The Function and Evolution of the Response Regulator CtrA in *Rhodobacter capsulatus* and Alphaproteobacteria

by Roshni S. Kollipara

A thesis submitted to the School of Graduate Studies

in partial fulfilment of the requirements for the degree of

Master of Science

Department of Biology

Memorial University of Newfoundland

November 2020

St. John's, Newfoundland and Labrador, Canada

Abstract

Rhodobacter capsulatus is a model organism for studying gene transfer agents (GTAs). GTAs are a unique facilitator of gene transfer in prokaryotes. The DNA binding response regulator CtrA plays a key role in modulating GTA activity in *R. capsulatus*, as well as flagellar biosynthesis and cell motility. CtrA is an OmpR/PhoB response regulator with an N-terminal receiver domain and a C-terminal transcriptional regulator domain. One unusual aspect of CtrA function in *R. capsulatus* is that it regulates gene expression in both the phosphorylated and non-phosphorylated forms. Using overlap extension PCR, the constructs for expression of three of different versions of *ctrA* in *R. capsulatus* were prepared: wild type, phosphomimetic, non-phosphorylatable. These constructs place the genes under the control of the *R. capsulatus puf* promoter for high level of expression and the encoded proteins have 6×-histidine tags for purification in studies aimed at determination of the DNA binding sites of the different versions of CtrA. Horizontal gene transfer is an interesting way that bacteria can increase their genetic diversity. In this work, the distribution of *ctrA* in the Alphaproteobacteria was examined and evidence of horizontal gene transfer of this gene was found. Using phylogenetic analyses, several instances of apparent misclassification of alphaproteobacteria to the wrong orders were found and one candidate *ctrA* horizontal gene transfer event that may have occurred in an ancestral bacterium that subsequently evolved into one lineage within the order Sphingomonadales was found.

Acknowledgements

I thank my lab members, especially Dr. Marta Canuti, for their support, guidance, and comradery. Of course, I thank my supervisor Dr. Andrew Lang for his encouragement, patience, and guidance. His support throughout my degree truly made a positive impact. This thesis is an accomplishment not only for me but also for my husband Evan, and I thank him for his relentless encouragement and belief.

Co-Authorship Statement

I am the primary author of all chapters in this thesis, for which I performed the associated laboratory work, data analysis, and bioinformatic analysis. I authored, in completeness, all text for the presented thesis, which then received editorial input from my supervisor Dr. Andrew Lang. In Chapter 3, I collected the sequence data and conducted the bioinformatic analysis. The results were interpreted by Dr. Lang, Dr. Marta Canuti, and me.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

16S rRNA	16S (Svedburg) ribosomal RNA
6×His tag	Six-histidine tagged protein
AA	Amino acid
AcOH	Glacial acetic aid
APS	Ammonium persulfate
bis	N,N'-Methylenebisacrylamide
BLAST	Basic local alignment search tool
BLASTn	Basic local alignment search tool for nucleotides
BLASTp	Basic local alignment search tool for proteins
bp	basepair
bs	bootstraps
ChIP-Seq	Chromatin immunoprecipitation sequencing
ClustalW	Computer program for multiple sequence alignment
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
EDTA	Ethylene diamine tetraacetic acid
FASTA	Fast-all (file containing nucleotide or protein sequence)
GTA	Gene transfer agent
HGT	Horizontal gene transfer
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kb	Kilo bases
kDa	kilodalton
LB	Luria Bertani medium

MeOH	Methanol
NCBI	National center for biotechnology information
Ni-NTA	Nickel-nitrilotriacetic acid resin
nr	Non-redundant
ORF	Open reading frame
ORI	Origin of replication
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
RBS	Ribosome binding site
RcGTA	*Rhodobacter capsulatus* gene transfer agent
RCV	*Rhodobacter capsulatus* medium five
RNA	Ribonucleic acid
rpm	Revolutions per minute
RpoB	RNA polymerase subunit-beta
s	seconds
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TAE	Tris-acetate EDTA buffer
TBST	Tris buffered saline-tween 20
TE	Tris-EDTA buffer
TEMED	N,N,N',N'-tetramethylethane-1,2-diamine
TrimAl	Tool for cleanup of multiple sequence alignment
Tris	tris(hydroxymethyl)aminomethane

TSS	Transcription start site
TTS	Transcription termination site
YPS	Yeast extract/peptone/salts medium

Chapter 1. Introduction and Overview

1.1 Alphaproteobacteria

Alphaproteobacteria is a class of Gram-negative bacteria within the phylum Proteobacteria. It contains ten orders and a large number of described species. They are found in most environments, some are clinically relevant, and many are agriculturally important (1). Several species form symbioses with eukaryotic organisms (2). The majority of phototrophic proteobacteria are in the class Alphaproteobacteria (3). Some notable members include *Rhodobacter capsulatus*, a model anoxygenic phototroph and the first organism in which gene transfer agents (GTAs) were discovered, *Caulobacter crescentus*, an organism continues to be thoroughly studied for its cell-cycle due to its interesting dual-cell life cycle, *Pelagibacter ubique*, possibly the most abundant bacterium in the ocean, and *Bartonella henselae*, the pathogen responsible for cat scratch fever.

1.2 *Rhodobacter capsulatus*

Rhodobacter capsulatus a purple nonsulfur alphaproteobacterium that is a facultative photosynthetic anaerobe. It is commonly found in aquatic environments such as sewage and waste water lagoons (4), and was first reported in 1907 by Molisch, renamed to *Rhodopseudomonas capsulata* by van Niel in 1944, and renamed again to its current name in 1984 (5, 6). Purple bacteria were originally named due to a slight purple pigmentation observed in the colonies of the first species discovered (1). Purple bacteria are divided into purple sulfur and purple nonsulfur bacteria as some species do not produce sulfur by-products (1). *R. capsulatus* is able to grow in both aerobic and anaerobic conditions, but it is worth noting that culture density is significantly higher under photosynthetic anaerobic conditions.

R. capsulatus was originally studied due to its very versatile metabolic capabilities as it can grow in five different situations: photoautotrophically, anaerobically with light as its energy source and CO_2 as the carbon source; photoheterotrophically, anaerobically with light as its energy source and using organic compounds as the carbon source; chemoorganotrophically, aerobically in the dark with organic compounds for energy and carbon sources; chemoautotrophically, aerobically in the dark with inorganic compounds for energy and CO_2 as the carbon source; and fermentatively, anaerobically in the dark with organic compounds as the energy and carbon sources (3, 7). In 1974 GTAs were discovered in *R. capsulatus*, introducing a new process of horizontal gene transfer (8).

1.3 Horizontal Gene Transfer

The primary transfer of genetic information occurs between a parent and an offspring, which can be considered "vertical" when looking at generational relatedness. Horizontal gene transfer (HGT) is when genetic material is transferred from one organism to another, where the organisms are not in a parent-offspring relationship (9). HGT is a fascinating process of genetic exchange that has mostly been observed among prokaryotes but there is also evidence of gene transfer between prokaryotes and eukaryotes (10). As detailed by Soucy and colleagues, HGT can help us understand phylogenies and evolution by conceptualizing a "web of life" instead of the traditional tree paradigm (9). An additional benefit to understanding HGT in prokaryotes is due to the importance of the transfer of genes that affect pathogenicity such as those that confer antibiotic resistance.

The method of HGT that most wet lab scientists are familiar with is transformation. Transformation is the uptake of external free DNA into a bacterial or archaeal cell (9), which can be forced in the lab with heat, electrical, or chemical shocks. Another important horizontal gene

transfer process that has been adopted by scientists is conjugation. A donor bacterium, most commonly *E. coli*, uses a pilus to draw itself and another bacterium together so that they have wall-to-wall contact, and DNA in the form of a plasmid (usually) is passed through a protein channel through the cell walls (11–13). Conjugation requires physical contact and therefore proximity is a major consideration in conjugation events. Interestingly, *Agrobacterium* spp. have been observed to transfer genetic material to plants via conjugation (14).

Two methods of HGT that are somewhat similar are transduction and DNA transfer by GTAs. Transduction can be understood as phage infection gone awry: when packaging DNA into assembling virions, a phage accidentally packages some of the host genome and transfers it to another cell. If the host DNA that is accidentally packaged is random, it is called generalized transduction, or if the host DNA that is packaged only originates from adjacent to the specific site where the phage was integrated as a prophage, it is called specialized transduction (9). GTAs are virus-like particles, believed to have evolved from prophages, which some prokaryotes produce (15). GTAs produced by *R. capsulatus* randomly package genomic DNA and transfer it to recipient cells, usually of the same species (16). A key point to distinguish between transduction and GTAs is that transduction is an artifact of a phage trying to propagate its own genome, whereas it seems some GTAs are biased against packaging the genes that encode their proteins (17).

1.4 Gene Transfer Agents

First observed in 1974 in the alphaproteobacterium *R. capsulatus*, GTAs are structures that are released from bacterial cells and facilitate horizontal gene transfer between cells of the same species (8, 16). Cross-species GTA activity has not yet been observed but the evidence of cross-species horizontal gene transfer suggests that it is not impossible (9). GTAs have been

observed in both prokaryotic domains, *Bacteria* and *Archaea* (15, 16). The structure of a GTA particle is similar to a tailed bacteriophage and the DNA is contained within the head (18). Similar to a virus or bacteriophage, DNA is packaged into the particles and then delivered to another bacterium after the GTAs are released from the host cell (8). Production of GTAs is culture density-dependent, as transcription of GTA-related genes is increased through a quorum-sensing system (19). There are no proven evolutionary benefits to GTAs, as a simulation showed that the conditions in which populations producing GTAs were more fit than populations that did not produce GTAs could only happen if the efficiency of GTA production and recombination were improbably high (20). Additionally, the cells that produce GTAs must die in order to release them, yet GTAs have been maintained in some lineages of bacteria for long evolutionary periods (21). Modelling studies have been conducted to determine the aspects of fitness related to GTAs that showed the cost of cell-lysis is difficult to counteract, and most simulations that showed an increase in fitness only in non-natural conditions such as strong synergistic epistasis (20).

GTAs are similar to bacteriophages in terms of morphology and activities, though there are some key differences. The major difference is that GTAs do not appear to act as selfish genetic elements (15). This is because they do not act towards the propagation of their own genome. Bacteriophages on the other hand replicate, package, and transfer their own DNA, and in some cases only accidentally package their host cell's DNA. Though GTAs have been detected in other Alphaproteobacteria such as *Dinoroseobacter shibae* and *Bartonella grahamii*, *R. capsulatus* is the model for GTA studies (22, 23). The GTA in *R. capsulatus* (RcGTA) has been shown to be biased against packaging the GTA genome (17). In addition, a single RcGTA particle is only capable of holding 4 kb, whereas its gene cluster is about 14 kb, and therefore it is not only a bias that prevents RcGTA from transferring its own genome, but also a packaging

limitation (15). Another significant difference between GTAs and bacteriophages is that GTAs are regulated by bacterial regulatory systems (24, 25).

1.5 Regulation of Gene Transfer Agents and the Response Regulator CtrA

In addition to regulation by quorum sensing, GTA gene transcription is regulated by the DNA binding protein CtrA (24). This is a response regulator protein found to be widely conserved within alphaproteobacteria (26). A member of the OmpR/PhoB family of response regulators, CtrA has an N-terminal receiver domain and a C-terminal transcriptional regulator domain (27). CtrA has been thoroughly studied in the Caulobacterales *Caulobacter crescentus*, also known as *C. vibrioides* (28). In *C. crescentus*, CtrA is part of a multi-component histidyl-aspartyl-phosphorelay system (29, 30) that appears to be unique to alphaproteobacteria (31). In the CtrA phosphorelay, an unknown signal causes the membrane-bound histidine kinase CckA to autophosphorylate, with subsequent phosphorylation of the phosphotransferase ChpT, which then transfers the phosphate to an aspartic acid residue in CtrA (D51) (32, 33). This phosphorelay exists in *R. capsulatus*, though the role of CtrA appears to be somewhat different than found for *C. crescentus* (32, 34).

The orders Caulobacterales and Rhodobacterales have a different scheme for cell cycle regulation (31), and interestingly these two orders also differ in the roles of their CtrA. In *C. crescentus*, CtrA is a cell cycle regulator and is necessary for survival (35), with the essential roles of coordinating DNA replication, cell division, and polar morphogenesis (31, 35). It has been observed in *C. crescentus* that phosphorylated CtrA autoregulates itself (36).

Although the CtrA proteins in the different species have some different functions, the model for the regulation of CtrA in *C. crescentus* can be extended to *R. capsulatus* (37). Unexpectedly, CtrA does not have as essential a role in *R. capsulatus* though it has been shown

to directly and indirectly regulate over 200 genes (38), including genes involved in incorporation of DNA from GTAs by recipient cells (39), flagellar motility (24), and GTA production (24). While it is the phosphorylated form of CtrA induces gene transcription in *C. crescentus* (40), there is evidence that the phosphorylated and unphosphorylated forms of CtrA affect the transcription of different genes in *R. capsulatus* (32) (Figure 1.1). This is an interesting observation because it is generally thought that only the phosphorylated form of a DNA-binding response regulator is active for transcriptional regulation (33), though there are some exceptions that have been reported in the literature (41).

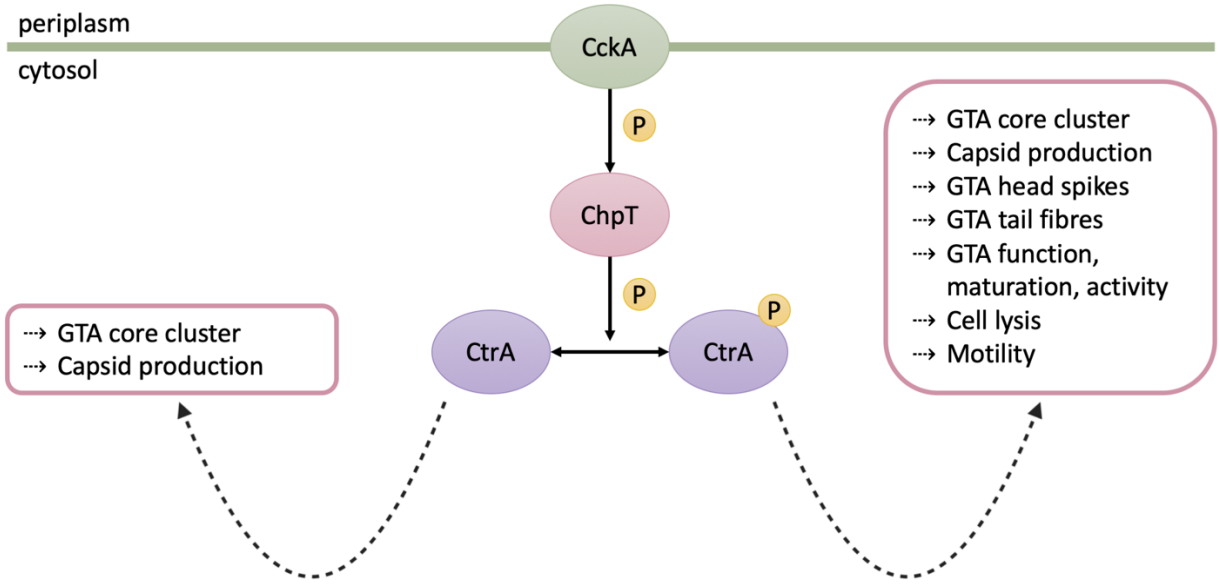


Figure 1.1 Representation of the histidyl-aspartyl phosphorelay involving CtrA and the roles that the phosphorylated and non-phosphorylated versions of CtrA in *R. capsulatus*.

In the *R. capsulatus* GTA system, GTAs are regulated by the producing cells (9, 16). In *R. capsulatus* it has been shown that GTA gene expression is regulated by quorum sensing via acyl-homoserine lactones, a partner-switching phosphorelay, and the CckA-ChpT-CtrA histidyl-

aspartyl phosphorelay (16) and there is evidence to suggest that these aspects of GTA function are also regulated within additional species such as *Dinoroseobacter shibae* (22, 42). Production of GTAs in *R. capsulatus* is highest when a culture is stationary phase, wherein a subpopulation of cells (<3%) produce GTAs and the remaining ~97% act as recipients (43). The capability of cells to receive DNA from GTAs is regulated by some of the same systems that control GTA production (43). The CckA-ChpT-CtrA phosphorelay is very important for GTA activity in *R. capsulatus*, and loss of the different components of this regulatory system leads to major disruptions in GTA production (32) (Figure 1.1).

1.6 Conservation of CtrA

CtrA is a highly conserved response regulator in Alphaproteobacteria. Phylogenetic studies as of 2012 found no evidence of horizontal gene transfer of *ctrA* homologs within the Alphaproteobacteria (44). This is surprising as *ctrA* is not an essential gene in many alphaproteobacteria and non-essential genes are more likely to be associated with horizontal gene transfer events as compared to essential genes (9). Interestingly, it seems that CtrA was involved in regulation of motility in an ancestral alphaproteobacterium but not for essential functions (44). There is evidence that a divergence event occurred wherein the CtrA in one lineage of alphaproteobacteria (containing the orders Rhizobiales and Caulobacterales) acquired an essential role (44). This idea is somewhat supported by comparative genome analysis that found two main schemes of cell cycle regulation: one that was shared between Caulobacterales and Rhizobiales, and a simpler scheme found in the order Rhodobacterales (31).

CtrA has been extensively studied in *C. crescentus*, where it has the essential roles of coordinating DNA replication, cell division, and polar morphogenesis (31). The two alphaproteobacteria *Brucella abortus* (a Rhizobiales) and *C. crescentus* (a Caulobacterales) have

some conserved targets of CtrA (45), and the *B. abortus* phosphorylated CtrA autoregulates itself, similar to the *C. crescentus* CtrA (36). *Magnetospirillum magneticum* strain AMB-1 has a CtrA that is not essential but is required for motility, similar to *R. capsulatus* (44).

1.7 Thesis Overview

The details of how GTA production in *R. capsulatus* is regulated are becoming clearer but there is still much to discover. The response regulator CtrA has many interesting roles and understanding its behaviour can yield interesting insights into not only GTA regulation, but also bacterial regulatory systems as a whole. Understanding the prevalence of horizontal gene transfer of *ctrA* can lead to interesting information about the importance and evolution of *ctrA*, as well as the horizontal gene transfer as a whole.

In this thesis I present steps taken to create *ctrA* constructs for further investigation of CtrA binding sites and I identify horizontal gene transfer of *ctrA* within Alphaproteobacteria. The goal of my thesis work was to increase the overall understanding of *ctrA* within *R. capsulatus* and to evaluate the presence or absence of horizontal gene transfer events for this gene. In Chapter 2, I detail the creation of constructs of *ctrA* that can be expressed in *R. capsulatus* with the photosynthetic promoter *pufP* and with a 6×-histidine tag for protein detection and purification steps. In addition, I created two mutant versions of *ctrA*: a phosphomimetic version and a non-phosphorylatable version in order to study the activity of both phosphorylated and non-phosphorylated CtrA. In Chapter 3, I used phylogenetic analysis to discover horizontal gene transfer events of *ctrA* within Alphaproteobacteria. I found evidence of horizontal gene transfer between an ancestral bacterium of *Sandarakinorhabdus* and *Polymorphobacter* from a member of Rhodospirillales, as well as evidence of possible misclassification of several bacteria.

Chapter 2: The Response Regulator CtrA in *Rhodobacter capsulatus*

2.1 Abstract

Gene transfer agents (GTAs) are fascinating mediators of horizontal gene transfer that are produced by the bacterium *Rhodobacter capsulatus*. The DNA binding response regulator CtrA is responsible for regulating GTA activity in *R. capsulatus*. There are still gaps in understanding how CtrA is regulated and how it regulates GTA activity. Both phosphorylated and non-phosphorylated CtrA have important roles in GTA production in *R. capsulatus*. A loss in phosphorylation of CtrA leads to increased transcription of the GTA cluster, however CtrA phosphorylation is essential to the production of some GTA structural proteins. To further study CtrA and how its phosphorylation affects regulation of different genes, we aimed to study the DNA binding activity of CtrA. Mutants of *ctrA*, encoding proteins that act as the nonphosphorylated and phosphorylated forms, were prepared by site-directed mutagenesis. Using overlap extension PCR, three versions of *ctrA* were prepared in an expression vector for *R. capsulatus*, each with a 6×His tag: wild type *ctrA*, a version encoding a D51E substitution mutant (phosphomimetic), and a version encoding a D51A substitution mutant (non-phosphorylatable). These constructs will be useful for downstream investigations using ChIP-Seq.

2.2 Introduction

2.2.1 The Response Regulator CtrA

CtrA is an interesting DNA-binding response regulator found in many Alphaproteobacteria. In the purple non-sulfur bacterium *R. capsulatus*, CtrA has a non-essential role as a regulator of motility and GTA production. CtrA activity is regulated through a phosphorelay involving the histidine kinase CckA and the phosphotransferase ChpT. Both

phosphorylated and non-phosphorylated CtrA show evidence of regulating the expression of GTA related genes. Understanding the important role of phosphorylation in CtrA regulation can help us understand the greater system of GTA expression.

2.2.2 Phosphorylation of CtrA

In *R. capsulatus*, the non-phosphorylated form of CtrA promotes capsid protein production but is not sufficient for full RcGTA activity. This was shown when capsid protein production was rescued by the introduction of *ctrA-D51A* into a *ctrA* null mutant strain but this was not accompanied by a return of normal levels of gene transfer activity (32). Introduction of a gene encoding the phosphomimetic version, *ctrA-D51E*, into the *ctrA* null mutant strain restored both capsid protein production and gene transfer functionality. Therefore, the non-phosphorylated form of the protein can stimulate transcription of the RcGTA structural gene cluster but the phosphorylated form is required to produce fully mature and functional GTA particles and for their release from cells (32, 34).

Unexpectedly, it seems that CtrA in *R. capsulatus* contributes to regulation in both phosphorylated and non-phosphorylated forms. In *C. crescentus*, CtrA has the vital roles of coordinating DNA replication, cell division, and polar morphogenesis (31). In *R. capsulatus*, phosphorylation of CtrA is essential for motility, but both phosphorylated and non-phosphorylated CtrA induces expression of the RcGTA major structural gene cluster (32). The head spikes found on mature RcGTA particles are made of the proteins GhsA and GhsB (37, 46), and phosphorylation of CtrA is important for production of these head spike proteins (37). Phosphorylation of CtrA is necessary for capsid release through cell lysis, as this function was rescued by introduction of a phosphomimetic CtrA to a mutant that lacked *cckA*, the histidine kinase upstream of CtrA in the phosphorelay (32) and cell lysis was increased by the phosphomimetic CtrA (37).

There is still much to discover regarding how CtrA regulates the different aspects of RcGTA production. The aim of this work was to prepare different versions of the *R. capsulatus ctrA* that can be used to determine DNA binding sites of phosphorylated and unphosphorylated versions of the CtrA protein. This will allow a genome-wide evaluation of the activities of the two versions of the protein whereas previous studies have only evaluated specific genes of interest.

2.3 Materials and Methods

2.3.1 Bacterial Strains, Plasmids, and Culturing

R. capsulatus liquid cultures were grown at 35 °C with shaking at 200-250 rpm for 16-18 hours in RCV medium (47). *R. capsulatus* strains were maintained short-term on RCV agar and supplemented with appropriate antibiotics as needed (Table 1). *Escherichia coli* liquid cultures were grown at 37 °C with shaking at 200-250 rpm for 16-18 hours in LB medium. *E. coli* strains were maintained short-term on LB agar and supplemented with appropriate antibiotics (Table 1). Long-term storage of strains was achieved through use of freezer stocks (-80 °C), prepared by resuspending 8 mL of centrifuged (8,000 rpm, 10 min, 10 °C) culture in 1 mL 20% glycerol in LB medium for *E. coli* and RCV medium for *R. capsulatus*. All strains and plasmids used in this study are detailed in Table 2.

Table 1. Concentrations of antibiotics used throughout this study.

Antibiotic	Stock concentration (mg/mL, in water)	Working concentration (μg/mL)	
		E. coli	*R. capsulatus*
Ampicillin	100	100	Not used
Kanamycin	10	50	10
Gentamicin	100	10 or 50*	3
Tetracycline	100 or 10	10	0.5

*some strains required higher concentrations in order to ensure maintenance of a plasmid

Table 2. Strains and plasmids used in this study.

Strains and plasmids	Description	Reference or source
***R. capsulatus* strains**		
SB1003	Genome-sequenced strain	(48, 49)
SBRM1	SB1003 with disruption in *ctrA* gene	(38)
SB1003 (pRRR4:*ctrA*)	SB1003 with additional copy of *ctrA* in plasmid in pRR4	This study
***E. coli* strains**		
BL21(DE3)	Strain for expression of 6×His tagged proteins	New England Biolabs
S17-1	Plasmid-mobilizing strain	(50)
NEB5α	Used for cloning	New England Biolabs
Plasmids		
pRRR4:*ctrA*	Expression of CtrA driven by the *puf* promoter	
pET28a:*ctrA*	Expression of CtrA with 6×His tag driven by T7 promoter	Mercer and Lang, unpublished
pET28a:D51A *ctrA*	Expression of non-phosphorylatable *ctrA* with 6×His tag driven by T7 promoter	This study
pET28a:D51E *ctrA*	Expression of phosphomimetic CtrA with 6×His tag driven by T7 promoter	This study
pGEM-T Easy	TA PCR product cloning vector	Promega
pGEM: WT *ctrA*	Indirect clothing vector for wild-type CtrA	This study
pGEM: D51A *ctrA*	Indirect clothing vector for non-phosphorylatable CtrA	This study
pGEM: D51E *ctrA*	Indirect clothing vector for phosphomimetic CtrA	This study
pRR5C	Expression of genes in *R. capsulatus* driven by *puf* promoter	(51)
pRR5C:D51A *ctrA*	Expression of non-phosphorylatable CtrA in *R. capsulatus* driven by *puf* promoter	This study

pCM62:tetC	Broad host range vector; expression of genes in *E. coli* driven by *lac* promoter, tetracycline resistance	(52)
pCM62:*pufP*	*puf* promoter inserted into pCM62:tetC for expression of genes in *R. capsulatus*	Gift from J.T. Beatty; unpublished
pCM62:tetC:WT *ctrA*	Expression of CtrA in *R. capsulatus* driven by *puf* promoter	This study
pCM62:tetC:D51A *ctrA*	Expression of non-phosphorylatable CtrA in *R. capsulatus* driven by *puf* promoter	This study
pCM62:tetC:D51E *ctrA*	Expression of phosphomimetic CtrA in *R. capsulatus* driven by *puf* promoter	This study

2.3.2 Bacterial Conjugations

Using bacterial conjugation (12), the plasmid pRRR4:*ctrA* was inserted into the laboratory wild-type *R. capsulatus* strain SB1003, thereby creating a new strain which overproduces CtrA. This new strain, SB1003 (pRRR4:*ctrA*), was subjected to Western blot visualization.

