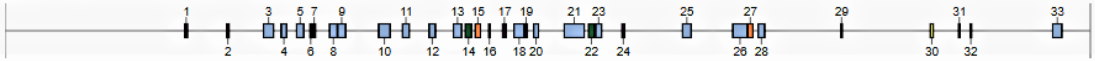
NZ_CP051006.1 (Streptomyces griseofuscus)

Region	Type	From	To	Most similar known cluster		Similarity
Region 1	T1PKS, NRPS, NRPS-like	112,034	176,874	nosiheptide	RiPP	8%
Region 2	NAPAA	315,216	347,707			
Region 3	siderophore	371,242	381,585	natamycin	Polyketide	9%
Region 4	other, lanthipeptide-class-iv	392,231	431,617	A-503083 A / A-503083 B / A-503083 E / A-503083 F	NRP	7%
Region 5	terpene	472,998	492,936	ebelactone	Polyketide	5%
Region 6	NRPS-like, NRPS, T1PKS	493,026	576,592	conglobatin	NRP	10%
Region 7	RiPP-like	748,506	756,410	informatipeptin	RiPP:Lanthipeptide	28%
Region 8	T1PKS	783,802	825,096	versipelostatin	Polyketide	5%
Region 9	lanthipeptide-class-i, lanthipeptide-class-ii, NRPS	896,103	956,839			
Region 10	T1PKS, NRPS-like, transAT-PKS-like	1,071,986	1,121,491	cinnabaramide A	NRP + Polyketide:Modular type I	18%
Region 11	thiopeptide, LAP, terpene	1,283,460	1,335,655	hopene	Terpene	92%
Region 12	NRPS, NRPS-like, betalactone	1,421,074	1,508,569	kirromycin	NRP + Polyketide:Modular type I + Polyketide:Trans-AT type I	16%
Region 13	lanthipeptide-class-i	1,676,556	1,701,597			
Region 14	NRPS-like	1,808,287	1,851,978			

Region 15	NRPS, siderophore	1,916,492	1,978,624	friulimicin A / friulimicin B / friulimicin C / friulimicin D	NRP	21%
Region 16	RRE-containing, NRPS, T1PKS	1,981,322	2,042,654	meilingmycin	Polyketide	2%
Region 17	terpene	2,150,766	2,171,477	geosmin	Terpene	100%
Region 18	RiPP-like	2,183,111	2,194,436			
Region 19	PKS-like, T1PKS, other	2,320,695	2,378,251	meilingmycin	Polyketide	5%
Region 20	siderophore, NRPS	2,550,673	2,598,163			
Region 21	terpene	3,159,315	3,179,207	julichrome Q3-3 / julichrome Q3-5	Polyketide	22%
Region 22	butyrolactone	4,004,072	4,015,004	thioviridamide	RiPP	10%
Region 23	T2PKS	4,642,100	4,714,611	spore pigment	Polyketide	83%
Region 24	siderophore	5,480,347	5,492,116	desferrioxamin B / desferrioxamine E	Other	83%
Region 25	melanin	5,564,462	5,575,058	melanin	Other	60%
Region 26	NRPS-like	5,828,286	5,871,217	SCO-2138	RiPP	64%
Region 27	ectoine	6,446,392	6,456,481	ectoine	Other	100%
Region 28	hglE-KS, T1PKS	7,049,867	7,101,670	cinnamycin	RiPP:Lanthipeptide	14%
Region 29	T3PKS	7,167,442	7,208,506	herboxidiene	Polyketide	7%
Region 30	NRPS, NAPAA, ectoine	7,212,385	7,324,494	stenothricin	NRP:Cyclic depsipeptide	27%
Region 31	T1PKS	7,342,702	7,389,296			
Region 32	NRPS, T1PKS	8,241,290	8,397,648	pentamycin	Polyketide	100%
Region 33	terpene	8,603,150	8,628,841	carotenoid	Terpene	63%

A.3.5.2: AntiSMASH predicted BGCs in the genome of *S. murinus*.