2.3.3 Protein Induction and Purification

Expression of CtrA that had an N-terminus 6×His tag was induced in *E. coli* strain BL21(DE3). A 10-mL culture at stationary phase was added to 200 mL of LB broth with kanamycin, and IPTG was added to a final concentration of 1 mM after one hour. The culture was grown for 4 hours and then the cells were pelleted by centrifugation at 5000 × *g* for 10 min at 10 °C. The cells were then lysed and the protein isolated with Ni-NTA agarose (Qiagen) gravity-flow chromatography, as per the manufacturer's instructions.

2.3.4 Antibody Testing and Western Blotting

Antibodies for the immunodetection of CtrA were produced by Pacific Immunology Corp. (California). White rabbits were injected four times with the peptide sequence

CHAIIRRSKGHSQSIIRTGK (CtrA residues 113-131) and the antibodies were purified using affinity purification. This sequence was chosen because it was predicted to be immunogenic and is located on the outside of the folded CtrA protein.

Evaluation of antibody efficiency and specificity was conducted using Western blots. Both *R. capsulatus* and *E. coli* bacterial cultures were grown and once they reached the log phase of their growth curve 500 µL of the culture was centrifuged at 12,000 rpm for one minute at room temperature. The cell pellet was resuspended in 500 µL TE buffer and cells were mixed (2:1) with 3X SDS-PAGE Sample Loading Buffer (NEB). All samples and a protein molecular weight ladder were boiled at 95-100 °C for 5 minutes prior to loading on the gel. A 12% separating gel was made by combining 938 µL 1.5 M Tris pH 8.8, 35 µL 10% w/v SDS, 35 µL 10% w/v ammonium persulfate, 1400 µL 30% acrylamide-bis (29:1) with 1090.25 µL sterile distilled water and cross-linked with 1.75 µL of the polymerising agent TEMED for 20-30 minutes. A 5% stacking gel was made by combining 180 µL 0.5 M Tris pH 6.8, 15 µL 10% w/v SDS, 15 µL 10% w/v APS, 250 µL 30% acrylamide-bis (29:1) with 1038.2 µL water and then cross-linking with 1.8 µL TEMED for 15 minutes. Electrophoresis was performed using an SDS-PAGE chamber (CBS Scientific) at 110 V for 80 minutes at room temperature in SDS-PAGE running buffer (25 mM tris base, 0.2 M glycine, and 3.5 mM SDS). Transfer of proteins from the gel to a nitrocellulose membrane was completed using the BioRad Mini Trans-Blot cell in transfer buffer (38 mM glycine, 50 mM tris base, 20% methanol), performed at 70 V for 70 minutes. The cell was chilled using ice packs during transfer. The membrane subsequently underwent three washes with TBST (150 mM NaCl, 20 mM tris base, 450 uM tween-20), each 15 minutes long. Approximately 7 mL of blocking buffer (10% w/v non-fat milk powder in TBST) was incubated with the membrane for 1 hour. Antibodies against CtrA were diluted in

TBST and incubated with the membrane for 1 hour with gentle rocking using a 2D shaker. The membrane was washed with TBST (3 x 15 minutes) and the secondary antibody, peroxidase-conjugated anti-rabbit IgG (Santa Cruz Biotechnology), was added (diluted 1:10000 in TBST). After one hour with gentle rocking, the membrane was subsequently washed three more times with TBST before being developed using the WestPico chemiluminescence kit (Thermo Fisher Scientific) and visualized with a GE ImageQuant LAS 4000 Digital Imaging System (GE Healthcare). The resulting images were analyzed using the ImageQuant TL software.

Various parameters were altered in order to optimize CtrA binding by the antibodies, including the concentrations of antibodies against CtrA (1:100, 1:500, 1:1000, 1:5000, 1:10000 in TBST), acrylamide/bis-acrylamide concentration of the gel, culture density of the samples, culture conditions, and amount of cells used for the blots. Immunoprecipitation of the cell samples using the Dynabeads™ Protein-A Immunoprecipitation Kit (Invitrogen) was also attempted as a pre-concentration step in some tests.

2.3.5 Site Directed Mutagenesis

Using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies Inc.), phosphomimetic and non-phosphorylatable mutants of CtrA were created using primers listed in Appendix A.1. The protocol was followed exactly as written by the manufacturer. Constructs were verified by Sanger sequencing at The Centre for Applied Genomics (TCAG) DNA Sequencing Facility in Toronto, ON. Sequences were analysed using Geneious R11 (Biomatters).

2.3.6 Ligation-Based Cloning

Cloning was used to create plasmids that had both an *R. capsulatus* promoter and *ctrA* so that, ultimately, strains of *R. capsulatus* that produce 6×His tagged CtrA mutants could be used for ChIP-Seq experiments. Direct cloning was attempted using the restriction enzymes EcoRI

and SalI, and then attempted with only EcoRI. The plasmid containing *R. capsulatus* gene *ctrA*, pET28a:*ctrA* was isolated from NEB 5-alpha using the GeneJet Miniprep kit (Fisher Scientific). The *R. capsulatus ctrA* gene was excised out of pET28a using EcoRI, and then ligated with the plasmid pRR5C which was also cut with EcoRI. This method yielded low concentrations of the gene fragment, leading me to adopt an indirect cloning method. PCR was used to produce the *ctrA* fragment using pET28a:*ctrA* as the template (Phusion polymerase with High GC Buffer, 5% v/v DMSO), under the following conditions: 98°C 30 s, (98°C 7 s, 50-60°C 20 s, 72 °C 30 s) × 35, 72 °C 7 min. The PCR product was then A-tailed using Taq polymerase and ligated to pGEM T-Easy as per the manufacturer's instructions (Promega), cut with EcoRI and ligated with pRR5C or pCM62:*pufP*, which were also cut with EcoRI. The various concentrations of insert and vector that were tested, as well as the time and temperatures of ligations are detailed in Table 3.

Table 3. Ligation reaction conditions.

	pRR5C			pCM62:pufP
	SalI & EcoRI overhang	EcoRI overhang	T-A overhang	EcoRI overhang
Molar ratios	1:3, 1:5, 1:8	1:1, 1:3, 1:5, 1:7, 1:8, 1:10	1:3, 1:12	1:3, 1:8, 1:12
Temperatures and times	16C 1h – 4C O/N 4C O/N 16C O/N 25C 1h	16C O/N 4C O/N 25C 15 min 25C 1h 25C O/N 100C 30s – 20C O/N	25C 1h – 4C O/N	16C O/N 25C 15 min 25C 1h 25C O/N
Mass of vector	~33 ng	35 ng, 50 ng, 70 ng, 100 ng	30 ng, 50 ng	10ng, 35ng, 50ng, 70 ng, 170ng

2.3.6 Overlap PCR Extension Cloning

The method described by Bryskin and Matsumura in 2013 was used to design primers for overlap PCR extension (53). Two sequential PCRs were completed to form a final product of pCM62:pufP:ctrA. Primers were designed based on the *ctrA* insert sequence, with a 25-40 bp tail complimentary to the recipient vector pCM62:*pufP*. The primers were used with DreamTaq Green PCR Master Mix (Fisher Scientific) to amplify *ctrA* using pGEM:*ctrA* as the template using the PCR conditions: 98 °C 30 s, (98 °C 7 s, 60 °C 15 s, 72 °C 30 s) × 35, 72 °C 7 min. The PCR product was cleaned with AMPure XP magnetic beads (Beckman-Coulter) and eluted in a 1/20 equivalent volume of water. This clean PCR product was then incubated with T4 DNA polymerase (New England Biolabs Ltd.) for 15 minutes at 12 °C in order to remove the A overhangs left by Taq polymerase. After a second cleaning with AMPure XP and elution with a 1/4 equivalent volume of water, the product was used as "megaprimers" in the second PCR, where pCM62:pufP was the template. In a 30-µL reaction, 6 µL 5× HF Buffer, 0.6 µL 10 mM dNTPs, 0.9 µL DMSO, 250 ng megaprimers, 3.25 ng template, and 0.6 µL of Phusion were added and subjected to the following PCR conditions: 98 °C 1 min, (98 °C 7 s, 60 °C 20 s, 72 °C 1 min 20 s) × 25, 72 °C 7 min. A user-friendly protocol can be found in the appendix (Appendix A.3).

The reactions were visualized using gel electrophoresis (0.8% w/v agarose in TAE, v/v 0.005% GelGreen) and those that showed a final product at the expected size were treated with the enzyme DpnI for 1 hour at 37 °C to remove the template DNA. Two microliters of the reaction were then transformed into NEB 5-alpha as per the supplier's instructions and plated on LB agar with tetracycline. After 18-24 hours, colonies were picked and screened using colony PCR. A colony was resuspended in 50 µL of LB media, and 2.5 µL added to a 30-µL PCR with

DreamTaq Green with the following reaction conditions: 98 °C 30 s, (98 °C 7 s, 60 °C 15 s, 72 °C 30 s) × 35, 72 °C 7 min. After confirmation using gel electrophoresis, positive colonies were used to inoculate 10 mL of LB broth with tetracycline. Plasmids were isolated and sequenced at The Centre for Applied Genomics (TCAG) DNA Sequencing Facility in Toronto, ON. Sequences were analysed using Geneious R11 (Biomatters).

2.4 Results

2.4.1 Antibody Testing and Western Blotting

The antibodies designed against CtrA amino acid residues 113-131 were assayed through western blotting with samples from *R. capsulatus* cultures grown under a number of conditions. Different concentrations of primary antibody were used to detect a band with the expected molecular weight of CtrA (26.74 kDa). The *R. capsulatus* strain SBRM1 has a *ctrA* knockout and therefore was used as a negative control (Figure 2.1).

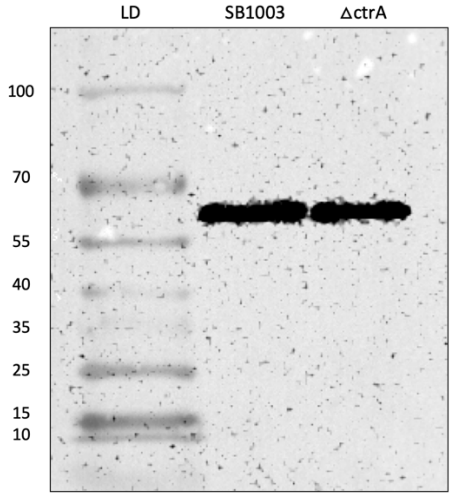


Figure 2.1 Western blot of *R. capsulatus* strains SB1003 and Δ*ctrA*. Ladder (LD) range is 10-100 kDa.

Despite varying the concentrations of primary antibody used, a band at the expected size of CtrA was not visualized in the *R. capsulatus* cells. This led to a number of optimization steps: increasing the volume of cell lysate loaded on the gel, longer culturing time, and adjustments to the visualization protocol. None of these steps yielded visualization of the expected band so a CtrA-overproducing strain was created by conjugation of pRRR4:*ctrA* from *E. coli* S17-1 into *R. capsulatus* SB1003 (Figure 2.2). As the positive control, 6×His tagged CtrA was induced and purified from *E. coli* and was expected to be visualized at a slightly greater size than the native CtrA.

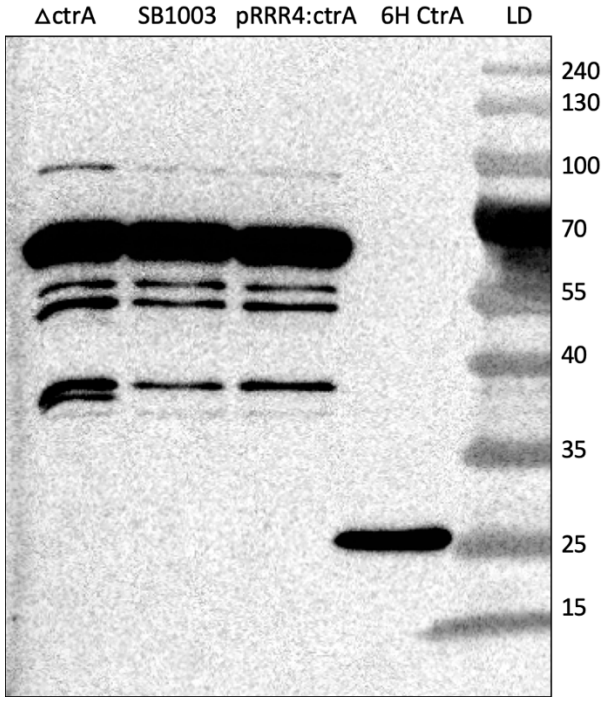


Figure 2.2 Western blot of CtrA over-expression strain pRRR4:*ctrA* as compared to SB1003, Δ*ctrA* and the 6×His tagged isolated CtrA (6H CtrA). Ladder (LD) range is 15-240 kDa.

Even with what should have been an overproduction of CtrA, visualization of a band at the expected size of approximately 26 kDa was difficult to detect. By increasing the exposure time when taking an image, I was able to see a faint band at the expected size in the sample of the culture of SB1003 (pRRR4:*ctrA*) (Figure 2.3), however the sensitivity was clearly very poor.

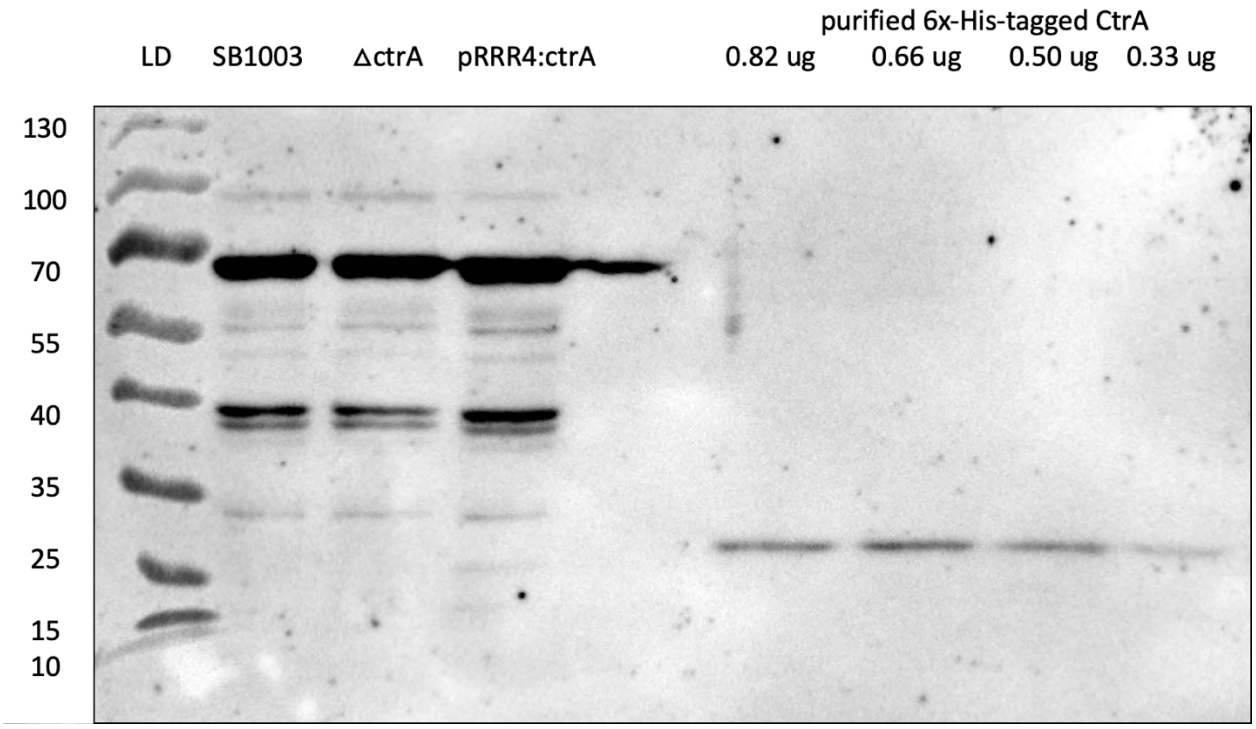


Figure 2.3 Western blot of wild-type *R. capsulatus* strain SB1003, *ctrA* knockout strain Δ*ctrA*, CtrA over-expression strain pRRR4:*ctrA*, and dilutions of purified 6×His tagged CtrA. Ladder (LD) range is 15-240 kDa.

Although CtrA was visualized in the overproducing strain, further consideration of the planned use of these antibodies for ChIP-Seq indicated that the non-specific binding exhibited was too high to be considered for these experiments.

2.4.2 Cloning

With the goal of ChIP-Seq experiments in mind, we opted to use a commonly available antibody against a 6×His tag. To do this, I needed strains of *R. capsulatus* that expressed a 6×His tag version of CtrA. In addition, in order to study how phosphorylation affects the role of CtrA, point mutations were introduced to produce two mutants of CtrA that mimic different phosphorylation states. Using SDM mutants of *R. capsulatus ctrA* in pET28a as well as the wild-type *ctrA* in pET28a, direct and indirect directional cloning was attempted using restriction enzymes. Out of the numerous ligation reactions conducted, the only one to work was with 6.9 ng of vector (pRR5C), 36 ng of insert (*ctrA* D51A), in a 25 µL reaction, subjected to a 1-hour incubation at 16 °C, followed by incubation at 4 °C overnight. Subsequent reactions under the same conditions (n=3) did not yield a positive cloning result. After exhausting possible variables to adjust with no success, I decided to opt for a different approach.

Overlap extension PCR uses two consecutive PCRs to amplify a vector with the desired insert as the primer. When using the procedure outlined in the paper describing this method (53), I found successful PCR conditions were difficult to reproduce. Due to irreproducibility of the reaction I wanted to optimize the protocol and so I conducted a reaction volume study (Table 4). Other parameters tested were annealing temperature through a gradient method and the DNA polymerase used.

Table 4. Optimizing PCR conditions for Part A of overlap extension PCR cloning.

		Annealing Temperature (°C)											
		57	57.9	58.8	59.7	60.6	61.5	62.5	63.4	64.3	65.2	66.1	67
Reaction Volume (µL)	10	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++
	20	++	+++	++	++	++	++	++	++	++	++	++	+++
	20†	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++
	30	+	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++
	40	+	++	++	++	++	++	++	++	++	+++	+++	+++
	50	+	++	++	+	++	++	++	++	++	+++	+++	+++

(+) Product detected, (++) product clearly visible, (+++) optimal conditions, (-) no product detected, (†) DreamTaqGreen used as polymerase instead of Phusion HF.

The suggested reaction volume in the literature was 50 µL, however my results showed that smaller reaction volumes yield more concentrated products. The 10-µL and 30-µL reaction volumes had the greatest yield, though the intensity of the PCR products on the agarose gels were greater in the 30-µL reactions. In addition, a reaction volume of 10 µL would not provide a sufficient amount of megaprimer to continue to Part B of the overlap extension method. It was also shown that DreamTaq Green performed better than Phusion HF in the same reaction volume, but primer dimers formed at lower annealing temperatures. For ease and reduction of user error, DreamTaq Green was chosen for Part A of the overlap extension. Though a reaction volume of only 20 µL was considered for DreamTaq Green, 30-µL reactions were conducted so as to produce ample material for the following steps. This PCR method was relatively consistent across annealing temperatures as compared to reaction volumes. Every annealing temperature produced some product, and in optimal reaction volumes all annealing temperatures yielded equal results. All subsequent Part A reactions were conducted with an annealing temperature of 60 °C.

One issue with using DreamTaq Green is that it leaves a 3’ overhanging A on the PCR product which could affect the binding of the megaprimers to the plasmid template in part B and

the sequence of the final produced plasmid. In order to overcome this drawback, the concentrated and cleaned product from part A was incubated with T4 DNA polymerase to remove the A overhang.

Part B of overlap extension requires an unfavourable reaction with megaprimers that were 881 bp long. The exonuclease activity of Phusion is essential to Part B of overlap extension, therefore only Phusion and not DreamTaq Green was used in the optimization steps for Part B. The masses of vector and megaprimers in the reaction had to be optimized, as well as the amount of Phusion polymerase (Table 5).

Table 5. Optimizing PCR components for Part B of overlap extension PCR cloning.

	Mass of megaprimers (ng)					
		10	50	100	250	
Mass of vector (ng)	0.1	-	-	-	-	0.15
	1	-	-	+	++	0.15
	3.25	-	-	+	++	0.15
	10	+	+	++	++	0.15
	3.25	-	-	+	++	0.225
	3.25	+	+	++	+++	0.30
	3.25	+	++	+++	+++	0.90

Volume of Phusion (µL)

(+) Product detected, (++) product clearly visible, (+++) optimal conditions, (-) no product detected.

Eighty-eight clones in total were sequenced with Sanger sequencing. Many of the clones had point mutations, but I obtained at least one clone for each version of CtrA (wild-type, phosphomimetic, and non-phosphorylatable) that did not have any mutations resulting in amino acid substitutions (discussed further below). Some of these undesired mutations probably arose from the use of DreamTaq Green to amplify the original fragment in part A of the overlap extension PCR because the enzyme is non-proofreading. The Phusion enzyme used in part B of the overlap extension has high fidelity.

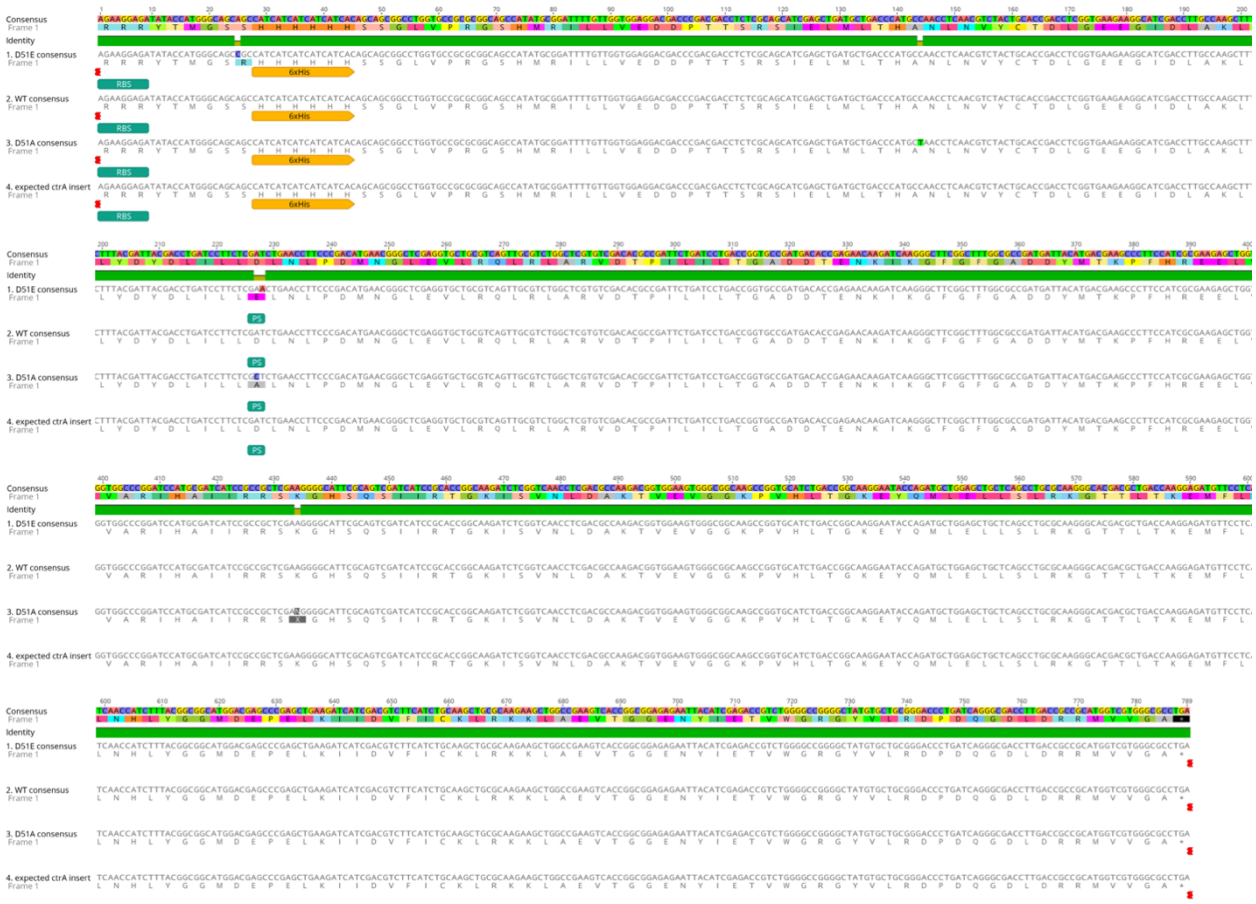


Figure 2.4 Nucleotide alignment of *ctrA* constructs made using overlap extension PCR. "D51E consensus" is the sequence for the phosphomimetic *ctrA* construct. "WT consensus" is the sequence for the wild-type *ctrA* construct. "D51A consensus" is the sequence for the non-phosphorylatable *ctrA* construct. Finally, "expected ctrA insert" is the expected wild type sequence after overlap extension PCR.

The wild-type *ctrA* construct was created without any point mutations (Figure 2.4). The phosphomimetic *ctrA* construct, D51E, has a point mutation before the 6×-Histidine tag leading to a codon change from AGC to CGC and the adenine at position 25 was replaced with a cytosine (Figure 2.4), and an amino acid change from serine to arginine. Though this substitution is before the 6×-histidine tag and is not in the part of the sequence that codes for CtrA, this

construct may exhibit a different protein behaviour and therefore needs to be further studied or cannot be used. The non-phosphorylatable *ctrA* construct, D51A, has a silent point mutation within the coding region for *ctrA* where the cytosine at position 144 in Figure 2.4 was replaced with a thymine, leading to a codon change from GCC to GCT. Fortunately, both codons code for the amino acid alanine. The D51A mutant also has one base noted as "N" that was not resolved in the sequencing data (position 434 in Figure 2.4). Unfortunately, due to the pandemic the sample was not able to be sequenced again before this thesis was submitted.

2.5 Discussion

The original experimental plan for this chapter was to determine the DNA binding sites of CtrA in *R. capsulatus*, using a wild-type version, a phosphomimetic version, and a non-phosphorylatable version. The proposed methodology for determining the DNA binding sites was ChIP-Seq. As a culture of bacteria grows, the DNA and DNA binding proteins are naturally "incubated together". For this method a DNA-protein complex is made by cross-linking the two biomolecules with formaldehyde. The DNA is sheared to form small fragments of 300-500 bp, some of which are bound to the protein of interest. Using an antibody specific for the protein, one can precipitate the protein and the DNA fragment to which it is bound along with it. The protein can subsequently be detached from the DNA using heat, and the remaining DNA can be sequenced using a next generation sequencing technology. The resulting sequences represent the fragments of DNA that the protein was bound to. These sequences can be aligned with the genome of the organism in order to find the locations of DNA binding sites (54–56).

The unexpected activity observed in non-phosphorylated CtrA requires more investigation to understand how CtrA regulates GTA production and activity. The aspartic acid residue at position 51 in the *R. capsulatus* CtrA is the phosphorylation site for the protein. I

created a point mutation in which the aspartic acid was substituted for glutamic acid, thereby creating a phosphomimetic protein (57, 58). Likewise, I effected a point mutation that caused the aspartic acid to become an alanine residue, ensuring the protein cannot be phosphorylated at this position (27). Using ChIP-Seq, we can determine the DNA-binding sites of CtrA and compare the sites of the phosphorylated and non-phosphorylated versions. This can lead to insight into the regulatory roles of CtrA.

The immunoprecipitation aspect of ChIP-Seq was expected to be completed with the antibodies created against the peptide sequence from CtrA (residues 113-131) , which was chosen because it was predicted to be immunogenic and located on the outside of the folded protein. This sequence was also compared to all *R. capsulatus* proteins and observed to have small regions of similarity with other proteins such as a branched-chain amino acid ABC transporter permease and a putative insertion element, with 7 and 6 amino acids conserved respectively. These small conserved regions were not expected to impact immunoprecipitation, however, these antibodies performed poorly and demonstrated nonspecific binding through Western blotting. By comparing BLASTp results of the immunogenic peptide to Western blots (Figures 2.2 and 2.3), it is possible the strong species at approximately 60 kDa is encoded by the genes *rcc01774* and *rcc02621*, as their estimated sizes are 59.5 and 60.5 kDa respectively. Additional proteins identified on the BLASTp result as having 5 and 6 amino acid matches to the peptide are estimated to be around 40 kDa and therefore could be responsible for the bands seen around 39 kDa on the Westerns. If any of these unwanted proteins were also DNA-binding this would lead to nonspecific and difficult to interpret ChIP-Seq data. This is what lead to the need to devise an alternative strategy to use different antibodies to immunoprecipitate CtrA. Antibodies targeting histidine tags are well vetted and available commercially, which lead us to this approach.

The plasmid pET28a is a low-copy plasmid, which is the reason why an indirect cloning method was chosen. The numerous ligations conducted were seldom successful in creating a recombinant plasmid. The potential reasons for this are numerous. For example, the size of the vector and the size of the insert were quite large, making it unfavourable for the ligation reaction to occur. Related to the size of the DNA fragments, the concentration and reaction volume are also important variables. Long DNA fragments need space to move in order for the ends to find each other and close to a circle. I was able to successfully clone one version of the *ctrA* gene but could not duplicate this success with the other versions. Spatial conformation is also likely the reason why the overlap extension PCR with a volume of 50 µL did not yield as much product as volumes of 10-30 µL.

The constructs created using overlap extension PCR are within NEB 5α *E. coli* cells and will need to be transformed into S17-1 *E. coli* cells, so that they can be conjugated into *R. capsulatus*. The *R. capsulatus* strain SBRM1 is an ideal recipient for these constructs as its own *ctrA* is disrupted, therefore it will only produce 6× His-tagged CtrA. Once there are three *R. capsulatus* strains that each hold a different *ctrA* construct, the phenotype of these mutants needs to be tested. Experiments need to be conducted to ensure that the 6× His-tag does not interfere with CtrA function and that the modified CtrA is not toxic to the cell. It is expected that the strain producing wild-type CtrA will have a wild-type level of motility and GTA production. The strain that produces phosphomimetic CtrA is expected to have normal to increased motility and GTA production, and the strain that produces non-phosphorylated CtrA is expected to have no motility and reduced GTA production. Once the activity of the modified CtrAs has been established, a prokaryotic ChIP-Seq protocol can be followed.

2.6 Conclusion

Several cloning methods were attempted to generate the three constructs desired for the planned ChIP-Seq experiment. Ligation-based cloning methods did not yield successful clones. With optimization, overlap extension PCR cloning yielded three plasmids that carried the desired versions of *ctrA*. The construct containing phosphomimetic *ctrA* has a substitution and needs to be further studied or created once more. These constructs will aid in further work in investigating the DNA binding sites of CtrA in *R. capsulatus*.

Chapter 3: Evaluation of the Distribution and Potential Horizontal Gene Transfer of *ctrA* within Alphaproteobacteria

3.1 Abstract

The DNA binding response regulator CtrA is a highly conserved protein in the class Alphaproteobacteria. In many species it serves an essential role as a cell cycle regulator. In others, such as *Rhodobacter capsulatus*, it is not essential, but does have the important role of regulating flagellar biosynthesis, cell motility, and gene transfer agent production. Gene transfer agents are a vehicle for horizontal gene transfer and genes for their production are encoded in many alphaproteobacterial genomes. In this work, I looked at the distribution of *ctrA* within alphaproteobacteria and for evidence of horizontal gene transfer of this gene. Phylogenetic trees were constructed using a large dataset consisting of about 700 sequences and 13 potential cases of HGT were identified using phylogenetic conflict analyses. Several of these were inferred to represent species misclassifications, and these were mostly species from the order Rhizobiales that were misclassified as belonging to the order Rhodobacterales. One species, *Sandarakinorhabdus cyanobacteriorum*, seems to have a *ctrA* gene that was acquired by horizontal gene transfer in an ancestral species. *S. cyanobateriorum* is classified as a member of the order Sphingomonodales, which was supported by my analysis of RpoB sequences, but its CtrA groups closely with those from the order Rhodospirillales in the phylogenetic analyses. Additionally, I found that the orders Magnetococcales, Holosporales, and Pelagibacterales within alphaproteobacteria do not have a *ctrA* homolog.

3.2 Introduction

3.2.1 Horizontal Gene Transfer

Horizontal gene transfer (HGT) occurs when genetic material is transferred between organisms that in a manner different from reproduction. Common examples of HGT that have been co-opted by laboratory scientists are transformation and conjugation (9, 12, 13). Another commonly known example of HGT is transduction (9), which is similar to gene transfer agent (GTA) mediated HGT (15). HGT has been most frequently observed between bacteria, though it has also been observed between a bacterium and a eukaryote (14). Alphaproteobacteria are an interesting class in which to study HGT as there are several species which produce GTAs and most Alphaproteobacteria have a homolog of the response regulator CtrA. By understanding more about the evolutionary history of CtrA and the influence of HGT, we can understand more about HGT and its impact on evolution as a whole.

3.2.2 Conservation of CtrA

CtrA is a highly conserved DNA-binding response regulator in Alphaproteobacteria. Most orders in Alphaproteobacteria have a *ctrA* homolog, with only the order Pelagibacterales lacking a CtrA as of previous analyses. Greene et al. found no evidence of HGT of *ctrA* in Alphaproteobacteria as of 2012. A major change that has happened since the conservation of CtrA in alphaproteobacteria was previously studied is the massive increase in sequenced genomes in the NCBI database. As of 2009, there were more than 500 complete or in progress proteobacterial genomes available in NCBI (3), with alphaproteobacterial sequences being a subset within the 500. At the time of this thesis in 2020, there are thousands of genomes available to study. In addition, there has been recent reclassification and new orders created within the Alphaproteobacteria, which may lead to different conclusions from phylogenetic analyses presented in literature.

3.3 Materials and Methods

3.3.1 Identification of CtrA Sequences

CtrA amino acid sequences were identified using the BLASTp online tool (59) with the default parameters. Searches were performed within the "Non-redundant protein sequences (nr)" database with the *R. capsulatus* CtrA (accession number: AAF13177) as the query. Search results were filtered for a ≥89% query coverage and a percent identity ≥30% as compared to the *R. capsulatus* CtrA sequence.

Sequence alignments were produced using CLUSTAL W (60) in Geneious R11 (Biomatters) and alignments with small datasets (<400 sequences) were polished using TrimAl (61) through the online platform Phylemon 2.0 (62). Large datasets (>400 sequences) were polished manually using Geneious.

3.3.2 Phylogenetic Analyses

The modeltest function in MEGA X (63) was used to determine the best model for distance estimation. The best distance model was selected to perform subsequent phylogenetic analyses. Phylogenetic trees were built using MEGA X with the maximum-likelihood method (64) and bootstrap analyses were performed to evaluate cluster robustness (65).

At first, all of the CtrA BLASTp hits within the percent query coverage and percent identity limits were considered. This dataset consisted of over 3000 sequences and was narrowed down to approximately 700 sequences by removing duplicates and visually inspecting sequences from the same species (sequences were removed if they had less than 5 amino acids different from another sequence). Sequences that were annotated as specific proteins other than CtrA were also removed. Furthermore, this dataset was reduced by removing entries from multispecies and undetermined species through the use of phylogenetic analyses as well as the online tool NCBI Conserved Domain Search (66).

Using the CtrA phylogenetic tree and a literature search, a smaller representative dataset was made with CtrA sequences from 3-10 species from each order in the Alphaproteobacteria. Phylogenetic analyses were also performed using RNA polymerase subunit beta (RpoB) sequences obtained using the BLASTp online tool with default parameters, using the *R. capsulatus* sequence as the query. Accession numbers of sequences used in this study are available in the appendix (Appendix A.4).

Possible HGT events were identified through phylogenetic conflicts within the tree of over 700 CtrA sequences in comparison to the expected taxonomic order for the originating species, excluding those in the outgroup. The expected positions of sequences were determined based on details provided in the NCBI Taxonomy Browser (67, 68). Due to the nature of NCBI-BLAST multiple proteins from the same species will appear in search results if they are closely related. The non-target proteins must be manually filtered out of the dataset as they can interfere with tests for phylogenetic conflict. In this work, proteins that are not CtrA grouped unexpectedly and result in erroneous conclusions about the potential phylogenetic conflict. Within the CtrA search results there were sequences that were the single entry from a species, and there were sequences that were part of multiple entries for a single species, as described above. Sequences that were the single entry from a species were considered to be HGT candidates and underwent further phylogenetic analysis. HGT candidates that continued to show phylogenetic conflict were further investigated for conservation in gene neighbourhoods by using the online tool BLASTn (69) to find *ctrA* in each species.

3.4 Results

Phylogenetic analysis of CtrA sequences that met the criteria of ≥89% query coverage and ≥30% identity indicated that some of the sequences initially included were not in fact CtrA but were a different DNA-binding response regulator (Appendix A4). After assigning every species to an order and determining which clades represented each order, 50 CtrA sequences were found to have potential phylogenetic conflict. However, some of these sequences clustered in an outgroup and were determined to not be genuine CtrA sequences. This was done by adding a known non-CtrA protein from *R. capsulatus* (the highest blast match to the *R. capsulatus* CtrA) to the phylogenetic tree and observing which of the sequences with potential conflict grouped with this non-CtrA sequence. The sequences that did were determined to not be genuine CtrA and therefore removed, paring the list of potential HGT candidates down to 13.

The large number of total sequences yielded trees that were difficult to visually inspect. Therefore, I selected 3-5 CtrA sequences from each order to create a representative subset that was easier to interpret. Phylogenetic analysis of this dataset with the remaining 13 HGT candidate sequences allowed for more manageable phylogenetic trees (Figure 3.1).

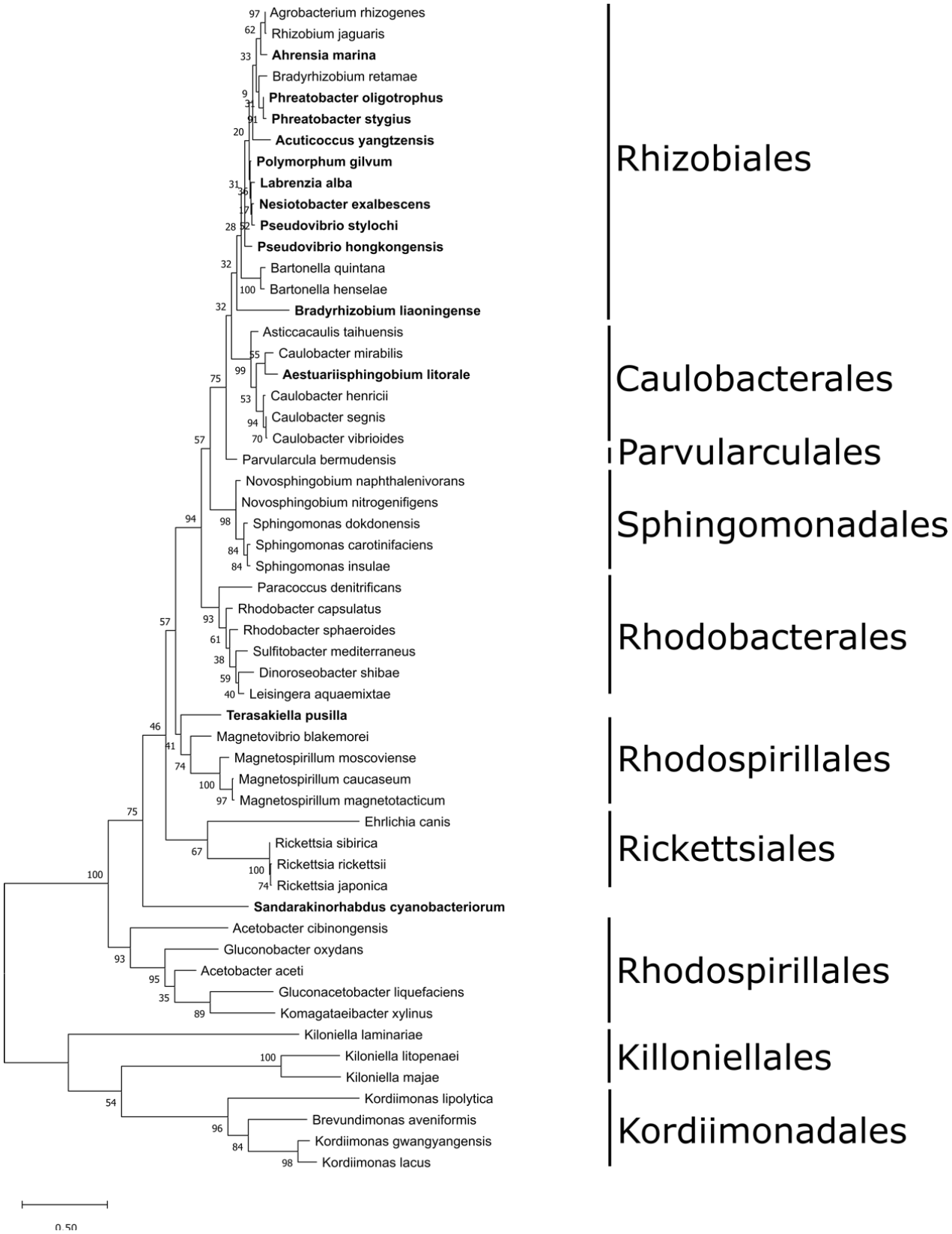

Agrobacterium rhizogenes
Rhizobium jaguaris
Ahrensia marina
Bradyrhizobium retamae
Phreatobacter oligotrophus
Phreatobacter stygius
Acuticoccus yangtzensis
Polymorphum gilvum
Labrenzia alba
Nesiotobacter exalbescens
Pseudovibrio stylochi
Pseudovibrio hongkongensis
Bartonella quintana
Bartonella henselae
Bradyrhizobium liaoningense
Asticcacaulis taihuensis
Caulobacter mirabilis
Aestuariisphingobium litorale
Caulobacter henricii
Caulobacter segnis
Caulobacter vibrioides
Parvularcula bermudensis
Novosphingobium naphthalenivorans
Novosphingobium nitrogenifigens
Sphingomonas dokdonensis
Sphingomonas carotinifaciens
Sphingomonas insulae
Paracoccus denitrificans
Rhodobacter capsulatus
Rhodobacter sphaeroides
Sulfitobacter mediterraneus
Dinoroseobacter shibae
Leisingera aquaemixtae
Terasakiella pusilla
Magnetovibrio blakemorei
Magnetospirillum moscoviense
Magnetospirillum caucaseum
Magnetospirillum magnetotacticum
Ehrlichia canis
Rickettsia sibirica
Rickettsia rickettsii
Rickettsia japonica
Sandarakinorhabdus cyanobacteriorum
Acetobacter cibinongensis
Gluconobacter oxydans
Acetobacter aceti
Gluconacetobacter liquefaciens
Komagataeibacter xylinus
Kiloniella laminariae
Kiloniella litopenaei
Kiloniella majae
Kordiimonas lipolytica
Brevundimonas aveniformis
Kordiimonas gwangyangensis
Kordiimonas lacus
0.50
Rhizobiales
Caulobacterales
Parvularculales
Sphingomonadales
Rhodobacterales
Rhodospirillales
Rickettsiales
Rhodospirillales
Killoniellales
Kordiimonadales

Figure 3.1. Unrooted maximum-likelihood phylogenetic tree of putative CtrA amino acid sequences, with representatives from each order found in the original dataset. Species exhibiting phylogenetic conflict based on the NCBI Taxonomy Browser are shown in bold. The tree was constructed with the Le Gascuel 2008 model and 1000 bootstraps. Branch lengths are proportional to genetic distance as indicated by the scale bar. The labels on the right indicate the order in which most or all species in a clade are classified within.

The CtrA sequences for *Bradyrhizobium liaoningense*, *Phreatobacter oligotrophus*, and *Phreatobacter stygius* did not exhibit phylogenetic conflict in this reduced dataset as they were most closely related to sequences from the order Rhizobiales (Figure 3.1), which is also their classification as supported by the RpoB tree (Figure 3.2). Similarly, the CtrA and RpoB sequences for *Acuticoccus yangtzensis, Ahrensia marina, Labrenzia alba, Nesiotobacter exalbescens, Polymorphum gilvum, Pseudovibrio hongkongensis,* and *Pseudovibrio stylochi* all grouped with others from the Rhizobiales. This was interesting and unexpected as all the aforementioned species are classified as belonging to the order Rhodobacterales, except for *P. gilvum* that is classified as a novel genus *Polymorphum* within "unclassified alphaproteobacteria" (70). *Terasakiella pusilla* also grouped unexpectedly as it is classified within the order Rhizobiales but its sequences clustered with those from the Rhodospirillales in both the CtrA and RpoB phylogenetic analyses.

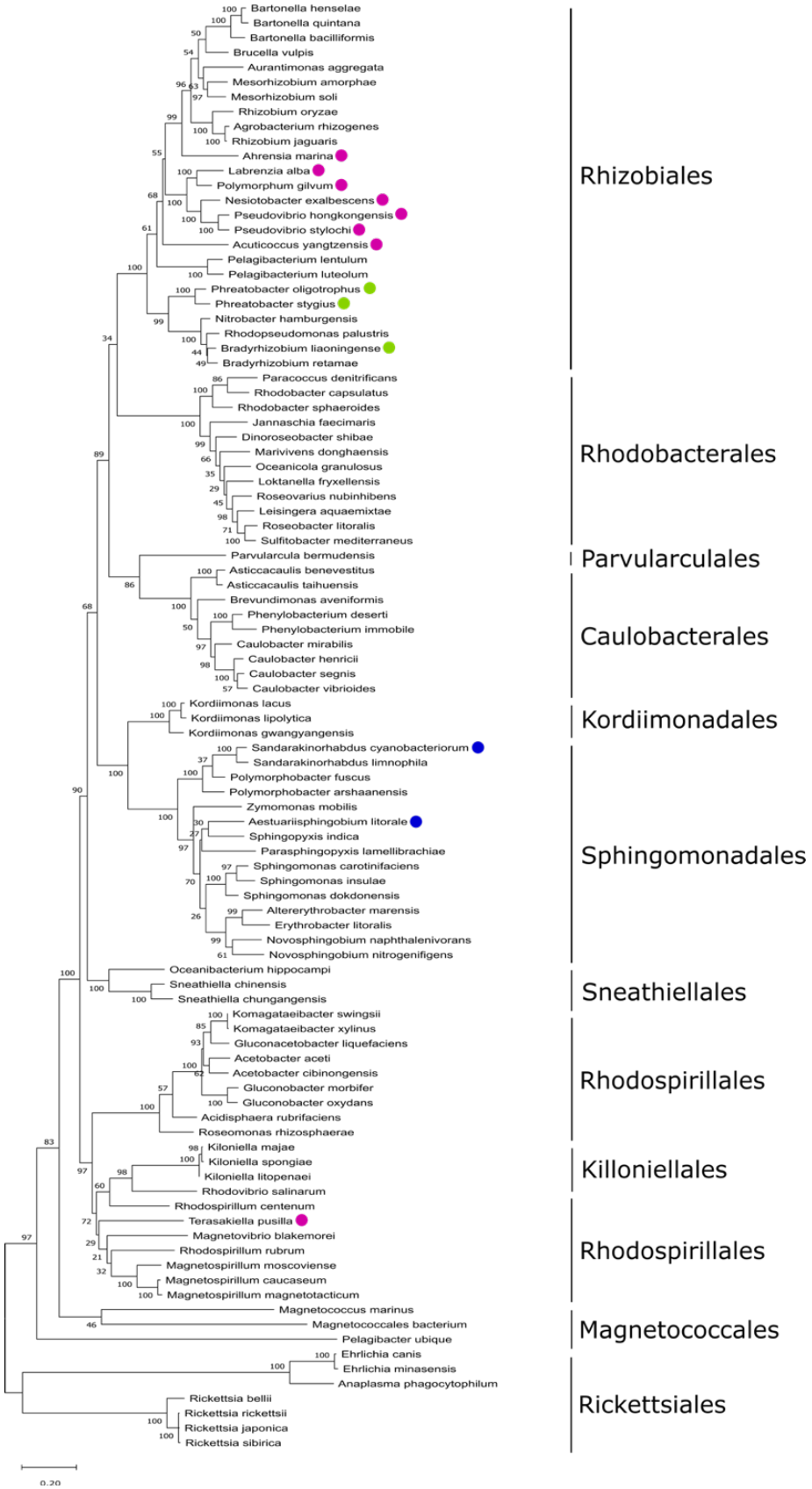

Bartonella henselae
Bartonella quintana
Bartonella bacilliformis
Brucella vulpis
Aurantimonas aggregata
Mesorhizobium amorphae
Mesorhizobium soli
Rhizobium oryzae
Agrobacterium rhizogenes
Rhizobium jaguaris
Ahrensia marina
Labrenzia alba
Polymorphum gilvum
Nesiotobacter exalbescens
Pseudovibrio hongkongensis
Pseudovibrio stylochi
Acuticoccus yangtzensis
Pelagibacterium lentulum
Pelagibacterium luteolum
Phreatobacter oligotrophus
Phreatobacter stygius
Nitrobacter hamburgensis
Rhodopseudomonas palustris
Bradyrhizobium liaoningense
Bradyrhizobium retamae
Paracoccus denitrificans
Rhodobacter capsulatus
Rhodobacter sphaeroides
Jannaschia faecimaris
Dinoroseobacter shibae
Marivivens donghaensis
Oceanicola granulosus
Loktanella fryxellensis
Roseovarius nubinhibens
Leisingera aquaemixtae
Roseobacter litoralis
Sulfitobacter mediterraneus
Parvularcula bermudensis
Asticcacaulis benevestitus
Asticcacaulis taihuensis
Brevundimonas aveniformis
Phenylobacterium deserti
Phenylobacterium immobile
Caulobacter mirabilis
Caulobacter henricii
Caulobacter segnis
Caulobacter vibrioides
Kordiimonas lacus
Kordiimonas lipolytica
Kordiimonas gwangyangensis
Sandarakinorhabdus cyanobacteriorum
Sandarakinorhabdus limnophila
Polymorphobacter fuscus
Polymorphobacter arshaanensis
Zymomonas mobilis
Aestuariisphingobium litorale
Sphingopyxis indica
Parasphingopyxis lamellibrachiae
Sphingomonas carotinifaciens
Sphingomonas insulae
Sphingomonas dokdonensis
Altererythrobacter marensis
Erythrobacter litoralis
Novosphingobium naphthalenivorans
Novosphingobium nitrogenifigens
Oceanibacterium hippocampi
Sneathiella chinensis
Sneathiella chungangensis
Komagataeibacter swingsii
Komagataeibacter xylinus
Gluconacetobacter liquefaciens
Acetobacter aceti
Acetobacter cibinongensis
Gluconobacter morbifer
Gluconobacter oxydans
Acidisphaera rubrifaciens
Roseomonas rhizosphaerae
Kiloniella majae
Kiloniella spongiae
Kiloniella litopenaei
Rhodovibrio salinarum
Rhodospirillum centenum
Terasakiella pusilla
Magnetovibrio blakemorei
Rhodospirillum rubrum
Magnetospirillum moscoviense
Magnetospirillum caucaseum
Magnetospirillum magnetotacticum
Magnetococcus marinus
Magnetococcales bacterium
Pelagibacter ubique
Ehrlichia canis
Ehrlichia minasensis
Anaplasma phagocytophilum
Rickettsia bellii
Rickettsia rickettsii
Rickettsia japonica
Rickettsia sibirica
0.20
Rhizobiales
Rhodobacterales
Parvularculales
Caulobacterales
Kordiimonadales
Sphingomonadales
Sneathiellales
Rhodospirillales
Killoniellales
Rhodospirillales
Magnetococcales
Rickettsiales

Figure 3.2 Unrooted maximum-likelihood phylogenetic tree of RNA polymerase subunit β (RpoB) amino acid sequences with representatives from every order within the class alphaproteobacteria. Species with candidate *ctrA* horizontal gene transfer events are noted with a coloured circle. A blue circle represents phylogenetic conflict between this tree and a maximum-likelihood phylogenetic tree made with CtrA amino acid sequences. A pink circle represents no phylogenetic conflict, but the results contradict in which order the species is classified. Finally, a green circle represents no phylogenetic conflict, in contrast to the previous phylogenetic analysis with 55 RpoB sequences. The tree was constructed with the Le Gascuel 2008 model, +G and 1000 bootstraps. Branch lengths are proportional to genetic distance as indicated by the scale bar. The labels on the right indicate the order in which most or all species in a clade are classified within.