NZ_CP066774.1 (Streptomyces murinus)



Region	Type	From	To	Most similar known cluster		Similarity
Region 1	terpene	1,415,244	1,435,309	julichrome Q3-3 / julichrome Q3-5	Polyketide	25%
Region 2	terpene	1,743,048	1,762,753			
Region 3	NRPS, T1PKS, siderophore	2,033,372	2,112,355	kinamycin	Polyketide	22%
Region 4	PKS-like, NRPS-like, terpene	2,172,479	2,215,397	phenalinolactone A	Terpene + Saccharide:Hybrid/tailoring	94%
Region 5	other, T1PKS, PKS-like	2,293,820	2,350,506	meilingmycin	Polyketide	5%
Region 6	RiPP-like	2,399,516	2,410,341			
Region 7	terpene	2,420,993	2,442,439	geosmin	Terpene	100%
Region 8	T1PKS, NRPS, RRE-containing	2,553,201	2,613,349	meilingmycin	Polyketide	2%
Region 9	siderophore, NRPS	2,617,364	2,679,661	friulimicin A / friulimicin B / friulimicin C / friulimicin D	NRP	21%
Region 10	NRPS, NRPS-like, betalactone	2,939,557	3,029,950	kirromycin	NRP + Polyketide:Modular type I + Polyketide:Trans-AT type I	16%
Region 11	terpene, thiopeptide, LAP	3,127,522	3,180,293	hopene	Terpene	92%
Region 12	transAT-PKS-like, NRPS-like, T1PKS	3,336,566	3,386,364	cinnabaramide A	NRP + Polyketide:Modular type I	18%
Region 13	NRPS, lanthipeptide-class-i, lanthipeptide-class-ii	3,527,781	3,592,060	bleomycin	NRP:Glycopeptide + Polyketide:Modular type I + Saccharide:Hybrid/tailoring	12%
Region 14	NRPS	3,621,606	3,671,116	mirubactin	NRP	78%

Region 15	T1PKS	3,699,365	3,740,844	versipelostatin	Polyketide	5%
Region 16	RiPP-like	3,805,892	3,816,107	informatipeptin	RiPP:Lanthipeptide	28%
Region 17	betalactone	3,914,206	3,943,661	fusaricidin B	Polyketide + NRP:Lipopeptide	25%
Region 18	NRPS , T1PKS , NRPS-like	4,004,151	4,086,302	conglobatin	NRP	10%
Region 19	terpene	4,089,318	4,110,045	ebelactone	Polyketide	5%
Region 20	other , lanthipeptide-class-iv	4,159,254	4,197,035	A-503083 A / A-503083 B / A-503083 E / A-503083 F	NRP	7%
Region 21	T1PKS , NRPS	4,399,957	4,557,322	pentamycin	Polyketide	100%
Region 22	NRPS	4,591,579	4,634,440	diisonitrile antibiotic SF2768	NRP	66%
Region 23	T1PKS , NRPS-like	4,646,766	4,693,709	borrelidin	Polyketide:Modular type I	9%
Region 24	terpene	4,854,234	4,875,922			
Region 25	other , NRPS	5,330,762	5,398,215	actinomycin D	NRP	89%
Region 26	ectoine , NRPS , NAPAA	5,726,015	5,839,731	stenothricin	NRP:Cyclic depsipeptide	27%
Region 27	T3PKS	5,843,100	5,884,164	herboxidiene	Polyketide	7%
Region 28	T1PKS , hglE-KS	5,925,431	5,976,996	cinnamycin	RiPP:Lanthipeptide	14%
Region 29	ectoine	6,576,827	6,587,237	ectoine	Other	100%
Region 30	lanthipeptide-class-iii	7,280,477	7,303,167			
Region 31	melanin	7,502,940	7,513,536	melanin	Other	60%
Region 32	siderophore	7,593,123	7,604,892	desferrioxamin B / desferrioxamine E	Other	83%
Region 33	T2PKS , T1PKS	8,242,983	8,315,594	spore pigment	Polyketide	83%