This left only 2 candidates for HGT with phylogenetic conflict between their CtrA and RpoB sequences. *Aestuariisphingobium litorale* and *Sandarakinorhabdus cyanobacteriorum* are classified within the order Sphingomonadales. The CtrA sequence of *A. litorale* was closely related to those from the order Caulobacterales, with a bootstrap value of 99 separating the Caulobacter clade from the Rhizobiales (Figure 3.1). The CtrA sequence of *S. cyanobacteriorum* was individually branched, away from the Sphingomonadales clade (Figure 3.1).

Based on previous analyses of the alphaproteobacterial phylogeny and the expected distribution of *ctrA* in this phylum, the members of the order Rickettsiales should have the most basal *ctrA* and be the most distantly related to the Rhodobacterales (71). The genetic distances among the orders in the produced trees were not supported by this previous literature, and so the dataset was expanded to include 3-10 representative sequences from each order to further resolve relationships and hopefully provide more statistical confidence (Figure 3.3). Due to the placement of the bottommost clade containing sequences from members of the orders Kiloniellales and Kordiimonadales, we suspected that these sequences were not truly CtrA. To determine the identity of these sequences, phylogenetic analysis was completed with an additional sequence of an *R. capsulatus* DNA-binding response regulator, selected by searching the *R. capsulatus* CtrA sequence on BLAST against *R. capsulatus*, and selecting the non-CtrA sequence with the highest identity. The identified *R. capsulatus* DNA-binding response regulator was closely related to the putative CtrA sequences from the Kiloniellales and Kordiimonadales members in the bottommost clade (Figure 3.3). The putative CtrA sequences from the Kiloniellales and Kordiimonadales members were selected based on percent identity to the *R. capsulatus* CtrA sequence. Since the proteins in Kiloniellales and Kordiimonadales that were most similar to *R. capsulatus* CtrA were not true CtrAs, we expected that these bacteria do not contain a true *ctrA* homolog and these sequences were removed from the phylogenetic analysis

(Figure 3.4). However, further literature search found some identified CtrA sequences from members of Kiloniellales and Kordiimonadales and these new sequences were added to subsequent phylogenetic analyses (Figure 3.5).

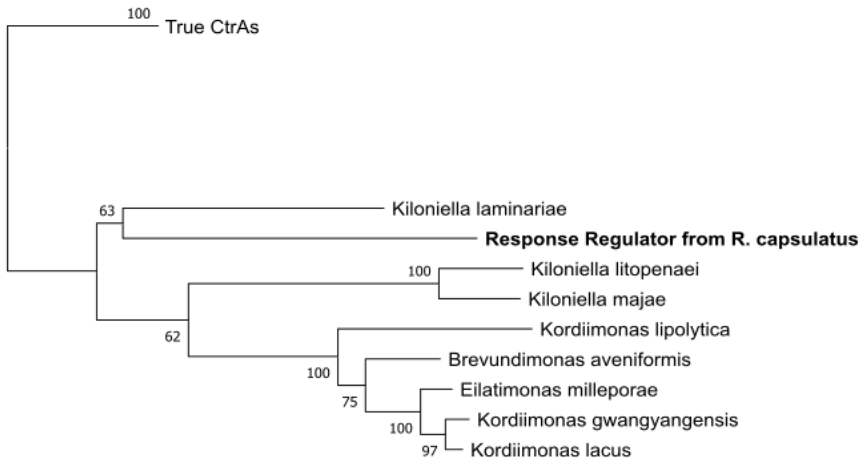


Figure 3.3. Section of a maximum-likelihood phylogenetic tree of CtrA amino acid sequences focusing on the outgroup. Phylogenetic analysis was performed with inclusion of an additional protein sequence of a non-CtrA response regulator from *R. capsulatus*, shown in bold. The clustering of this non-CtrA protein with the putative CtrAs from the other species suggests that these other sequences in the same clade are also not CtrA but another response regulator. The additional clades of true CtrAs have been cropped and are represented by "True CtrAs" and are separated from the clade shown by a bootstrap value of 100. The tree was constructed with the Le Gascuel 2008 model, +G and 100 bootstraps.

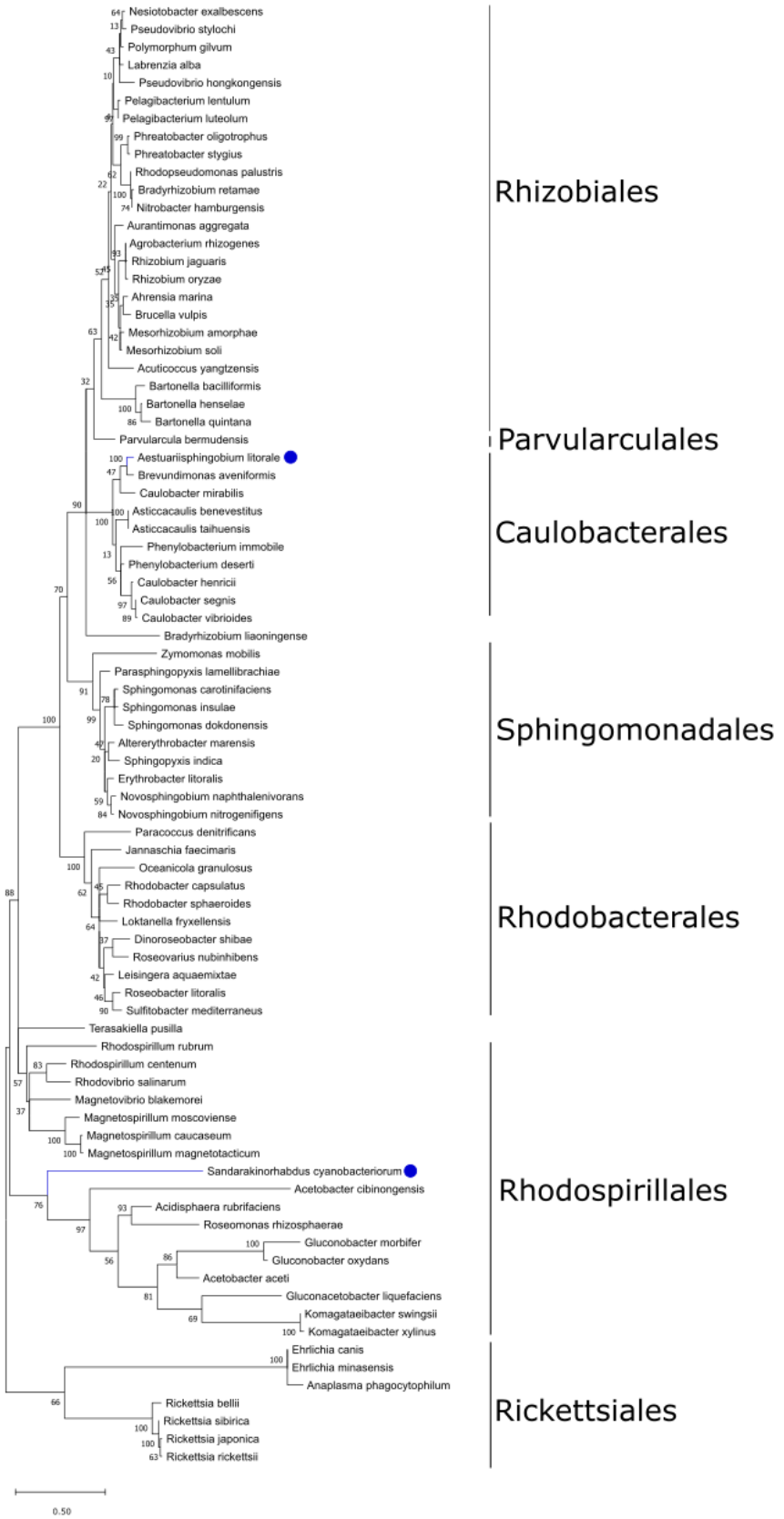

Nesiotobacter exalbescens
Pseudovibrio stylochi
Polymorphum gilvum
Labrenzia alba
Pseudovibrio hongkongensis
Pelagibacterium lentulum
Pelagibacterium luteolum
Phreatobacter oligotrophus
Phreatobacter stygius
Rhodopseudomonas palustris
Bradyrhizobium retamae
Nitrobacter hamburgensis
Aurantimonas aggregata
Agrobacterium rhizogenes
Rhizobium jaguaris
Rhizobium oryzae
Ahrensia marina
Brucella vulpis
Mesorhizobium amorphae
Mesorhizobium soli
Acuticoccus yangtzensis
Bartonella bacilliformis
Bartonella henselae
Bartonella quintana
Parvularcula bermudensis
Aestuariisphingobium litorale
Brevundimonas aveniformis
Caulobacter mirabilis
Asticcacaulis benevestitus
Asticcacaulis taihuensis
Phenylobacterium immobile
Phenylobacterium deserti
Caulobacter henricii
Caulobacter segnis
Caulobacter vibrioides
Bradyrhizobium liaoningense
Zymomonas mobilis
Parasphingopyxis lamellibrachiae
Sphingomonas carotinifaciens
Sphingomonas insulae
Sphingomonas dokdonensis
Altererythrobacter marensis
Sphingopyxis indica
Erythrobacter litoralis
Novosphingobium naphthalenivorans
Novosphingobium nitrogenifigens
Paracoccus denitrificans
Jannaschia faecimaris
Oceanicola granulosus
Rhodobacter capsulatus
Rhodobacter sphaeroides
Loktanella fryxellensis
Dinoroseobacter shibae
Roseovarius nubinhibens
Leisingera aquaemixtae
Roseobacter litoralis
Sulfitobacter mediterraneus
Terasakiella pusilla
Rhodospirillum rubrum
Rhodospirillum centenum
Rhodovibrio salinarum
Magnetovibrio blakemorei
Magnetospirillum moscoviense
Magnetospirillum caucaseum
Magnetospirillum magnetotacticum
Sandarakinorhabdus cyanobacteriorum
Acetobacter cibinongensis
Acidisphaera rubrifaciens
Roseomonas rhizosphaerae
Gluconobacter morbifer
Gluconobacter oxydans
Acetobacter aceti
Gluconacetobacter liquefaciens
Komagataeibacter swingsii
Komagataeibacter xylinus
Ehrlichia canis
Ehrlichia minasensis
Anaplasma phagocytophilum
Rickettsia bellii
Rickettsia sibirica
Rickettsia japonica
Rickettsia rickettsii
0.50
Rhizobiales
Parvularculales
Caulobacterales
Sphingomonadales
Rhodobacterales
Rhodospirillales
Rickettsiales

Figure 3.4. Unrooted maximum-likelihood phylogenetic tree of CtrA amino acid sequences with the non-CtrA out-group removed. Species exhibiting phylogenetic conflict are marked with a blue circle. The tree was constructed with the Le Gascuel 2008 model, +G and 1000 bootstraps. Branch lengths are proportional to genetic distance as indicated by the scale bar. The labels on the right indicate the order in which most or all species in a clade are classified within.

Removal of the sequences suspected of not being CtrA yielded a phylogenetic tree (Figure 3.4) that more closely matched the expectations based on alphaproteobacterial phylogeny from the literature. *A. litorale* and *S. cyanobacteriorum* still exhibited phylogenetic conflict and so the gene neighbourhoods of the *ctrA* sequences were investigated within the whole genome shotgun contigs. This investigation lead to the discovery that the *A. litorale ctrA* gene in the NCBI database is 100% identical to that from the genome sequence of *Brevundimonas aventiformis*. In addition, the *A. litorale* contig which contains *ctrA* has high percent identities to sequences from several *Brevundimonas* species. The discovery of *A. litorale* and the subsequent shotgun sequencing was obtained by culturing a water sample from the Pearl River estuary (72), and it is possible that one or more *Brevundimonas* species were contaminants in the *A. litorale* culture used for genomic sequence determination.

The remaining candidate for HGT, *S. cyanobacteriorum*, was on its own branch in the CtrA phylogenetic tree, without a high bootstrap confidence (Figure 3.4). In order to better elucidate the relationships among these sequences, additional phylogenetic analyses were conducted with an additional species from the *Sandarakinorhabdus* genus, *S. limnophila*, and two members of the order Sphingomonadales that showed high percent identities to the *ctrA* from *S. cyanobacteriorum* on a BLASTn search, *Polymorphobacter fuscus* and *P. arshaanensis*.

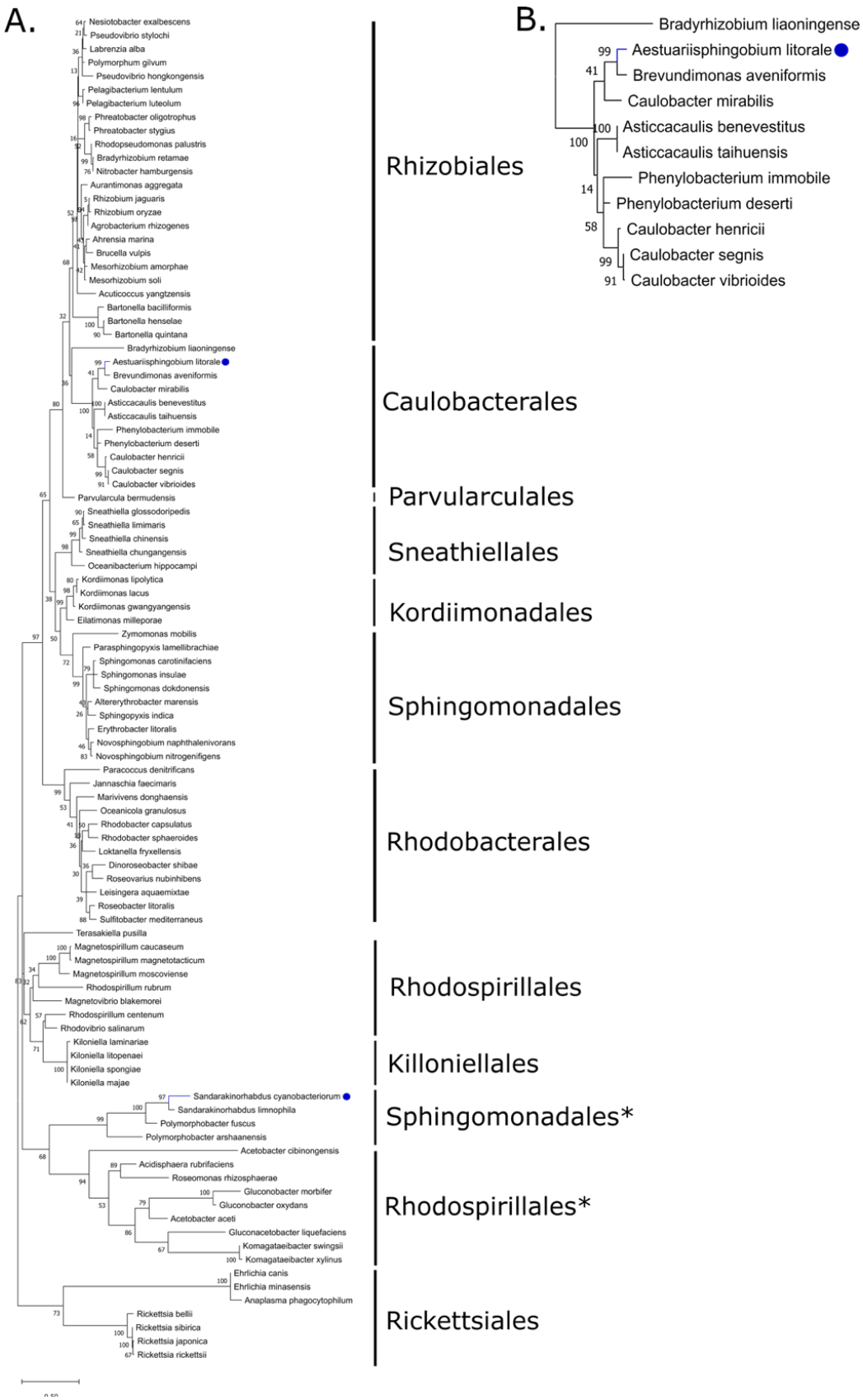

A.
Nesiotobacter exalbescens
Pseudovibrio stylochi
Labrenzia alba
Polymorphum gilvum
Pseudovibrio hongkongensis
Pelagibacterium lentulum
Pelagibacterium luteolum
Phreatobacter oligotrophus
Phreatobacter stygius
Rhodopseudomonas palustris
Bradyrhizobium retamae
Nitrobacter hamburgensis
Aurantimonas aggregata
Rhizobium jaguaris
Rhizobium oryzae
Agrobacterium rhizogenes
Ahrensia marina
Brucella vulpis
Mesorhizobium amorphae
Mesorhizobium soli
Acuticoccus yangtzensis
Bartonella bacilliformis
Bartonella henselae
Bartonella quintana
Bradyrhizobium liaoningense
Aestuariisphingobium litorale
Brevundimonas aveniformis
Caulobacter mirabilis
Asticcacaulis benevestitus
Asticcacaulis taihuensis
Phenylobacterium immobile
Phenylobacterium deserti
Caulobacter henricii
Caulobacter segnis
Caulobacter vibrioides
Parvularcula bermudensis
Sneathiella glossodoripedis
Sneathiella limimaris
Sneathiella chinensis
Sneathiella chungangensis
Oceanibacterium hippocampi
Kordiimonas lipolytica
Kordiimonas lacus
Kordiimonas gwangyangensis
Eilatimonas milleporae
Zymomonas mobilis
Parasphingopyxis lamellibrachiae
Sphingomonas carotinifaciens
Sphingomonas insulae
Sphingomonas dokdonensis
Altererythrobacter marensis
Sphingopyxis indica
Erythrobacter litoralis
Novosphingobium naphthalenivorans
Novosphingobium nitrogenifigens
Paracoccus denitrificans
Jannaschia faecimaris
Marivivens donghaensis
Oceanicola granulosus
Rhodobacter capsulatus
Rhodobacter sphaeroides
Loktanella fryxellensis
Dinoroseobacter shibae
Roseovarius nubinhibens
Leisingera aquaemixtae
Roseobacter litoralis
Sulfitobacter mediterraneus
Terasakiella pusilla
Magnetospirillum caucaseum
Magnetospirillum magnetotacticum
Magnetospirillum moscoviense
Rhodospirillum rubrum
Magnetovibrio blakemorei
Rhodospirillum centenum
Rhodovibrio salinarum
Kiloniella laminariae
Kiloniella litopenaei
Kiloniella spongiae
Kiloniella majae
Sandarakinorhabdus cyanobacteriorum
Sandarakinorhabdus limnophila
Polymorphobacter fuscus
Polymorphobacter arshaanensis
Acetobacter cibinongensis
Acidisphaera rubrifaciens
Roseomonas rhizosphaerae
Gluconobacter morbifer
Gluconobacter oxydans
Acetobacter aceti
Gluconacetobacter liquefaciens
Komagataeibacter swingsii
Komagataeibacter xylinus
Ehrlichia canis
Ehrlichia minasensis
Anaplasma phagocytophilum
Rickettsia bellii
Rickettsia sibirica
Rickettsia japonica
Rickettsia rickettsii
Rhizobiales
Caulobacterales
Parvularculales
Sneathiellales
Kordiimonadales
Sphingomonadales
Rhodobacterales
Rhodospirillales
Killoniellales
Sphingomonadales*
Rhodospirillales*
Rickettsiales
B.
Bradyrhizobium liaoningense
Aestuariisphingobium litorale
Brevundimonas aveniformis
Caulobacter mirabilis
Asticcacaulis benevestitus
Asticcacaulis taihuensis
Phenylobacterium immobile
Phenylobacterium deserti
Caulobacter henricii
Caulobacter segnis
Caulobacter vibrioides

Figure 3.5. Unrooted maximum-likelihood phylogenetic tree of CtrA amino acid sequences with a larger subset of representatives from each order within the Alphaproteobacteria. (A) Total tree. Species exhibiting phylogenetic conflict are marked with a blue circle. The second clades of Sphingomonadales and Rhodospirillales are marked by an asterisk (*). The tree was constructed with the Le Gascuel 2008 model, +G and 1000 bootstraps. Branch lengths are proportional to genetic distance as indicated by the scale bar. The labels on the right indicate the order in which most or all species in a clade are classified within. (B) A section of the larger tree shown in (A), focusing on the Caulobacterales clade. *A. litorale* is marked with a blue circle on both trees.

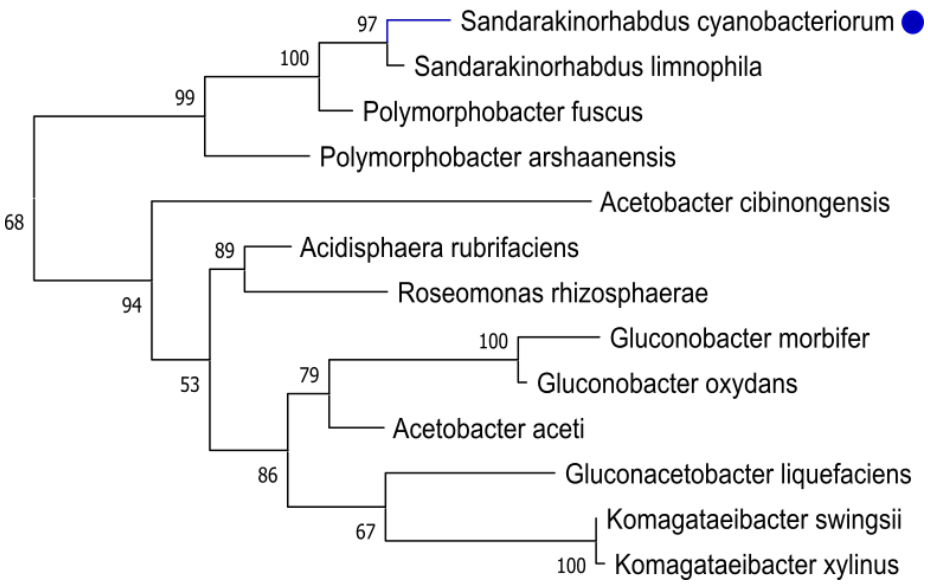


Figure 3.6. Section of a maximum-likelihood phylogenetic tree of CtrA amino acid sequences focusing on the clades of Sphingomonadales and Rhodospirillales. The focal clades are those marked by asterisks in Figure 3.5. *S. cyanobacteriorum* is marked with a blue circle and shows phylogenetic conflict with its placement on the RpoB phylogenetic tree (Figure 3.2). The tree was constructed with the Le Gascuel 2008 model, +G and 1000 bootstraps.

The sequences from *S. limnophila*, *P. fuscus,* and *P. arshaanensis* clustered in the same clade as *S. cyanobacteriorum* on the CtrA (Figure 3.5) and RpoB (Figure 3.2) trees. In Figure 3.2 the four species clustered with the other Sphingomonadales, albeit in their own subclade. This subclade was within a larger clade containing sequences from the Rhodospirillales (Figures 3.5 and 3.6) and this showed phylogenetic conflict. This indicates that the CtrAs of these four species are more closely related to each other and to those from the Rhodospirillales than to those from the other members of the Sphingomonadales. This phylogenetic conflict may indicate there has been a past horizontal gene transfer event.

3.5 Discussion

It is unlikely that genes that are essential for survival are transferred through horizontal gene transfer (9). This likely explains minimal horizontal gene transfer of the DNA-binding response regulator CtrA within alphaproteobacteria. Of the 13 candidates initially identified as potential HGT events, it was subsequently determined that the majority were results of misclassification, rather than a true HGT event. *Phreatobacter oligotrophus* and *Phreatobacter stygius* are part of the genus *Phreatobacter* in the order Rhizobiales. This genus has a distinct lineage based on the 16S rRNA gene (73) that is related to, but distinctly separate from, other members of the order Rhizobiales (74), and this is also supported by my CtrA phylogenetic analysis. *Ahrensia marina* is classified as a member of the Rhizobiales (75) although it is listed as a member of the Rhodobacterales on the NCBI taxonomy browser. It is unclear which classification is correct based on the previous literature, but my phylogenetic analyses of CtrA and RpoB suggest *A. marina* belongs within the order Rhizobiales, supporting the classification by Liu and coworkers (75). We selected RpoB as the protein to compare to the analysis of CtrA sequences as it is an essential gene found in all alphaproteobacteria and a common bacterial marker gene for phylogenies (76).

Additional HGT candidates show strong evidence of being misclassified in the literature. *Acuticoccus yangtzensis* is classified as a member of the family Rhodobacteraceae within the order Rhodobacterales (77), *Labrenzia alba* is indicated as a member of the Rhodobacterales on NCBI and in the literature (78, 79), and *Nesiotobacter exalbescens* is classified as a member of the Rhodobacterales because it is closely related to, but distinct from, *Stappia* (now *Labrenzia*) and *Roseibium* (80). *Polymorphum gilvum* is classified as a member of the Rhodobacterales but shares many gene homologs with the Rhizobiales (81, 82), and *Pseudovibrio hongkongensis* and *Pseudovibrio stylochi* are said to belong to the order Rhodobacterales (83). Therefore, all of

these species have been classified as Rhodobacterales but in my phylogenetic analyses of CtrA and RpoB sequences, they belong within the Rhizobiales.

Terasakiella pusilla, previously *Oceanospirillum pusillum* (84), is classified as a Rhizobiales. Similar to the above examples, it has likely been misclassified. The RpoB sequence of *T. pusilla* is most closely related to those from the Rhodospirillales. Phylogenetic analysis of the CtrA sequences shows that the *T. pusilla* CtrA also clusters in a clade with members of the Rhodospirillales and not in the Rhizobiales clade, which has a bootstrap confidence of 100. In order to support a reclassification of the aforementioned species, further genetic and biochemical tests should be conducted. The classification of these species mostly occurred before high-throughput sequencing technologies became readily accessible and revisiting the classification could lead to further insight into the relationships among Alphaproteobacteria.

Aestuariisphingobium litorale seemed to be a strong candidate for *ctrA* HGT. Its CtrA clustered within the clade of Caulobacterales with a bootstrap value of 100 whereas its RpoB clusters within the Sphingomonadales. Upon further investigation, the entire contig in which the *A. litorale ctrA* was from is highly similar to *Brevundimonas* spp. from the Caulobacterales and the CtrA sequence was identical to that found in *Brevundimonas diminuta*. The sample from which the *A. litorale* sequence was obtained was a freshwater sample from a river (72), and since *Brevunidmonas* species also live in freshwater (85) it is possible that there was contamination in the bacterial culture used for sequencing. Due to these findings, *A. litorale* was subsequently disqualified from being an HGT candidate.

Sandarakinorhabdus cyanobacteriorum is classified as a Sphingomonadales based on its 16S rRNA sequence (86), and this classification is also supported by my RpoB analysis (Figure 3.2). In the phylogenetic analysis of CtrA sequences however, the CtrA of *S. cyanobacteriorum* appears in the same clade as those from the Rhodospirillales. In the 16S rRNA phylogenetic tree

in the literature (86) and as seen in the RpoB analysis in this chapter (Figure 7), Sphingomonadales have two main branches, one of which includes the genera *Sandarakinorhabdus* and *Polymorphobacter*. Interestingly, this branch in the CtrA-based phylogenetic tree does not fall as expected, i.e. next to the corresponding Sphingomonadales branch, but instead is joined to the Rhodospirillales clade with a bootstrap confidence of 68 (Figure 3.6). This indicates that there could have been horizontal gene transfer in the past from a member of Rhodospirillales to an ancestor of *Sandarakinorhabdus* and *Polymorphobacter*. Using *Rhodospirillum centenum* as a model for CtrA function in the order Rhodospirillales, it seems that CtrA does not have an essential role in these organisms. An *R. centenum ctrA* null mutant is viable, although it is less motile (87). Likewise, using *Sphingomonas melonis* as a model for CtrA function in Sphingomonadales, CtrA does not seem to be essential as *ctrA* mutants of *S. melonis* are viable, although they may have some disruptions to cell division (88). These two examples of *ctrA* not being an essential gene in members of the Sphingomonadales and Rhodospirillales fit with the notion that HGT is much more likely to occur with genes that are not essential.

It was previously noted that members of the order Pelagibacterales do not have a *ctrA* homolog. In this study, there were no *ctrA* homologs found in the additional orders Holosporales and Magnetococcales. This could be due to the evolutionary history of *ctrA* within Alphaproteobacteria, where it is suspected that the orders Magnetococcales, Holosporales, and Pelagibacterales were the first to diverge from the ancestral Alphaproteobacteria (Figure 3.7).

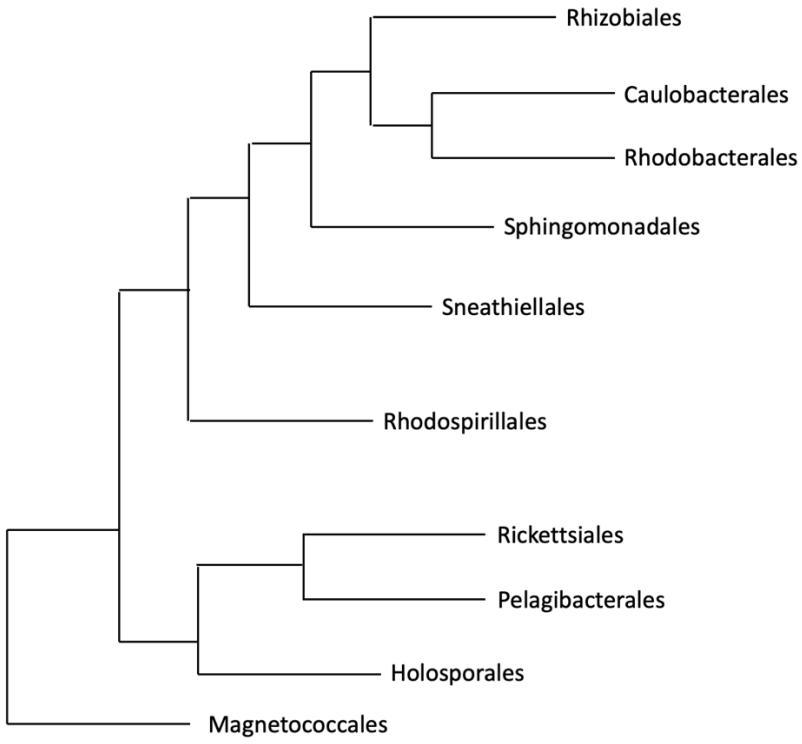


Figure 3.7. A stylized phylogenetic tree representing the general relatedness of orders within Alphaproteobacteria. Branch lengths do not represent evolutionary distance. The tree is based on published phylogenetic analyses (71).

3.6 Conclusion

Using phylogenetic analyses of the CtrA and RpoB protein sequences, the possibility of *ctrA* HGT was investigated within the class Alphaproteobacteria. Phylogenetic conflict between these two analyses was used as an indication of a possible *ctrA* HGT event. Of the 13 HGT candidate sequences initially identified, 3 did not display phylogenetic conflict when a smaller dataset was analyzed, 7 appeared to be mis-classified taxa, and 2 seemed to be true candidate HGT events. *A. litorale* was subsequently disqualified from having an HGT event due to discrepancies in the literature and likely contamination in the genomic sequencing data. However, an ancestor of *S. cyanobacteriorum* seems to have experienced a *ctrA* HGT event as

the phylogenetic conflict exhibited by this species is also found for its closest relatives, *Polymorphum* spp. HGT was only observed between the orders Rhodospirillales and Sphingomonadales, both of which appear to have CtrA with non-essential roles, supporting the notion that HGT generally occurs with non-essential genes. Additionally, the orders Magnetococcales, Holosporales, and Pelagibacterales were found to not have a CtrA homolog.

Chapter 4: Conclusions

Response regulators have very important roles in the functioning of a bacterium. Everything within a cell is regulated either by a pathway or regulatory proteins. Within the class Alphaproteobacteria, the DNA binding response regulator CtrA is highly conserved and is often involved in regulating motility. In some species, CtrA also has an essential role as a cell cycle regulator. In *R. capsulatus*, a model organism for studying gene transfer agents (GTAs), CtrA has the important role of regulating GTA production. GTAs are mediators of gene transfer, in which one bacterium can give genetic information to another bacterium. CtrA regulates and induces expression of various genes depending on its phosphorylation state, and it is clear that both phosphorylated and non-phosphorylated CtrA contribute to functions of the cell in *R. capsulatus* as modification of the phosphorylation states of CtrA has been observed to lead to a loss of a phenotype. For example, a loss of GTA capsid protein production can be observed if there is a lack of non-phosphorylated CtrA in *R. capsulatus* cells (32).

In this work, I used overlap extension PCR to create three constructs of *ctrA*, all with a 6×-histidine tag. In the constructs, three versions of CtrA were created, one is the wild type with aspartic acid as the 51st residue, which is the site of phosphorylation for its regulation in the cell. The second encodes a non-phosphorylatable version of CtrA with alanine in the 51st residue. Finally, a phosphomimetic version was generated that has glutamic acid at the 51st position. CtrA plays a role in the regulation of gene expression in both phosphorylated and non-phosphorylated forms. The purpose of these constructs is to generate strains which will be used for further studies to determine the DNA binding sites of CtrA as well as if the binding sites change depending on phosphorylation state. With this information we can better understand the role CtrA plays in the regulation of GTA expression and production.

GTAs are one of many modes of horizontal gene transfer that can be observed in bacteria. In recent years, there has been a substantial increase in the number of alphaproteobacterial genomes sequenced. This advance has allowed for larger and more comprehensive datasets in which to study horizontal gene transfer. There were CtrA homologs found in all orders of Alphaproteobacteria except Magnetococcales, Holosporales, and Pelagibacterales. Using the principle of phylogenetic conflict, I looked for evidence of horizontal gene transfer of *ctrA* within the Alphaproteobacteria. This analysis revealed multiple species misclassifications and one potential *ctrA* horizontal gene transfer event. *Sandarakinorhabdus cyanobacteriorum* is classified as a Sphingomonadales, as supported by phylogenetic analysis of RpoB, but its CtrA is more similar to those of the order Rhodospirillales. The same is true of the related species *Sandarakinorhabdus limnophila*, *Polymorphobacter fuscus* and *Polymorphobacter arshaanensis*, suggesting that a horizontal gene transfer occurred in a bacterium ancestral to this lineage. Horizontal gene transfer occurs more frequently with non-essential genes and since CtrA has an essential role in many species, this may be why we do not see more horizontal gene transfer events.

This work has laid the foundation for further work on investigating the roles of CtrA and the effects of phosphorylation on its activity. Three constructs producing three versions of CtrA were prepared for further investigation of the DNA binding sites of each version of the protein. A bioinformatics approach was employed to conduct phylogenetic analysis on a large dataset from the Alphaproteobacteria, looking for *ctrA* horizontal gene transfer events.

References

1. Berman JJ. 2012. The Alpha Proteobacteria, p. 25–31. *In* Taxonomic Guide to Infectious Diseases. Elsevier.

2. Stackebrandt E. 2001. Bacterial Biodiversity, p. 307–316. *In* Encyclopedia of Biodiversity. Elsevier.

3. Swingley WD, Blankenship RE, Raymond J. 2009. Evolutionary Relationships Among Purple Photosynthetic Bacteria and the Origin of the Proteobacterial Photosynthetic Systems, p. 17–29. *In* The purple phototrophic bacteria. Springer, Dordrecht, The Netherlands.

4. Madigan MT, Jung DO. 2009. An Overview of Purple Bacteria: Systematics, Physiology, and Habitats, p. 1–15. *In* The purple phototrophic bacteria. Springer, Dordrecht, The Netherlands.

5. Imhoff JF, Trüper HG, Pfennig N. 1984. Rearrangement of the Species and Genera of the Phototrophic "Purple Nonsulfur Bacteria." International Journal of Systematic Bacteriology 34:340–343.

6. van Niel CB. 1944. The culture, general physiology, morphology, and classification of the non-sulfur purple and brown bacteria. Bacteriological Reviews 8:1–118.

7. Madigan MT, Gest H. 1979. Growth of the photosynthetic bacterium *Rhodopseudomonas capsulata* chemoautotrophically in darkness with H2 as the energy source. Journal of Bacteriology 137:524–530.

8. Marrs B. 1974. Genetic Recombination in *Rhodopseudomonas capsulata*. Proceedings of the National Academy of Sciences 71:971–973.

9. Soucy SM, Huang J, Gogarten JP. 2015. Horizontal gene transfer: building the web of life. Nature Reviews Genetics 16:472–482.

10. Huang J. 2013. Horizontal gene transfer in eukaryotes: The weak-link model: Insights & Perspective. BioEssays 868–875.

11. Hobot JA. 2015. Bacterial Ultrastructure, p. 7–32. *In* Molecular Medical Microbiology, 2nd ed. Academic Press.

12. Leung MM, Beatty JT. 2013. Bacterial Conjugation in *Rhodobacter capsulatus*. Bio-protocol 3(11).

13. Norman A, Hansen LH, Sørensen SJ. 2009. Conjugative plasmids: vessels of the communal gene pool. Philosophical Transactions of the Royal Society B: Biological Sciences 364:2275–2289.

14. Kyndt T, Quispe D, Zhai H, Jarret R, Ghislain M, Liu Q, Gheysen G, Kreuze JF. 2015. The genome of cultivated sweet potato contains *Agrobacterium* T-DNAs with expressed genes: An example of a naturally transgenic food crop. Proceedings of the National Academy of Sciences 112:5844–5849.

15. Lang AS, Zhaxybayeva O, Beatty JT. 2012. Gene transfer agents: phage-like elements of genetic exchange. Nature Reviews Microbiology 10:472–482.

16. Lang AS, Westbye AB, Beatty JT. 2017. The Distribution, Evolution, and Roles of Gene Transfer Agents in Prokaryotic Genetic Exchange. Annual Review of Virology 4:87–104.

17. Hynes AP, Mercer RG, Watton DE, Buckley CB, Lang AS. 2012. DNA packaging bias and differential expression of gene transfer agent genes within a population during production and release of the *Rhodobacter capsulatus* gene transfer agent, RcGTA: Gene transfer agent gene expression and DNA packaging. Molecular Microbiology 85:314–325.

18. Yen HC, Hu NT, Marrs BL. 1979. Characterization of the gene transfer agent made by an overproducer mutant of *Rhodopseudomonas capsulata*. Journal of Molecular Biology 131:157–168.

19. Leung MM, Brimacombe CA, Spiegelman GB, Beatty JT. 2012. The GtaR protein negatively regulates transcription of the gtaRI operon and modulates gene transfer agent (RcGTA) expression in *Rhodobacter capsulatus*: Quorum-sensing in *Rhodobacter capsulatus*. Molecular Microbiology 83:759–774.

20. Redfield RJ, Soucy SM. 2018. Evolution of Bacterial Gene Transfer Agents. Frontiers in Microbiology 9:1–14.

21. Lang AS, Beatty JT. 2007. Importance of widespread gene transfer agent genes in a-proteobacteria. Trends in Microbiology 15:9.

22. Wang H, Ziesche L, Frank O, Michael V, Martin M, Petersen J, Schulz S, Wagner-Döbler I, Tomasch J. 2014. The CtrA phosphorelay integrates differentiation and communication in the marine alphaproteobacterium *Dinoroseobacter shibae*. BMC Genomics 15:130.

23. Berglund EC, Frank AC, Calteau A, Vinnere Pettersson O, Granberg F, Eriksson A-S, Näslund K, Holmberg M, Lindroos H, Andersson SGE. 2009. Run-Off Replication of Host-Adaptability Genes Is Associated with Gene Transfer Agents in the Genome of Mouse-Infecting Bartonella grahamii. PLoS Genetics 5:e1000546.

24. Lang AS, Beatty JT. 2000. Genetic analysis of a bacterial genetic exchange element: The gene transfer agent of *Rhodobacter capsulatus*. Proceedings of the National Academy of Sciences 97:859–864.

25. Schaefer AL, Taylor TA, Beatty JT, Greenberg EP. 2002. Long-Chain Acyl-Homoserine Lactone Quorum-Sensing Regulation of Rhodobacter capsulatus Gene Transfer Agent Production. Journal of Bacteriology 184:6515–6521.

26. Leung MM, Brimacombe CA, Beatty JT. 2013. Transcriptional regulation of the *Rhodobacter capsulatus* response regulator CtrA. Microbiology 159:96–106.

27. Ryan KR, Judd EM, Shapiro L. 2002. The CtrA Response Regulator Essential for *Caulobacter crescentus* Cell-cycle Progression Requires a Bipartite Degradation Signal for Temporally Controlled Proteolysis. Journal of Molecular Biology 324:443–455.

28. Abraham W-R, Strömpl C, Meyer H, Lindholst S, Moore ERB, Christ R, Vancanneyt M, Tindall BJ, Bennasar A, Smit J, Tesar M. 1999. Phylogeny and polyphasic taxonomy of Caulobacter species. Proposal of Maricaulis gen. nov. with Maricaulis maris (Poindexter) comb. nov. as the type species, and emended description of the genera Brevundirnonas and Caulobacter. International Journal of Systematic Bacteriology 49:1053–1073.

29. Biondi EG, Reisinger SJ, Skerker JM, Arif M, Perchuk BS, Ryan KR, Laub MT. 2006. Regulation of the bacterial cell cycle by an integrated genetic circuit. Nature 444:899–904.

30. Chen YE, Tsokos CG, Biondi EG, Perchuk BS, Laub MT. 2009. Dynamics of Two Phosphorelays Controlling Cell Cycle Progression in *Caulobacter crescentus*. Journal of Bacteriology 191:7417–7429.

31. Brilli M, Fondi M, Fani R, Mengoni A, Ferri L, Bazzicalupo M, Biondi EG. 2010. The diversity and evolution of cell cycle regulation in alpha-proteobacteria: a comparative genomic analysis. BMC Systems Biology 4:16.

32. Mercer RG, Quinlan M, Rose AR, Noll S, Beatty JT, Lang AS. 2012. Regulatory systems controlling motility and gene transfer agent production and release in *Rhodobacter capsulatus*. FEMS Microbiology Letters 331:53–62.

33. Ryan KR, Judd EM, Shapiro L. 2002. The CtrA Response Regulator Essential for Caulobacter crescentus Cell-cycle Progression Requires a Bipartite Degradation Signal for Temporally Controlled Proteolysis. Journal of Molecular Biology 324:443–455.

34. Westbye AB, Leung MM, Florizone SM, Taylor TA, Johnson JA, Fogg PC, Beatty JT. 2013. Phosphate Concentration and the Putative Sensor Kinase Protein CckA Modulate Cell Lysis and Release of the *Rhodobacter capsulatus* Gene Transfer Agent. Journal of Bacteriology 195:5025–5040.

35. Quon KC, Marczynski GT, Shapiro L. 1996. Cell Cycle Control by an Essential Bacterial Two-Component Signal Transduction Protein. Cell 84:83–93.

36. Domian IJ, Reisenauer A, Shapiro L. 1999. Feedback control of a master bacterial cell-cycle regulator. Proceedings of the National Academy of Sciences 96:6648–6653.

37. Westbye AB, Kater L, Wiesmann C, Ding H, Yip CK, Beatty JT. 2018. The Protease ClpXP and the PAS Domain Protein DivL Regulate CtrA and Gene Transfer Agent Production in *Rhodobacter capsulatus*. Applied and Environmental Microbiology 84.

38. Mercer RG, Callister SJ, Lipton MS, Pasa-Tolic L, Strnad H, Paces V, Beatty JT, Lang AS. 2010. Loss of the Response Regulator CtrA Causes Pleiotropic Effects on Gene Expression but Does Not Affect Growth Phase Regulation in *Rhodobacter capsulatus*. J Bacteriol 192:2701–2710.

39. Brimacombe CA, Ding H, Beatty JT. 2014. *Rhodobacter capsulatus* DprA is essential for RecA-mediated gene transfer agent (RcGTA) recipient capability regulated by quorum-sensing and the CtrA response regulator: DprA, CtrA, and RecA are required for gene transfer agent recipient capability. Molecular Microbiology 92:1260–1278.

40. Chen YE, Tsokos CG, Biondi EG, Perchuk BS, Laub MT. 2009. Dynamics of Two Phosphorelays Controlling Cell Cycle Progression in Caulobacter crescentus. Journal of Bacteriology 191:7417–7429.

41. Lin W, Wang Y, Han X, Zhang Z, Wang C, Wang J, Yang H, Lu Y, Jiang W, Zhao G-P, Zhang P. 2014. Atypical OmpR/PhoB Subfamily Response Regulator GlnR of Actinomycetes Functions as a Homodimer, Stabilized by the Unphosphorylated Conserved Asp-focused Charge Interactions. Journal of Biological Chemistry 289:15413–15425.

42. Koppenhöfer S, Wang H, Scharfe M, Kaever V, Wagner-Döbler I, Tomasch J. 2019. Integrated Transcriptional Regulatory Network of Quorum Sensing, Replication Control, and SOS Response in *Dinoroseobacter shibae*. Frontiers in Microbiology 10.

43. Westbye AB, Beatty JT, Lang AS. 2017. Guaranteeing a captive audience: coordinated regulation of gene transfer agent (GTA) production and recipient capability by cellular regulators. Current Opinion in Microbiology 38:122–129.

44. Greene SE, Brilli M, Biondi EG, Komeili A. 2012. Analysis of the CtrA Pathway in *Magnetospirillum* Reveals an Ancestral Role in Motility in Alphaproteobacteria. Journal of Bacteriology 194:2973–2986.

45. Bellefontaine A-F, Pierreux CE, Mertens P, Vandenhaute J, Letesson J-J, Bolle XD. 2002. Plasticity of a transcriptional regulation network among alpha-proteobacteria is supported by the identification of CtrA targets in *Brucella abortus*. Molecular Microbiology 43:945–960.

46. Westbye AB, Kuchinski K, Yip CK, Beatty JT. 2016. The Gene Transfer Agent RcGTA Contains Head Spikes Needed for Binding to the *Rhodobacter capsulatus* Polysaccharide Cell Capsule. Journal of Molecular Biology 428:477–491.

47. Beatty JT, Gest H. 1981. Generation of succinyl-coenzyme A in photosynthetic bacteria. Archives of Microbiology 129:335–340.

48. Strnad H, Lapidus A, Paces J, Ulbrich P, Vlcek C, Paces V, Haselkorn R. 2010. Complete Genome Sequence of the Photosynthetic Purple Nonsulfur Bacterium *Rhodobacter capsulatus* SB 1003. Journal of Bacteriology 192:3545–3546.

49. Yen HC, Marrs B. 1976. Map of genes for carotenoid and bacteriochlorophyll biosynthesis in *Rhodopseudomonas capsulata*. Journal of Bacteriology 126:619–629.

50. Simon R, Priefer U, Pühler A. 1983. A broad host range mobilization system for in vivo genetic engineering: Transposon mutagenesis in gram negative bacteria. Bio/Technol 784–791.

51. Young CS, Reyes RC, Beatty JT. 1998. Genetic Complementation and Kinetic Analyses of *Rhodobacter capsulatus* ORF1696 Mutants Indicate that the ORF1696 Protein Enhances Assembly of the Light-Harvesting I Complex. Journal of Bacteriology 180:1759–1765.

52. Marx CJ, Lidstrom ME. 2015. Development of improved versatile broad-host- range vectors for use in methylotrophs and other Gram-negative bacteria 11.

53. Bryksin A, Matsumura I. 2013. Overlap Extension PCR Cloning, p. 31–42. *In* Polizzi, KM, Kontoravdi, C (eds.), Synthetic Biology. Humana Press, Totowa, NJ.

54. Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-Resolution Profiling of Histone Methylations in the Human Genome. Cell 129:823–837.

55. Bonocora RP, Wade JT. 2015. ChIP-Seq for Genome-Scale Analysis of Bacterial DNA-Binding Proteins, p. 327–340. *In* Artsimovitch, I, Santangelo, TJ (eds.), Bacterial Transcriptional Control: Methods and Protocols. Springer Science+Business Media New York.

56. Grainger DC, Overton TW, Reppas N, Wade JT, Tamai E, Hobman JL, Constantinidou C, Struhl K, Church G, Busby SJW. 2004. Genomic Studies with *Escherichia coli* MelR Protein: Applications of Chromatin Immunoprecipitation and Microarrays. Journal of Bacteriology 186:6938–6943.

57. Domian IJ, Quon KC, Shapiro L. 1997. Cell Type-Specific Phosphorylation and Proteolysis of a Transcriptional Regulator Controls the G1-to-S Transition in a Bacterial Cell Cycle. Cell 90:415–424.

58. Siam R, Marczynski GT. 2003. Glutamate at the phosphorylation site of response regulator CtrA provides essential activities without increasing DNA binding. Nucleic Acids Research 31:1775–1779.

59. Gish W, States DJ. 1993. Identification of protein coding regions by database similarity search. Nature Genetics 3:266–272.

60. Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22:4673–4680.

61. Capella-Gutiérrez S, Silla-Martínez J, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973.

62. Sanchez R, Serra F, Tarraga J, Medina I, Carbonell J, Pulido L, de Maria A, Capella-Gutierrez S, Huerta-Cepas J, Gabaldon T, Dopazo J, Dopazo H. 2011. Phylemon 2.0: a suite of web-tools for molecular evolution, phylogenetics, phylogenomics and hypotheses testing. Nucleic Acids Research 39:W470–W474.

63. Stecher G, Tamura K, Kumar S. 2020. Molecular Evolutionary Genetics Analysis (MEGA) for macOS. Molecular Biology and Evolution 37:1237–1239.

64. Felsenstein J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17:368–376.

65. Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783.

66. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Marchler GH, Song JS, Thanki N, Yamashita RA, Yang M, Zhang D, Zheng C, Lanczycki CJ, Marchler-Bauer A. 2020. CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Research 48:D265–D268.

67. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2009. GenBank. Nucleic Acids Research 37:D26–D31.

68. Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Edgar R, Federhen S, Feolo M, Geer LY, Helmberg W, Kapustin Y, Landsman D, Lipman DJ, Madden TL, Maglott DR, Miller V, Mizrachi I, Ostell J, Pruitt KD, Schuler GD, Sequeira E, Sherry ST, Shumway M, Sirotkin K, Souvorov A, Starchenko G, Tatusova TA, Wagner L, Yaschenko E, Ye J. 2009. Database resources of the National Center for Biotechnology Information. Nucleic Acids Research 37:D5–D15.

69. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

70. Li S-G, Tang Y-Q, Nie Y, Cai M, Wu X-L. 2011. Complete Genome Sequence of *Polymorphum gilvum* SL003B-26A1T, a Crude Oil-Degrading Bacterium from Oil-Polluted Saline Soil. Journal of Bacteriology 193:2894–2895.

71. Muñoz-Gómez SA, Hess S, Burger G, Lang BF, Susko E, Slamovits CH, Roger AJ. 2019. An updated phylogeny of the Alphaproteobacteria reveals that the parasitic *Rickettsiales* and *Holosporales* have independent origins. eLife 8:1–23.

72. Li X, Li J-L, Zhang X-T, Duan L, Asem MD, Xiao M, Mou X, Salam N, Li W-J. 2019. *Aestuariisphingobium litorale* gen. nov., sp. nov., a novel proteobacterium isolated from a water sample of Pearl River estuary. Antonie van Leeuwenhoek 112:1357–1367.

73. Tóth EM, Vengring A, Homonnay ZG, Kéki Zs, Spröer C, Borsodi AK, Márialigeti K, Schumann P. 2014. *Phreatobacter oligotrophus* gen. nov., sp. nov., an alphaproteobacterium isolated from ultrapure water of the water purification system of a power plant. International Journal of Systematic and Evolutionary Microbiology 64:839–845.

74. Lee SD, Joung Y, Cho J-C. 2017. *Phreatobacter stygius* sp. nov., isolated from pieces of wood in a lava cave and emended description of the genus *Phreatobacter*. International Journal of Systematic and Evolutionary Microbiology 67:3296–3300.

75. Liu J, Wang Y, Liu Y, Zhang X-H. 2016. *Ahrensia marina* sp. nov., a dimethylsulfoniopropionate-cleaving bacterium isolated from seawater, and emended descriptions of the genus *Ahrensia* and *Ahrensia kielensis*. International Journal of Systematic and Evolutionary Microbiology 66:874–880.

76. Case RJ, Boucher Y, Dahllöf I, Holmström C, Doolittle WF, Kjelleberg S. 2007. Use of 16S rRNA and rpoB Genes as Molecular Markers for Microbial Ecology Studies. Applied and Environmental Microbiology 73:278–288.

77. Hou L, Zhang Y, Sun J, Xie X. 2015. *Acuticoccus yangtzensis* gen. nov., sp. nov., a Novel Member in the Family *Rhodobacteraceae*, Isolated from the Surface Water of the Yangtze Estuary. Current Microbiology 70:176–182.

78. Biebl H, Pukall R, Lünsdorf H, Schulz S, Allgaier M, Tindall BJ, Wagner-Döbler I. 2007. Description of *Labrenzia alexandrii* gen. nov., sp. nov., a novel alphaproteobacterium containing bacteriochlorophyll a, and a proposal for reclassification of *Stappia aggregata* as *Labrenzia aggregata* comb. nov., of *Stappia marina* as *Labrenzia marina* comb. nov. and of *Stappia alba* as *Labrenzia alba* comb. nov., and emended descriptions of the genera *Pannonibacter*, *Stappia* and *Roseibium*, and of the species *Roseibium denhamense* and *Roseibium hamelinense*. International Journal of Systematic and Evolutionary Microbiology 57:1095–1107.

79. Pujalte MJ, Carmen Macián M, Arahal DR, Garay E. 2005. *Stappia alba* sp. nov., isolated from Mediterranean oysters. Systematic and Applied Microbiology 28:672–678.

80. Donachie SP, Bowman JP, Alam M. 2006. *Nesiotobacter exalbescens* gen. nov., sp. nov., a moderately thermophilic alphaproteobacterium from an Hawaiian hypersaline lake. International Journal of Systematic and Evolutionary Microbiology 56:563–567.

81. Cai M, Wang L, Cai H, Li Y, Tang Y-Q, Wu X-L. 2011. *Rubrimonas shengliensis* sp. nov. and *Polymorphum gilvum* gen. nov., sp. nov., novel members of Alphaproteobacteria from crude oil contaminated saline soil. Systematic and Applied Microbiology 34:321–327.

82. Nie Y, Tang Y-Q, Li Y, Chi C-Q, Cai M, Wu X-L. 2012. The Genome Sequence of *Polymorphum gilvum* SL003B-26A1T Reveals Its Genetic Basis for Crude Oil Degradation and Adaptation to the Saline Soil. PLoS ONE 7:e31261.

83. Shieh WY, Lin Y-T, Jean WD. 2004. *Pseudovibrio denitrificans* gen. nov., sp. nov., a marine, facultatively anaerobic, fermentative bacterium capable of denitrification. International Journal of Systematic and Evolutionary Microbiology 54:2307–2312.

84. Satomi M. 2002. Phylogenetic study of the genus *Oceanospirillum* based on 16S rRNA and gyrB genes: emended description of the genus *Oceanospirillum*, description of *Pseudospirillum* gen. nov., *Oceanobacter* gen. nov. and *Terasakiella* gen. nov. and transfer of *Oceanospirillum jannaschii* and *Pseudomonas stanieri* to *Marinobacterium* as *Marinobacterium jannaschii* comb. nov. and *Marinobacterium stanieri* comb. nov. International Journal of Systematic and Evolutionary Microbiology 52:739–747.

85. Abraham W-R, Estrela AB, Nikitin DI, Smit J, Vancanneyt M. 2010. *Brevundimonas halotolerans* sp. nov., *Brevundimonas poindexterae* sp. nov. and *Brevundimonas staleyi* sp. nov., prosthecate bacteria from aquatic habitats. International Journal of Systematic and Evolutionary Microbiology 60:1837–1843.

86. Cai H, Cui H, Zeng Y, An M, Jiang H. 2018. *Sandarakinorhabdus cyanobacteriorum* sp. nov., a novel bacterium isolated from cyanobacterial aggregates in a eutrophic lake. International Journal of Systematic and Evolutionary Microbiology 68:730–735.

87. Bird TH, MacKrell A. 2011. A CtrA homolog affects swarming motility and encystment in *Rhodospirillum centenum*. Archives of Microbiology 193:451–459.

88. Francez-Charlot A, Kaczmarczyk A, Vorholt JA. 2015. The branched CcsA/CckA-ChpT-CtrA phosphorelay of *Sphingomonas melonis* controls motility and biofilm formation: CtrA of *Sphingomonas melonis*. Molecular Microbiology 97:47–63.

Appendix

A.1 Primers

Primer name	Use	Primer sequence	Reference
SDM D51A ctrA F	Create non-phosphorylatable point mutation	CTGATCCTTCTCGCTCTGAACCTTCCCG AC	(32)
SDM D51A ctrA R	Create non-phosphorylatable point mutation	GTCGGGAAGGTTCAGAGCGAGAAGGA TCAG	(32)
SDM D51E ctrA F	Create phosphomimetic point mutation	CTGATCCTTCTCGAACTGAACCTTCCC GAC	(32)
SDM D51E ctrA R	Create phosphomimetic point mutation	GTCGGGAAGGTTCAGTTCGAGAAGGA TCAG	(32)
ctrA F	Amplify insert for ligation	AAAGTCGACCTTTAAGAAGGAGAATG G	This study
ctrA R	Amplify insert for ligation	ACGGAGCTCGAATTCTCAGGC	This study
OverFctrA	Overlap extension part A	GGCTGCAGGTCGACTCTAGAGGATCCC CGGGGACCTTTAAGAAGGAGATATAC CATGGGCAGC	This study
OverRctrA	Overlap extension part A	CCTCTTCGCTATTACGCCAGCTGGCGA AAGCGAATTCTCAGAGAAGCGGAGAC TCAGG	This study
ctrA middle F	Sequencing	CATCGCGAAGAGCTGGTGG	This study
T7	Sequencing	TAATACGACTCACTATAGGG	Invitrogen
T7term	Sequencing	GCTAGTTATTGCTCAGCGG	Invitrogen
M13F (-20) Invitrogen	Sequencing	GTAAAACGACGGCCAGT	Invitrogen
M13 Rev (-27) – Invitrogen	Sequencing	CAGGAAACAGCTATGAC	Invitrogen
pRR5C F	Sequencing	AATTGCCTACTGAGCGCTGC	Invitrogen
pRR5C R	Sequencing	GGGGGCTGAAAGAGTGGTTC	Invitrogen
pRR5C EcoRI F	Sequencing	TTCGCGCTGGTCGGCTGG	This study
pRR5C EcoRI R	Sequencing	ATGCCGACGGATTTGCACTG	This study
pGEM EcoRI R	Sequencing	ATGCATCCAACGCGTTGG	This study

A.2 Sequence Accession Numbers

A.2.1 RpoB Final Tree

WP_010666061.1_Acetobacter aceti
WP_048838093.1_Acetobacter cibinongensis
WP_048860015.1_Acidisphaera rubrifaciens
WP_075221078.1_Acuticoccus yangtzensis
RJT22700.1_Aestuariisphingobium litorale
WP_034486688.1_Agrobacterium rhizogenes
WP_054000503.1_Ahrensia marina
WP_047806934.1_Altererythrobacter marensis
WP_044143222.1_Anaplasma phagocytophilum
WP_018080321.1_Asticcacaulis benevestitus
WP_090642704.1_Asticcacaulis taihuensis
WP_163044881.1_Aurantimonas aggregata
WP_041843526.1_Bartonella bacilliformis
WP_082250848.1_Bartonella henselae
WP_014924148.1_Bartonella quintana
WP_061876887.1_Bradyrhizobium liaoningense
WP_057843569.1_Bradyrhizobium retamae
WP_029087435.1_Brevundimonas aveniformis
CUW50851.1_Brucella vulpis
WP_006996805.1_Candidatus Pelagibacter ubique
WP_035049873.1_Caulobacter henricii
WP_099620733.1_Caulobacter mirabilis
AVQ01060.1_Caulobacter segnis
WP_058350135.1_Caulobacter vibrioides
WP_012176950.1_Dinoroseobacter shibae
WP_011304298.1_Ehrlichia canis
WP_045170818.1_Ehrlichia minasensis
WP_011415806.1_Erythrobacter litoralis
WP_114727485.1_Gluconacetobacter liquefaciens
WP_008852414.1_Gluconobacter morbifer

WP_041242769.1_Gluconobacter oxydans
WP_092647930.1_Jannaschia faecimaris
WP_046510127.1_Kiloniella litopenaei
WP_085900485.1_Kiloniella majae
WP_047766118.1_Kiloniella spongiae
WP_110555357.1_Komagataeibacter swingsii
WP_025439071.1_Komagataeibacter xylinus
WP_020401500.1_Kordiimonas gwangyangensis
WP_068307855.1_Kordiimonas lacus
WP_068148674.1_Kordiimonas lipolytica
WP_055678485.1_Labrenzia alba
WP_141891412.1_Leisingera aquaemixtae
WP_089905567.1_Loktanella fryxellensis
MAF31044.1_Magnetococcales bacterium
WP_011712519.1_Magnetococcus marinus
WP_008620992.1_Magnetospirillum caucaseum
WP_041039820.1_Magnetospirillum magnetotacticum
WP_068503590.1_Magnetospirillum moscoviense
WP_069958751.1_Magnetovibrio blakemorei
WP_167636870.1_Marivivens donghaensis
PZO75651.1_Mesorhizobium amorphae
WP_106722469.1_Mesorhizobium soli
WP_102866716.1_Nesiotobacter exalbescens
WP_011510042.1_Nitrobacter hamburgensis
WP_067734696.1_Novosphingobium naphthalenivorans
WP_008069407.1_Novosphingobium nitrogenifigens
WP_085883568.1_Oceanibacterium hippocampi
WP_007255242.1_Oceanicola granulosus
WP_011747092.1_Paracoccus denitrificans
WP_116234607.1_Parasphingopyxis lamellibrachiae
WP_013300906.1_Parvularcula bermudensis
WP_127071683.1_Pelagibacterium lentulum

WP_090600311.1_Pelagibacterium luteolum
WP_111516535.1_Phenylobacterium deserti
WP_091738971.1_Phenylobacterium immobile
WP_108174373.1_Phreatobacter oligotrophus
WP_136964110.1_Phreatobacter stygius
WP_135244427.1_Polymorphobacter arshaanensis
WP_152578190.1_Polymorphobacter fuscus
WP_013652475.1_Polymorphum gilvum
WP_068314117.1_Pseudovibrio hongkongensis
WP_068082285.1_Pseudovibrio stylochi
WP_120703913.1_Rhizobium jaguaris
WP_085423131.1_Rhizobium oryzae
WP_136906555.1_Rhodobacter capsulatus
WP_119004049.1_Rhodobacter sphaeroides
WP_011158811.1_Rhodopseudomonas palustris
WP_012565924.1_Rhodospirillum centenum
WP_014626434.1_Rhodospirillum rubrum
WP_027289410.1_Rhodovibrio salinarum
WP_045798722.1_Rickettsia bellii
WP_014120533.1_Rickettsia japonica
WP_012150397.1_Rickettsia rickettsii
WP_004996628.1_Rickettsia sibirica
WP_013960393.1_Roseobacter litoralis
WP_099097358.1_Roseomonas rhizosphaerae
WP_009812405.1_Roseovarius nubinhibens
WP_094474600.1_Sandarakinorhabdus cyanobacteriorum
WP_156874416.1_Sandarakinorhabdus limnophila
WP_169562178.1_Sneathiella chinensis
WP_161339460.1_Sneathiella chungangensis
WP_149680933.1_Sphingomonas carotinifaciens
WP_088367960.1_Sphingomonas dokdonensis
WP_163958602.1_Sphingomonas insulae

WP_089215018.1_Sphingopyxis indica

WP_037911444.1_Sulfitobacter mediterraneus

WP_028880764.1_Terasakiella pusilla

WP_014848589.1_Zymomonas mobilis

A.2.2 CtrA sequences from representatives of each order in Alphaproteobacteria

WP_169568121.1_Sneathiella limimaris

WP_169559802.1_Sneathiella chinensis

WP_167638795.1_Marivivens donghaensis

WP_163959263.1_Sphingomonas insulae

WP_163045434.1_Aurantimonas aggregata

WP_161339525.1_Sneathiella chungangensis

WP_152576298.1_Polymorphobacter fuscus

WP_141920062.1_Zymomonas mobilis

WP_141890204.1_Leisingera aquaemixtae

WP_136958952_Phreatobacter stygius

WP_135245133.1_Polymorphobacter arshaanensis

WP_127073694.1_Pelagibacterium lentulum

WP_123155058_Aestuariisphingobium litorale

WP_121937372.1_Eilatimonas milleporae

WP_120705148.1_Rhizobium jaguaris

WP_117125908.1_Komagataeibacter xylinus

WP_116235910.1_Parasphingopyxis lamellibrachiae

WP_114726701.1_Gluconacetobacter liquefaciens

WP_112383408.1_Sphingomonas carotinifaciens

WP_111512978.1_Phenylobacterium deserti

WP_110557183.1_Komagataeibacter swingsii

WP_106727286.1_Mesorhizobium soli

WP_099621086.1_Caulobacter mirabilis

WP_099095590.1_Roseomonas rhizosphaerae

WP_096372719.1_Rickettsia japonica

WP_094473039_Sandarakinorhabdus cyanobacteriorum

WP_092644234.1_Jannaschia faecimaris
WP_091741635.1_Phenylobacterium immobile
WP_090647334.1_Asticcacaulis taihuensis
WP_090594763.1_Pelagibacterium luteolum
WP_089897464.1_Loktanella fryxellensis
WP_089215420.1_Sphingopyxis indica
WP_088365599.1_Sphingomonas dokdonensis
WP_085900876.1_Kiloniella majae
WP_075625826.1_Rhizobium oryzae
WP_075222955_Acuticoccus yangtzensis
WP_069957282.1_Magnetovibrio blakemorei
WP_068498488.1_Magnetospirillum moscoviense
WP_068313551_Pseudovibrio hongkongensis
WP_068150049.1_Kordiimonas lipolytica
WP_068087222_Pseudovibrio stylochi
WP_067732289.1_Novosphingobium naphthalenivorans
WP_057842910.1_Bradyrhizobium retamae
WP_055677578_Labrenzia alba
WP_055208583.1_Rhodobacter capsulatus
WP_053999890_Ahrensia marina
WP_050994895_Bradyrhizobium liaoningense
WP_048862760.1_Acidisphaera rubrifaciens
WP_048838918.1_Acetobacter cibinongensis
WP_047765977.1_Kiloniella spongiae
WP_046507927.1_Kiloniella litopenaei
WP_045895782.1_Anaplasma phagocytophilum
WP_045799065.1_Rickettsia bellii
WP_045170889.1_Ehrlichia minasensis
WP_041375382_Polymorphum gilvum
WP_041041883.1_Magnetospirillum magnetotacticum
WP_035042898.1_Caulobacter henricii
WP_034906536.1_Erythrobacter litoralis

WP_034476386.1_Agrobacterium rhizogenes
WP_029087941.1_Brevundimonas aveniformis
WP_028879163_Terasakiella pusilla
WP_028481094_Nesiotobacter exalbescens
WP_027289079.1_Rhodovibrio salinarum
WP_025898305.1_Sneathiella glossodoripedis
WP_025048148.1_Sulfitobacter mediterraneus
WP_023447480.1_Asticcacaulis benevestitus
WP_022679779.1_Sandarakinorhabdus limnophila
WP_020592224.1_Kiloniella laminariae
WP_020399773.1_Kordiimonas gwangyangensis
WP_016769841.1_Rickettsia sibirica
WP_014362718.1_Rickettsia rickettsii
WP_013961871.1_Roseobacter litoralis
WP_013300833.1_Parvularcula bermudensis
WP_012178180.1_Dinoroseobacter shibae
WP_011511754.1_Nitrobacter hamburgensis
WP_011388283.1_Rhodospirillum rubrum
WP_011304925.1_Ehrlichia canis
WP_011253215.1_Gluconobacter oxydans
WP_011181041.1_Bartonella henselae
WP_010667799.1_Acetobacter aceti
WP_008850534.1_Gluconobacter morbifer
WP_008616069.1_Magnetospirillum caucaseum
WP_008070048.1_Novosphingobium nitrogenifigens
WP_007254149.1_Oceanicola granulosus
WP_004622056.1_Caulobacter vibrioides
SLN69662.1_Oceanibacterium hippocampi
SDD94010.1_Kordiimonas lacus
PZR31394.1_Caulobacter segnis
PZQ99575.1_Rhodobacter sphaeroides
PZO77913.1_Mesorhizobium amorphae

PZO68350.1_Altererythrobacter marensis

PZO65986.1_Paracoccus denitrificans

PTM53524_Phreatobacter oligotrophus

HAR52580.1_Roseovarius nubinhibens

EYS89622.1_Bartonella bacilliformis San Pedro600-02

CUW43955.1_Brucella vulpis

CAE27073.1_Rhodopseudomonas palustris CGA009

AFR26609.1_Bartonella quintana RM-11

ACI99149.1_Rhodospirillum centenum SW

A.3 Overlap Extension PCR Cloning

BASED ON BRYSKIN & MATSUMURA (2013), CHAPTER 4, SYNTHETIC BIOLOGY, METHODS IN MOLECULAR BIOLOGY, VOL. 1073 AND BRYSKIN & MATSUMURA (2010), OVERLAP EXTENSION PCR CLONING: A SIMPLE AND RELIABLE WAY TO CREATE RECOMBINANT PLASMIDS, BIOTECHNIQUES 48:463-465

PROTOCOL BY ROSHNI KOLLIPARA

Basic Methodology

First, we amplify the insert using normal PCR. The primers we use to amplify the insert have tails that match your recipient plasmid.

Next, we use this PCR product (the amplified insert) as megaprimers for a second PCR in which we amplify the recipient plasmid.

The final product is the recipient plasmid containing the insert.

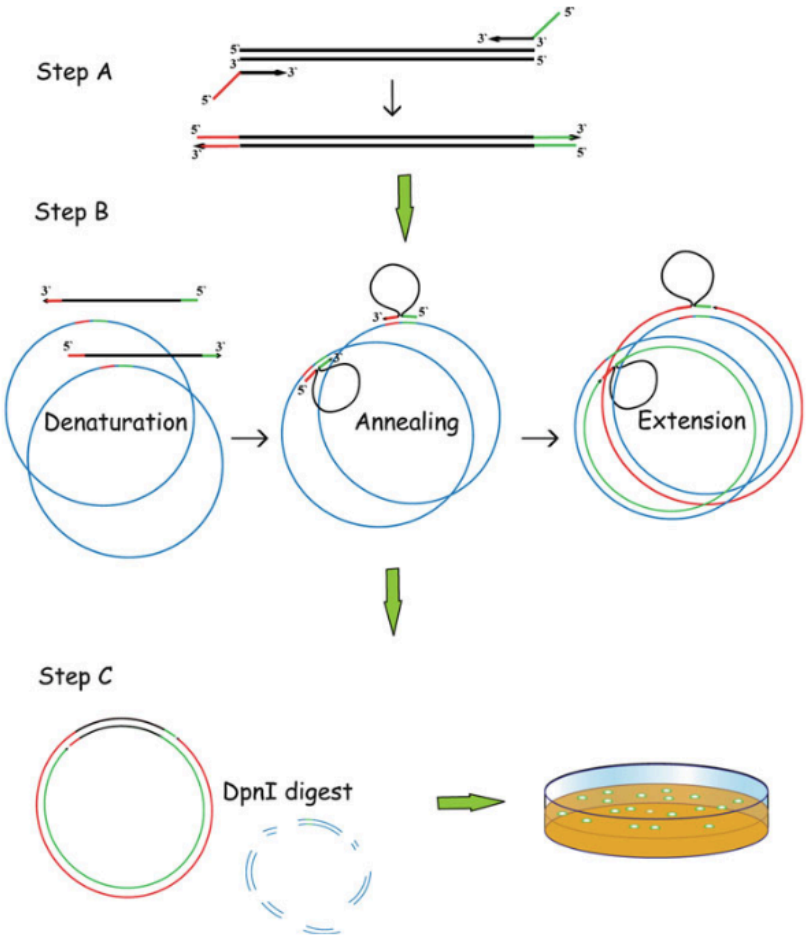


Figure 1. From Bryskin & Matsumura 2013. (Step A) The insert is PCR amplified with the chimeric primers. The ends of final PCR product overlap regions of the vector. (Step B) Vector and insert are mixed, denatured, and annealed. The hybridized insert is then extended by Phusion® polymerase using the vector as a template until polymerase reaches the 50-end of the insert. After several PCR cycles, the new plasmid with two nicks (one on each strand) accumulates as a product. (Step C) The parental plasmid is destroyed by DpnI digest, and new plasmid is used to transform *E. coli*.

Designing Primers

This method uses two primers, each one matches a section of the insert and a section of the recipient vector. Designing them in four parts is easiest – primers A and B will be on the insert, and primers C and D will be on the recipient plasmid.

1. Design appropriate primers A and B to PCR amplify the insert (like normal). A will be forward and B will be reverse, aim for a Tm of 65°C. Note that the Tm of both A and B should be as close as possible.
2. Select where you want your insert to be on your recipient plasmid. The two points of insertion should preferably be 50 to several hundred base pairs apart. The sequence between the two points of insertion will be lost.
3. Then select 30–40 bp upstream of the left point of insertion on the top strand of the plasmid. Copy this sequence and estimate its Tm using an online tool. If the Tm is 65-68°C or higher, save the sequence as primer C.
4. Next, select 30–40 bp downstream of the right point of insertion on the bottom strand of the plasmid. Copy the reverse sequence and analyze its Tm using an online tool (Oligo Analyzer); if the Tm is 65-68°C or higher, save the sequence as primer D. Note that the Tm of both C and D should be as close as possible.
5. Attach the sequence of primer C to the 5' end of primer A. Attach the sequence of primer D to the 5' end of primer B.

PCR Part A

Set-up multiple (at least 10) PCR reactions per insert. These must be pooled in order to have enough material for Part B.

Water	up to 20 uL*
Forward primer (2uM)	3 uL
Reverse primer (2uM)	3 uL
DreamTaq Green 2X	10 uL
Template	10 ng

*the best yield occurred with 20 uL reaction volume, do not exceed 40 uL/reaction

98°C 30s

98°C 7s

60°C* 15s } 35 cycles

72°C 30s

72°C 7 min

*this method did not seem to be temperature sensitive over a 10°C range

Run 2 uL of the PCR reaction on a 0.8% agarose gel. If a single product is confirmed with faint primer dimers, pool the reactions and clean with AmPure beads (elute ~50 uL). DreamTaq adds an A to the end of the PCR product so a digestion with T4 DNA Polymerase is necessary before proceeding to the next step.

Water	up to 40 uL
DNA	3 ug
dNTPs (10mM)	10 uL
NEBuffer 2.1	4 uL
T4 DNA polymerase	1 uL

12°C for 15 min, add EDTA to a final concentration of 10mM, 75°C for 20 min.
Clean with AmPure beads, elute in 10 uL.

PCR Part B

Some optimization may be required regarding the amount of megaprimers and Phusion added as well as the PCR method itself.
For an insert size of 0.9 kb and a recipient plasmid size of 7.3 kb, the following method worked.

Water	up to 30 uL
5X HF Buffer	6 uL
dNTPs (10 mM)	0.6 uL
DMSO	0.9 uL
megaprimers	250 ng

template	3.25 ng
Phusion	0.6 uL

98°C	1 min	
98°C	7s	25 cycles
60°C	20s	
72°C	1 min 20s	
72°C	7 min	

Add 1 uL DpnI to each reaction, incubate at 37°C for 1 hour. Transform 1 uL of the reaction into NEB5alpha competent cells.
Screen and sequence many colonies. Efficiency of recombination is quite high however only some will not have mutations.

A.4 Phylogenetic tree

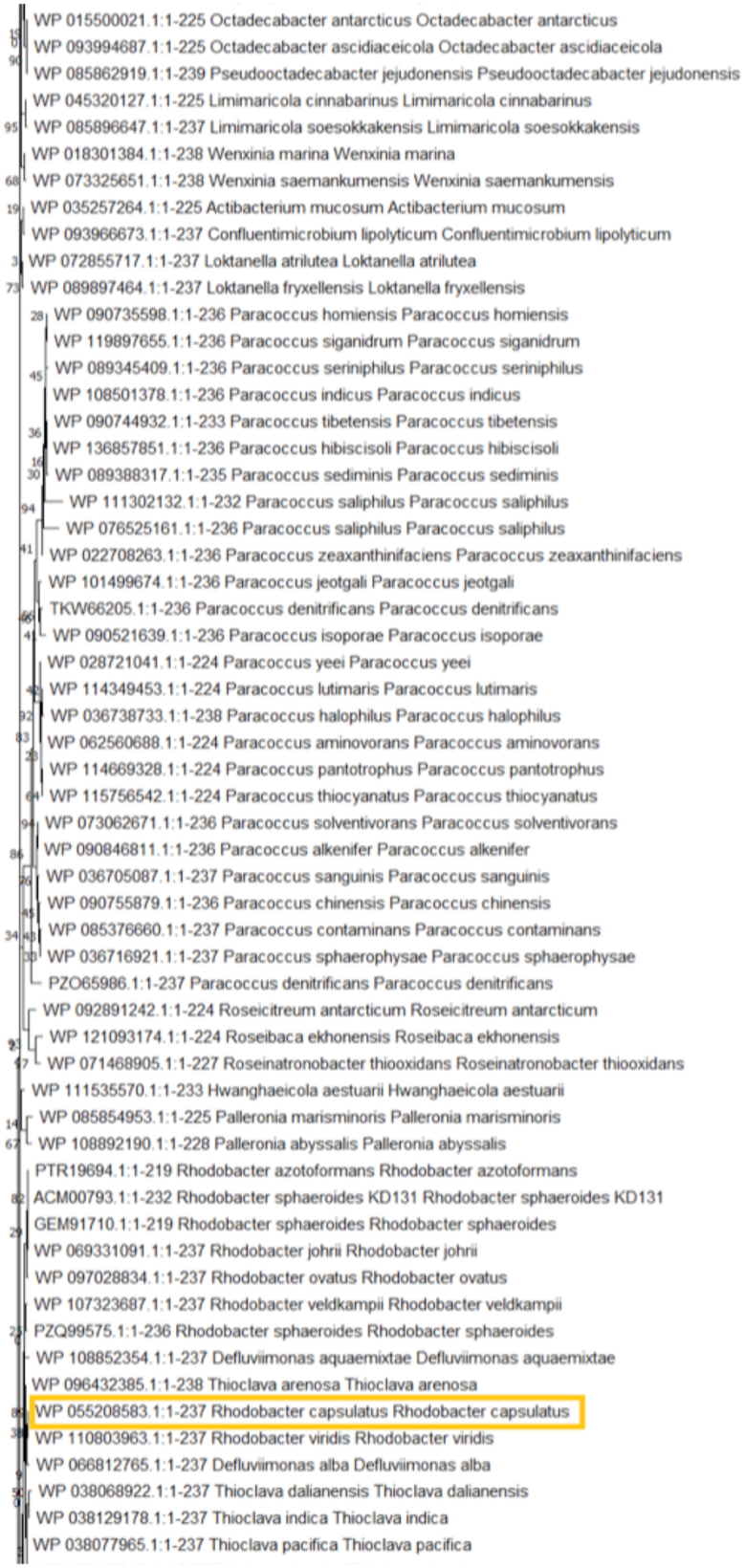

WP 027314840.1:1-228 Microvirga flocculans Microvirga flocculans
WP 091129617.1:1-228 Microvirga guangxiensis Microvirga guangxiensis
WP 114187467.1:1-228 Microvirga aerophila Microvirga aerophila
WP 114770484.1:1-233 Microvirga subterranea Microvirga subterranea
WP 046867319.1:1-228 Microvirga massiliensis Microvirga massiliensis
WP 029030502.1:1-228 Salinarimonas rosea Salinarimonas rosea
WP 023433916.1:1-226 Lutibaculum baratangense Lutibaculum baratangense
WP 026380392.1:1-229 Afifella pfennigii Afifella pfennigii
WP 069095031.1:1-226 Methyloligella halotolerans Methyloligella halotolerans
KWT64268.1:1-224 Hyphomicrobium sulfonivorans Hyphomicrobium sulfonivorans
WP 068465321.1:1-228 Hyphomicrobium sulfonivorans Hyphomicrobium sulfonivorans
WP 020086750.1:1-224 Hyphomicrobium zavarzinii Hyphomicrobium zavarzinii
WP 023787190.1:1-225 Hyphomicrobium nitrativorans Hyphomicrobium nitrativorans
WP 068313551.1:1-226 Pseudovibrio hongkongensis Pseudovibrio hongkongensis
WP 128775552.1:1-226 Hansschlegelia zhihuaiae Hansschlegelia zhihuaiae
WP 127186829.1:1-228 Arsenicitalea aurantiaca Arsenicitalea aurantiaca
WP 090594763.1:1-229 Pelagibacterium luteolum Pelagibacterium luteolum
WP 046107795.1:1-229 Devosia geojensis Devosia geojensis
WP 127073694.1:1-229 Pelagibacterium lentulum Pelagibacterium lentulum
WP 029040369.1:1-229 Cucumibacter marinus Cucumibacter marinus
WP 027834926.1:1-229 Maritalea myrionectae Maritalea myrionectae
WP 133571938.1:1-229 Maritalea mobilis Maritalea mobilis
WP 046169807.1:1-225 Devosia psychrophila Devosia psychrophila
WP 046133755.1:1-226 Devosia limi Devosia limi
WP 046140095.1:1-228 Devosia epidermidihirudinis Devosia epidermidihirudinis
WP 108397876.1:1-229 Devosia submarina Devosia submarina
WP 072340850.1:1-229 Devosia enhydra Devosia enhydra
WP 069909042.1:1-228 Devosia insulae Devosia insulae
WP 046105482.1:1-228 Devosia chinhatensis Devosia chinhatensis
WP 067457628.1:1-229 Devosia elaeis Devosia elaeis
WP 092419838.1:1-229 Devosia crocina Devosia crocina
WP 075222955.1:1-225 Acuticoccus yangtzensis Acuticoccus yangtzensis
WP 018700803.1:1-226 Amorphus coralli Amorphus coralli
WP 073629918.1:1-223 Pseudoxanthobacter soli Pseudoxanthobacter soli
WP 099557386.1:1-225 Hartmannibacter diazotrophicus Hartmannibacter diazotrophicus
WP 126539041.1:1-225 Oharaeibacter diazotrophicus Oharaeibacter diazotrophicus
WP 075215318.1:1-225 Mongoliimonas terrestris Mongoliimonas terrestris
WP 026784327.1:1-226 Pleomorphomonas koreensis Pleomorphomonas koreensis
WP 028481094.1:1-226 Nesiotobacter exalbescens Nesiotobacter exalbescens
WP 068087222.1:1-226 Pseudovibrio stylochi Pseudovibrio stylochi
WP 041375382.1:1-225 Polymorphum gilvum Polymorphum gilvum
WP 055677578.1:1-225 Labrenzia alba Labrenzia alba
WP 054360571.1:1-229 Prosthecomicrobium hirschii Prosthecomicrobium hirschii
WP 107990695.1:1-226 Breoghania corrubedonensis Breoghania corrubedonensis
WP 119309456.1:1-226 Cohaesibacter haloalkalitolerans Cohaesibacter haloalkalitolerans
WP 097155391.1:1-229 Cohaesibacter gelatinilyticus Cohaesibacter gelatinilyticus
WP 090073375.1:1-226 Cohaesibacter marisflavi Cohaesibacter marisflavi
WP 111436085.1:1-226 Rhodobium orientis Rhodobium orientis
SHG76900.1:21-246 Kaistia soli DSM 19436 Kaistia soli DSM 19436
WP 015273449.1:1-228 Liberibacter crescens Liberibacter crescens
WP 084412801.1:1-227 Fulvimarina manganoxydans Fulvimarina manganoxydans
WP 134764185.1:1-226 Jiella endophytica Jiella endophytica
HDZ75231.1:1-227 Aurantimonas coralicida Aurantimonas coralicida
WP 163045434.1:1-227 Aurantimonas aggregata Aurantimonas aggregata
WP 007066082.1:1-226 Fulvimarina pelagi Fulvimarina pelagi
WP 116681673.1:1-226 Fulvimarina endophytica Fulvimarina endophytica
WP 078708460.1:1-226 Consotaella salsifontis Consotaella salsifontis
WP 039189067.1:1-225 Aureimonas altamirensis Aureimonas altamirensis

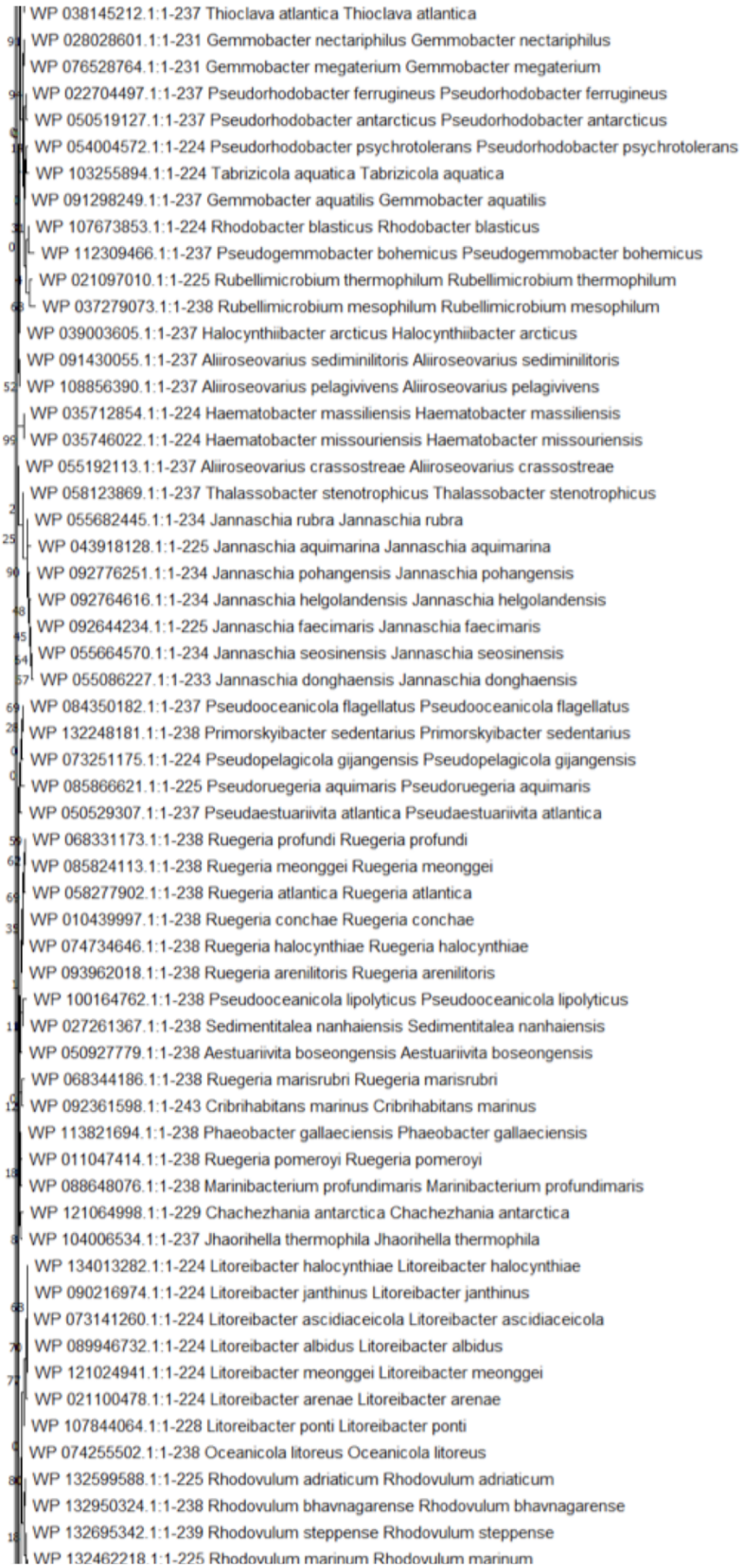

WP 062227404.1:1-225 Aureimonas frigidaquae Aureimonas frigidaquae
WP 027488658.1:1-225 Rhizobium undicola Rhizobium undicola
WP 146932962.1:1-228 Agrobacterium vitis Agrobacterium vitis
WP 105383392.1:1-228 Neorhizobium alkalisoli Neorhizobium alkalisoli
EKJ93400.1:1-225 Bradyrhizobium lupini HPC(L) Bradyrhizobium lupini HPC(L)
WP 027510136.1:1-225 Rhizobium sullae Rhizobium sullae
WP 120705148.1:1-229 Rhizobium jaguaris Rhizobium jaguaris
WP 066877537.1:1-228 Sinorhizobium saheli Sinorhizobium saheli
WP 097137390.1:1-225 Rhizobium subbaraonis Rhizobium subbaraonis
WP 108003612.1:1-229 Mycoplana dimorpha Mycoplana dimorpha
WP 075625826.1:1-228 Rhizobium oryzae Rhizobium oryzae
WP 018011827.1:1-225 Sinorhizobium medicae Sinorhizobium medicae
WP 034476386.1:1-229 Agrobacterium rhizogenes Agrobacterium rhizogenes
WP 119256748.1:1-228 Shinella zoogloeoides Shinella zoogloeoides
WP 111160890.1:1-225 Rhizobium tubonense Rhizobium tubonense
WP 133285425.1:1-229 Pseudohoeflea suaedae Pseudohoeflea suaedae
WP 007198768.1:1-229 Hoeflea phototrophica Hoeflea phototrophica
WP 110034415.1:1-229 Hoeflea marina Hoeflea marina
WP 099865361.1:1-229 Pararhizobium haloflavum Pararhizobium haloflavum
WP 039723750.1:1-225 Roseitalea porphyridii Roseitalea porphyridii
WP 053999890.1:1-229 Ahrensia marina Ahrensia marina
WP 131567405.1:1-226 Oricola cellulosilytica Oricola cellulosilytica
WP 105731966.1:1-229 Phyllobacterium myrsinacearum Phyllobacterium myrsinacearum
WP 106718609.1:1-229 Phyllobacterium endophyticum Phyllobacterium endophyticum
WP 114430330.1:1-229 Phyllobacterium bourgognense Phyllobacterium bourgognense
WP 106714380.1:1-229 Phyllobacterium brassicacearum Phyllobacterium brassicacearum
WP 077993064.1:1-229 Bartonella apis Bartonella apis
WP 119042654.1:1-225 Ochrobactrum haematophilum Ochrobactrum haematophilum
WP 094505307.1:1-229 Ochrobactrum thiophenivorans Ochrobactrum thiophenivorans
CUW43955.1:2-226 Brucella vulpis Brucella vulpis
WP 111574279.1:1-229 Falsochrobactrum ovis Falsochrobactrum ovis
WP 110752911.1:1-225 Phyllobacterium leguminum Phyllobacterium leguminum
WP 009449858.1:1-226 Nitratireductor indicus Nitratireductor indicus
WP 091524239.1:1-226 Aquamicrobium aerolatum Aquamicrobium aerolatum
WP 094077414.1:1-226 Notoacmeibacter marinus Notoacmeibacter marinus
WP 121645342.1:1-226 Notoacmeibacter ruber Notoacmeibacter ruber
WP 072602935.1:1-225 Mesorhizobium oceanicum Mesorhizobium oceanicum
WP 106727286.1:1-226 Mesorhizobium soli Mesorhizobium soli
WP 019173841.1:1-226 Pseudaminobacter salicylatoxidans Pseudaminobacter salicylatoxidans
WP 080920723.1:1-225 Pseudaminobacter manganicus Pseudaminobacter manganicus
PZO77913.1:1-226 Mesorhizobium amorphae Mesorhizobium amorphae
WP 091594645.1:1-225 Mesorhizobium muleiense Mesorhizobium muleiense
WP 123149336.1:1-225 Mesorhizobium delmotii Mesorhizobium delmotii
WP 036479330.1:1-226 Nitratireductor basaltis Nitratireductor basaltis
WP 035028413.1:1-225 Aquamicrobium defluvii Aquamicrobium defluvii
WP 105740573.1:1-229 Phyllobacterium phragmitis Phyllobacterium phragmitis
WP 114438686.1:1-229 Phyllobacterium salinisoli Phyllobacterium salinisoli
WP 045684182.1:1-232 Martelella endophytica Martelella endophytica
WP 106309770.1:1-232 Martelella mediterranea Martelella mediterranea
WP 015398284.1:1-224 Bartonella australis Bartonella australis
WP 013544644.1:1-224 Bartonella clarridgeiae Bartonella clarridgeiae
WP 010702718.1:1-231 Bartonella bovis Bartonella bovis
EYS89622.1:14-237 Bartonella bacilliformis San Pedro600-02 Bartonella bacilliformis San Pedro600-02
WP 019222331.1:1-224 Bartonella rattaustraliani Bartonella rattaustraliani
AFR26609.1:15-238 Bartonella quintana RM-11 Bartonella quintana RM-11
WP 004859645.1:1-224 Bartonella taylorii Bartonella taylorii
WP 007347872.1:1-224 Bartonella rattimassiliensis Bartonella rattimassiliensis
WP 100130280.1:1-224 Bartonella tribocorum Bartonella tribocorum

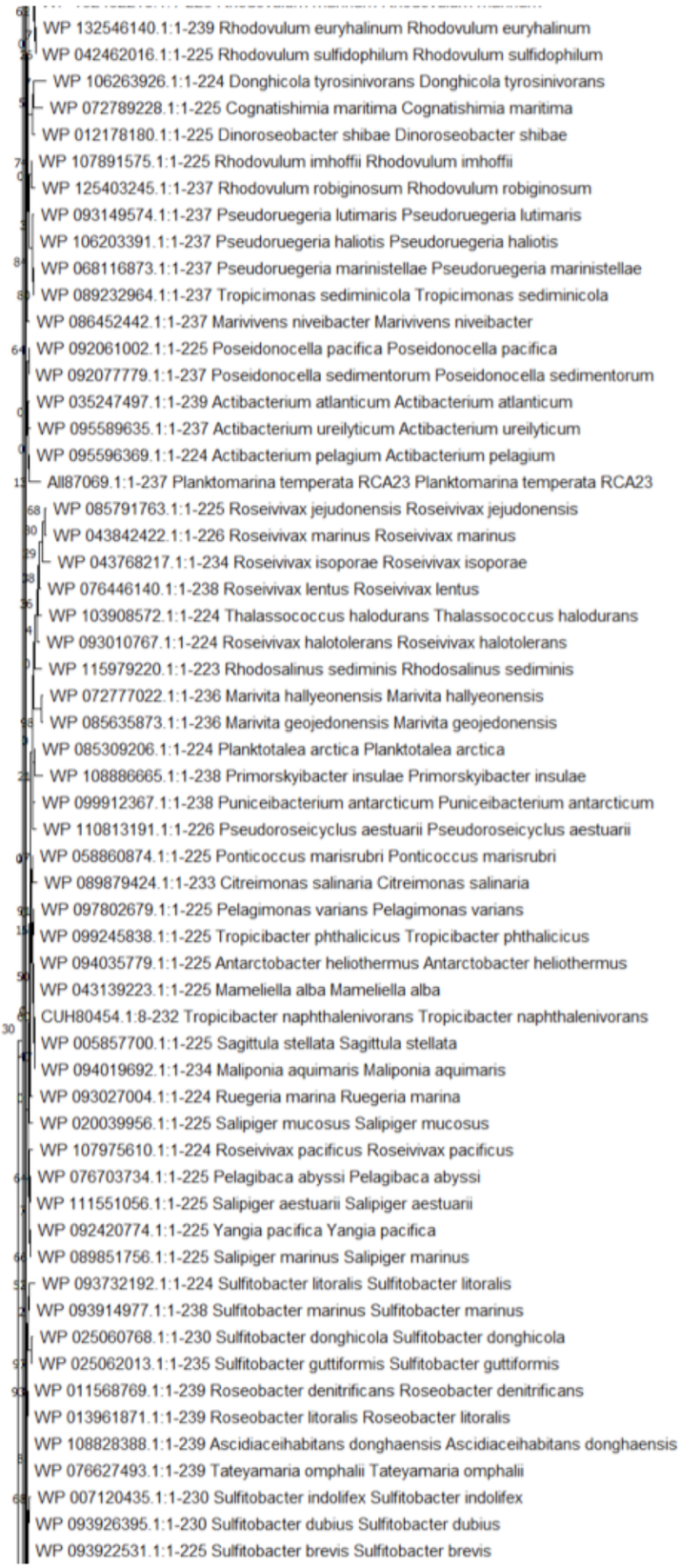

WP 005864769.1:1-224 Bartonella alsatica Bartonella alsatica
WP 006590160.1:1-231 Bartonella birtlesii Bartonella birtlesii
WP 019218966.1:1-231 Bartonella florencae Bartonella florencae
WP 010705499.1:1-224 Bartonella vinsonii Bartonella vinsonii
WP 011181041.1:1-224 Bartonella henselae Bartonella henselae
WP 034460287.1:1-224 Bartonella koehlerae Bartonella koehlerae
WP 034989129.1:1-231 Bartonella senegalensis Bartonella senegalensis
WP 063988126.1:1-224 Methylorubrum populi Methylorubrum populi
WP 119060436.1:1-224 Dichotomicrobium thermohalophilum Dichotomicrobium thermohalophilum
WP 012111093.1:1-229 Parvibaculum lavamentivorans Parvibaculum lavamentivorans
WP 013300833.1:1-226 Parvularcula bermudensis Parvularcula bermudensis
WP 017932262.1:1-224 Robiginitomaculum antarcticum Robiginitomaculum antarcticum
WP 026942254.1:1-224 Hellea balneolensis Hellea balneolensis
WP 022697208.1:1-227 Euryhalocaulis caribicus Euryhalocaulis caribicus
WP 011642609.1:1-224 Maricaulis maris Maricaulis maris
WP 091766370.1:1-224 Maricaulis salignorans Maricaulis salignorans
WP 135945136.1:1-226 Marinicauda pacifica Marinicauda pacifica
WP 135995941.1:1-226 Marinicauda algicola Marinicauda algicola
WP 019962281.1:1-232 Woodsholea maritima Woodsholea maritima
WP 127568283.1:1-224 Glycocaulis alkaliphilus Glycocaulis alkaliphilus
WP 050994895.1:1-224 Bradyrhizobium liaoningense Bradyrhizobium liaoningense
WP 018147981.1:1-232 Henriciella marina Henriciella marina
WP 084398142.1:1-232 Henriciella aquimarina Henriciella aquimarina
WP 084393320.1:1-232 Henriciella pelagia Henriciella pelagia
WP 034794409.1:1-224 Hyphomonas beringensis Hyphomonas beringensis
WP 015828052.1:1-224 Hirschia baltica Hirschia baltica
WP 023447480.1:1-226 Asticcacaulis benevestitus Asticcacaulis benevestitus
WP 090647334.1:1-230 Asticcacaulis taihuensis Asticcacaulis taihuensis
WP 121482329.1:1-227 Brevundimonas naejangsanensis Brevundimonas naejangsanensis
WP 123155058.1:1-227 Aestuariisphingobium litorale Aestuariisphingobium litorale
WP 114818632.1:1-227 Brevundimonas bullata Brevundimonas bullata
WP 013268628.1:1-229 Brevundimonas subvibrioides Brevundimonas subvibrioides
WP 092311341.1:1-227 Brevundimonas viscosa Brevundimonas viscosa
WP 021698001.1:1-226 Brevundimonas abyssalis Brevundimonas abyssalis
WP 029087941.1:1-226 Brevundimonas aveniformis Brevundimonas aveniformis
PZR31394.1:1-226 Caulobacter segnis Caulobacter segnis
WP 004622056.1:1-227 Caulobacter vibrioides Caulobacter vibrioides
WP 035042898.1:1-226 Caulobacter henricii Caulobacter henricii
WP 099621086.1:1-224 Caulobacter mirabilis Caulobacter mirabilis
WP 111455957.1:1-229 Phenylobacterium hankyongense Phenylobacterium hankyongense
WP 091741635.1:1-226 Phenylobacterium immobile Phenylobacterium immobile
WP 012521413.1:1-227 Phenylobacterium zucineum Phenylobacterium zucineum
WP 111512978.1:1-227 Phenylobacterium deserti Phenylobacterium deserti
WP 109922711.1:1-228 Zavarzinia compransoris Zavarzinia compransoris
WP 085884586.1:1-226 Oceanibacterium hippocampi Oceanibacterium hippocampi
WP 025898305.1:1-227 Sneathiella glossodoripedis Sneathiella glossodoripedis
WP 068150049.1:1-226 Kordiimonas lipolytica Kordiimonas lipolytica
WP 068303000.1:1-226 Kordiimonas lacus Kordiimonas lacus
WP 020399773.1:1-226 Kordiimonas gwangyangensis Kordiimonas gwangyangensis
WP 121937372.1:1-224 Eilatimonas milleporae Eilatimonas milleporae
WP 132708384.1:1-228 Rhodothalassium salexigens Rhodothalassium salexigens
WP 121047914.1:1-227 Sphingosinicella microcystinivorans Sphingosinicella microcystinivorans
EMD82167.1:8-232 Pacificimonas flava Pacificimonas flava
WP 011950580.1:1-226 Sphingomonas wittichii Sphingomonas wittichii
WP 095687359.1:1-226 Sphingobium hydrophobicum Sphingobium hydrophobicum
WP 141920062.1:1-232 Zymomonas mobilis Zymomonas mobilis
WP 072595494.1:1-227 Tardibacter chloracetimidivorans Tardibacter chloracetimidivorans
WP 012051027.1:1-229 Sphingomonas wittichii Sphingomonas wittichii

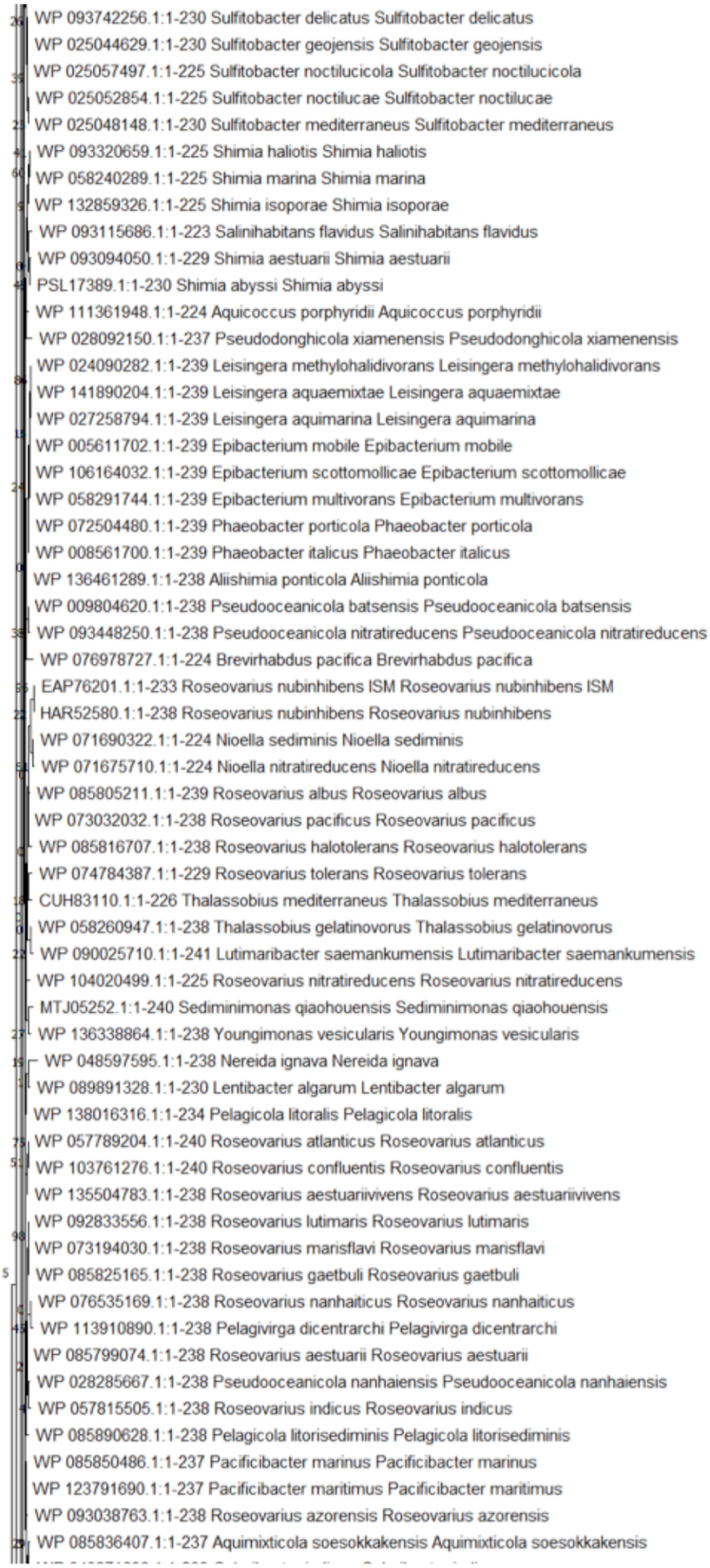

WP 066958925.1:1-229 Rhizorhabdus dicambivorans Rhizorhabdus dicambivorans
WP 076073160.1:1-231 Sphingomonas montana Sphingomonas montana
WP 093315344.1:1-231 Sphingomonas jatrophae Sphingomonas jatrophae
WP 085217952.1:1-226 Sphingomonas indica Sphingomonas indica
WP 037503850.1:1-225 Sphingomonas jaspsi Sphingomonas jaspsi
WP 029941205.1:1-228 Sphingomonas astaxanthinifaciens Sphingomonas astaxanthinifaciens
WP 089220731.1:1-225 Sphingomonas laterariae Sphingomonas laterariae
WP 057884566.1:1-226 Altererythrobacter troitsensis Altererythrobacter troitsensis
WP 116235910.1:1-226 Parasphingopyxis lamellibrachiae Parasphingopyxis lamellibrachiae
WP 125230727.1:1-225 Sphingorhabdus wooponensis Sphingorhabdus wooponensis
WP 069205166.1:1-225 Sphingomonas panacis Sphingomonas panacis
WP 025293910.1:1-225 Sphingomonas sanxanigenens Sphingomonas sanxanigenens
WP 058754888.1:1-227 Sphingomonas endophytica Sphingomonas endophytica
WP 066709257.1:1-226 Sphingomonas adhaesiva Sphingomonas adhaesiva
WP 022687090.1:1-226 Sphingomonas phyllosphaerae Sphingomonas phyllosphaerae
WP 119037574.1:1-234 Hephaestia caeni Hephaestia caeni
WP 066574571.1:1-227 Sphingomonas koreensis Sphingomonas koreensis
WP 135963236.1:1-224 Sphingomonas gei Sphingomonas gei
WP 135986875.1:1-225 Sphingomonas naasensis Sphingomonas naasensis
WP 066797183.1:1-228 Sphingomonas soli Sphingomonas soli
WP 044333454.1:1-223 Sphingomonas hengshuiensis Sphingomonas hengshuiensis
WP 097064441.1:1-227 Sphingomonas guangdongensis Sphingomonas guangdongensis
WP 010403489.1:1-226 Sphingomonas echinoides Sphingomonas echinoides
WP 093334083.1:1-230 Sphingomonas rubra Sphingomonas rubra
WP 088365599.1:1-226 Sphingomonas dokdonensis Sphingomonas dokdonensis
WP 076745448.1:1-226 Sphingomonas jeddahensis Sphingomonas jeddahensis
WP 088334626.1:1-226 Sphingomonas mucosissima Sphingomonas mucosissima
WP 112383408.1:1-226 Sphingomonas carotinifaciens Sphingomonas carotinifaciens
WP 058717908.1:1-226 Sphingomonas sanguinis Sphingomonas sanguinis
WP 163959263.1:1-226 Sphingomonas insulae Sphingomonas insulae
WP 066805668.1:1-226 Sphingomonas asaccharolytica Sphingomonas asaccharolytica
WP 066827241.1:1-226 Sphingomonas mali Sphingomonas mali
WP 096341451.1:1-226 Sphingomonas spermidinifaciens Sphingomonas spermidinifaciens
WP 109808644.1:1-229 Sphingosinithalassobacter portus Sphingosinithalassobacter portus
WP 129340665.1:1-226 Sphingomonas desiccabilis Sphingomonas desiccabilis
WP 088472792.1:1-234 Sphingopyxis witflariensis Sphingopyxis witflariensis
WP 089215420.1:1-234 Sphingopyxis indica Sphingopyxis indica
WP 021689354.1:1-225 Novosphingobium tardaugens Novosphingobium tardaugens
WP 133496217.1:1-234 Stakelama pacifica Stakelama pacifica
WP 121118800.1:1-223 Croceibacterium ferulae Croceibacterium ferulae
WP 039097367.1:1-229 Croceibacterium mercuriale Croceibacterium mercuriale
WP 046349406.1:1-226 Sphingomonas changbaiensis Sphingomonas changbaiensis
WP 067732289.1:1-226 Novosphingobium naphthalenivorans Novosphingobium naphthalenivorans
WP 058735574.1:1-229 Novosphingobium barchaimii Novosphingobium barchaimii
WP 008070048.1:1-225 Novosphingobium nitrogenifigens Novosphingobium nitrogenifigens
WP 028641381.1:1-226 Novosphingobium acidiphilum Novosphingobium acidiphilum
WP 068078585.1:1-225 Novosphingobium lentum Novosphingobium lentum
WP 110299461.1:1-230 Blastomonas natatoria Blastomonas natatoria
WP 068092971.1:1-227 Novosphingobium rosa Novosphingobium rosa
WP 100868079.1:1-225 Novosphingobium kunmingense Novosphingobium kunmingense
PZO68350.1:1-226 Altererythrobacter marensis Altererythrobacter marensis
WP 066846804.1:1-224 Croceicoccus marinus Croceicoccus marinus
WP 070155504.1:1-226 Sphingobium phenoxybenzoativorans Sphingobium phenoxybenzoativorans
WP 066699050.1:1-226 Sphingobium amiense Sphingobium amiense
WP 021319599.1:1-228 Sphingobium ummariense Sphingobium ummariense
WP 062065116.1:1-226 Sphingobium baderi Sphingobium baderi
WP 037488316.1:1-227 Sphingobium japonicum Sphingobium japonicum
WP 020821006.1:1-222 Sphingobium chinhatense Sphingobium chinhatense

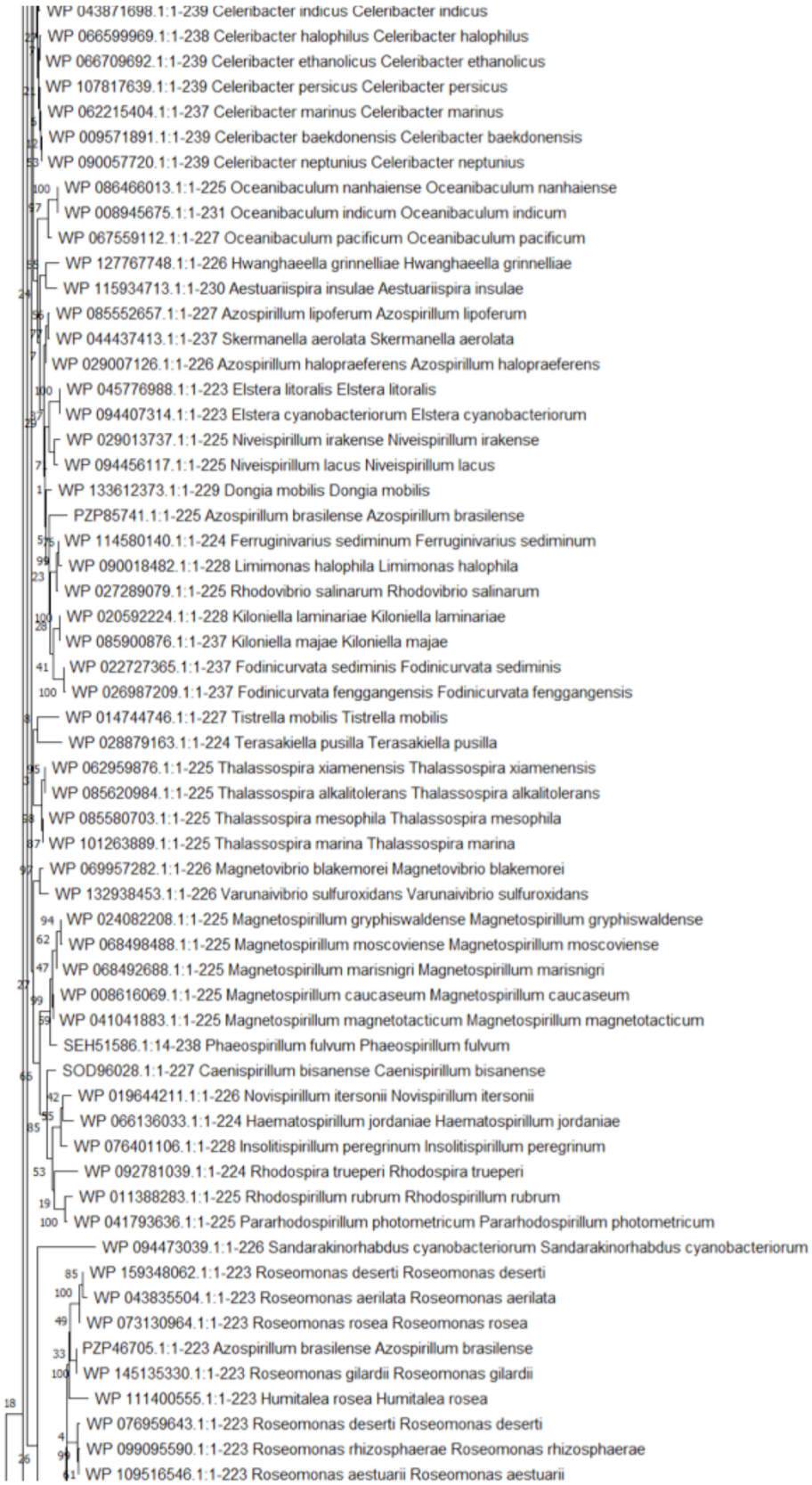

WP 072674154.1:1-226 Porphyrobacter sanguineus Porphyrobacter sanguineus
WP 027442434.1:1-226 Porphyrobacter cryptus Porphyrobacter cryptus
WP 067675125.1:1-224 Altererythrobacter dongtanensis Altererythrobacter dongtanensis
WP 126177351.1:1-226 Altererythrobacter rigui Altererythrobacter rigui
3
WP 133366001.1:1-226 Qipengyuania sediminis Qipengyuania sediminis
WP 034962326.1:1-226 Erythrobacter longus Erythrobacter longus
21
WP 067597397.1:1-227 Altererythrobacter ishigakiensis Altererythrobacter ishigakiensis
WP 034906536.1:1-226 Erythrobacter litoralis Erythrobacter litoralis
WP 115490757.1:1-226 Erythrobacter lutimaris Erythrobacter lutimaris
0
WP 068865824.1:1-226 Porphyrobacter dokdonensis Porphyrobacter dokdonensis
WP 120048734.1:1-225 Erythrobacter aquimixticola Erythrobacter aquimixticola
WP 047004150.1:1-225 Erythrobacter luteus Erythrobacter luteus
WP 047007059.1:1-225 Erythrobacter gangjinensis Erythrobacter gangjinensis
67
WP 046904133.1:1-226 Altererythrobacter atlanticus Altererythrobacter atlanticus
14
WP 119592451.1:1-226 Erythrobacter xanthus Erythrobacter xanthus
WP 047092762.1:1-225 Erythrobacter marinus Erythrobacter marinus
20
WP 067784887.1:1-226 Altererythrobacter namhicola Altererythrobacter namhicola
WP 120076434.1:1-225 Erythrobacter odishensis Erythrobacter odishensis
WP 048884763.1:1-225 Erythrobacter atlanticus Erythrobacter atlanticus
WP 009489600.1:1-224 Microvirga lotononidis Microvirga lotononidis
WP 147023481.1:1-224 Microvirga aerophila Microvirga aerophila
32
WP 114948006.1:1-224 Microvirga calopogonii Microvirga calopogonii
WP 082145892.1:6-239 Microvirga massiliensis Microvirga massiliensis
WP 018631868.1:1-223 Neomegalonema perideroedes Neomegalonema perideroedes
WP 072747388.1:1-224 Oceanicella actignis Oceanicella actignis
64
WP 092683496.1:1-224 Albimonas donghaensis Albimonas donghaensis
WP 092857989.1:1-224 Albimonas pacifica Albimonas pacifica
WP 165095699.1:1-224 Halovulum dunhuangense Halovulum dunhuangense
96
WP 097930030.1:1-224 Monaibacterium marinum Monaibacterium marinum
WP 108782138.1:1-224 Pontivivens insulae Pontivivens insulae
WP 069299104.1:1-237 Neptunicoccus sediminis Neptunicoccus sediminis
100
WP 101656770.1:1-237 Amylibacter cionae Amylibacter cionae
61
WP 025312677.1:1-224 Roseibacterium elongatum Roseibacterium elongatum
13
WP 109665574.1:1-227 Roseicyclus mahoneyensis Roseicyclus mahoneyensis
2
ARO14354.1:1-233 Ketogulonicigenium robustum Ketogulonicigenium robustum
94
WP 013384333.1:1-238 Ketogulonicigenium vulgare Ketogulonicigenium vulgare
WP 092501145.1:1-237 Albidovulum xiamenense Albidovulum xiamenense
0
SMY07477.1:1-221 Flavimaricola marinus Flavimaricola marinus
100
WP 093991673.1:1-238 Flavimaricola marinus Flavimaricola marinus
WP 093421264.1:1-237 Tranquillimonas alkanivorans Tranquillimonas alkanivorans
1
WP 092686857.1:1-225 Roseivivax roseus Roseivivax roseus
39
WP 099823683.1:1-237 Oceaniglobus indicus Oceaniglobus indicus
WP 102223274.1:1-224 Acidimangrovimonas sediminis Acidimangrovimonas sediminis
WP 007427489.1:1-225 Oceaniovalibus guishaninsula Oceaniovalibus guishaninsula
0
WP 097374166.1:1-237 Celeribacter manganoxidans Celeribacter manganoxidans
WP 093972223.1:1-225 Boseongicola aestuarii Boseongicola aestuarii
0
WP 036634037.1:1-237 Paenirhodobacter enshiensis Paenirhodobacter enshiensis
10
WP 106159526.1:1-224 Hasllibacter halocynthiae Hasllibacter halocynthiae
83
WP 007254149.1:1-238 Oceanicola granulosus Oceanicola granulosus
26
WP 085877167.1:1-238 Roseisalinus antarcticus Roseisalinus antarcticus
1
WP 007204864.1:1-237 Yoonia vestfoldensis Yoonia vestfoldensis
WP 090197976.1:1-225 Yoonia tamlensis Yoonia tamlensis
0
WP 106359048.1:1-225 Yoonia maritima Yoonia maritima
WP 108387231.1:1-237 Yoonia sediminilitoris Yoonia sediminilitoris
WP 090205327.1:1-237 Yoonia litorea Yoonia litorea
0
WP 055292785.1:1-237 Yoonia rosea Yoonia rosea
WP 100366699.1:1-237 Yoonia maricola Yoonia maricola
WP 089991178.1:1-237 Cognatiyoonia koreensis Cognatiyoonia koreensis

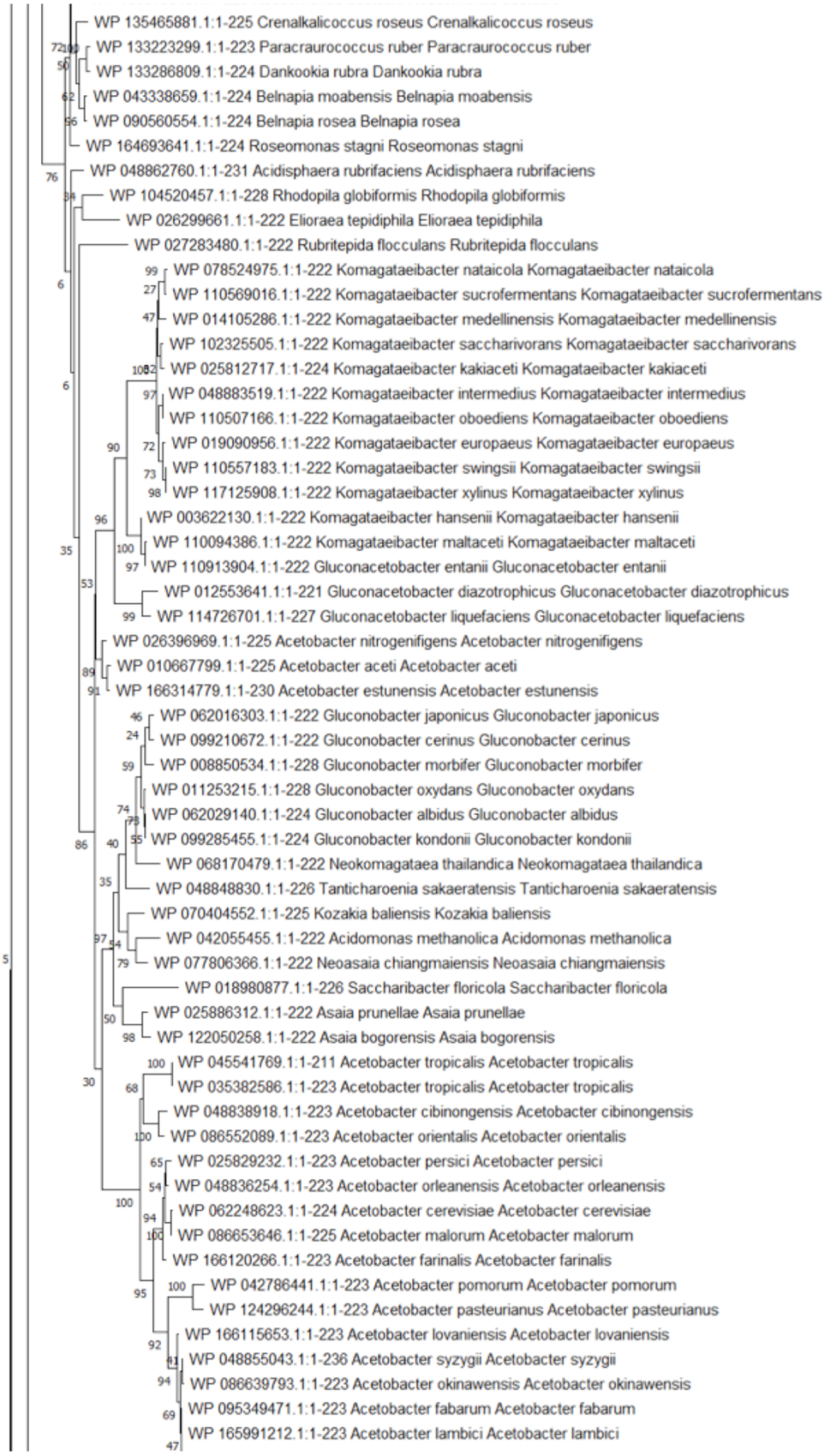

WP 015500021.1:1-225 Octadecabacter antarcticus Octadecabacter antarcticus
WP 093994687.1:1-225 Octadecabacter ascidiaceicola Octadecabacter ascidiaceicola
WP 085862919.1:1-239 Pseudooctadecabacter jejudonensis Pseudooctadecabacter jejudonensis
WP 045320127.1:1-225 Limimaricola cinnabarinus Limimaricola cinnabarinus
WP 085896647.1:1-237 Limimaricola soesokkakensis Limimaricola soesokkakensis
WP 018301384.1:1-238 Wenxinia marina Wenxinia marina
WP 073325651.1:1-238 Wenxinia saemankumensis Wenxinia saemankumensis
WP 035257264.1:1-225 Actibacterium mucosum Actibacterium mucosum
WP 093966673.1:1-237 Confluentimicrobium lipolyticum Confluentimicrobium lipolyticum
WP 072855717.1:1-237 Loktanella atrilutea Loktanella atrilutea
WP 089897464.1:1-237 Loktanella fryxellensis Loktanella fryxellensis
WP 090735598.1:1-236 Paracoccus homiensis Paracoccus homiensis
WP 119897655.1:1-236 Paracoccus siganidrum Paracoccus siganidrum
WP 089345409.1:1-236 Paracoccus seriniphilus Paracoccus seriniphilus
WP 108501378.1:1-236 Paracoccus indicus Paracoccus indicus
WP 090744932.1:1-233 Paracoccus tibetensis Paracoccus tibetensis
WP 136857851.1:1-236 Paracoccus hibiscisoli Paracoccus hibiscisoli
WP 089388317.1:1-235 Paracoccus sediminis Paracoccus sediminis
WP 111302132.1:1-232 Paracoccus saliphilus Paracoccus saliphilus
WP 076525161.1:1-236 Paracoccus saliphilus Paracoccus saliphilus
WP 022708263.1:1-236 Paracoccus zeaxanthinifaciens Paracoccus zeaxanthinifaciens
WP 101499674.1:1-236 Paracoccus jeotgali Paracoccus jeotgali
TKW66205.1:1-236 Paracoccus denitrificans Paracoccus denitrificans
WP 090521639.1:1-236 Paracoccus isoporae Paracoccus isoporae
WP 028721041.1:1-224 Paracoccus yeei Paracoccus yeei
WP 114349453.1:1-224 Paracoccus lutimaris Paracoccus lutimaris
WP 036738733.1:1-238 Paracoccus halophilus Paracoccus halophilus
WP 062560688.1:1-224 Paracoccus aminovorans Paracoccus aminovorans
WP 114669328.1:1-224 Paracoccus pantotrophus Paracoccus pantotrophus
WP 115756542.1:1-224 Paracoccus thiocyanatus Paracoccus thiocyanatus
WP 073062671.1:1-236 Paracoccus solventivorans Paracoccus solventivorans
WP 090846811.1:1-236 Paracoccus alkenifer Paracoccus alkenifer
WP 036705087.1:1-237 Paracoccus sanguinis Paracoccus sanguinis
WP 090755879.1:1-236 Paracoccus chinensis Paracoccus chinensis
WP 085376660.1:1-237 Paracoccus contaminans Paracoccus contaminans
WP 036716921.1:1-237 Paracoccus sphaerophysae Paracoccus sphaerophysae
PZO65986.1:1-237 Paracoccus denitrificans Paracoccus denitrificans
WP 092891242.1:1-224 Roseicitreum antarcticum Roseicitreum antarcticum
WP 121093174.1:1-224 Roseibaca ekhonensis Roseibaca ekhonensis
WP 071468905.1:1-227 Roseinatronobacter thiooxidans Roseinatronobacter thiooxidans
WP 111535570.1:1-233 Hwanghaeicola aestuarii Hwanghaeicola aestuarii
WP 085854953.1:1-225 Palleronia marisminoris Palleronia marisminoris
WP 108892190.1:1-228 Palleronia abyssalis Palleronia abyssalis
PTR19694.1:1-219 Rhodobacter azotoformans Rhodobacter azotoformans
ACM00793.1:1-232 Rhodobacter sphaeroides KD131 Rhodobacter sphaeroides KD131
GEM91710.1:1-219 Rhodobacter sphaeroides Rhodobacter sphaeroides
WP 069331091.1:1-237 Rhodobacter johrii Rhodobacter johrii
WP 097028834.1:1-237 Rhodobacter ovatus Rhodobacter ovatus
WP 107323687.1:1-237 Rhodobacter veldkampii Rhodobacter veldkampii
PZQ99575.1:1-236 Rhodobacter sphaeroides Rhodobacter sphaeroides
WP 108852354.1:1-237 Defluviimonas aquaemixtae Defluviimonas aquaemixtae
WP 096432385.1:1-238 Thioclava arenosa Thioclava arenosa
WP 055208583.1:1-237 Rhodobacter capsulatus Rhodobacter capsulatus
WP 110803963.1:1-237 Rhodobacter viridis Rhodobacter viridis
WP 066812765.1:1-237 Defluviimonas alba Defluviimonas alba
WP 038068922.1:1-237 Thioclava dalianensis Thioclava dalianensis
WP 038129178.1:1-237 Thioclava indica Thioclava indica
WP 038077965.1:1-237 Thioclava pacifica Thioclava pacifica

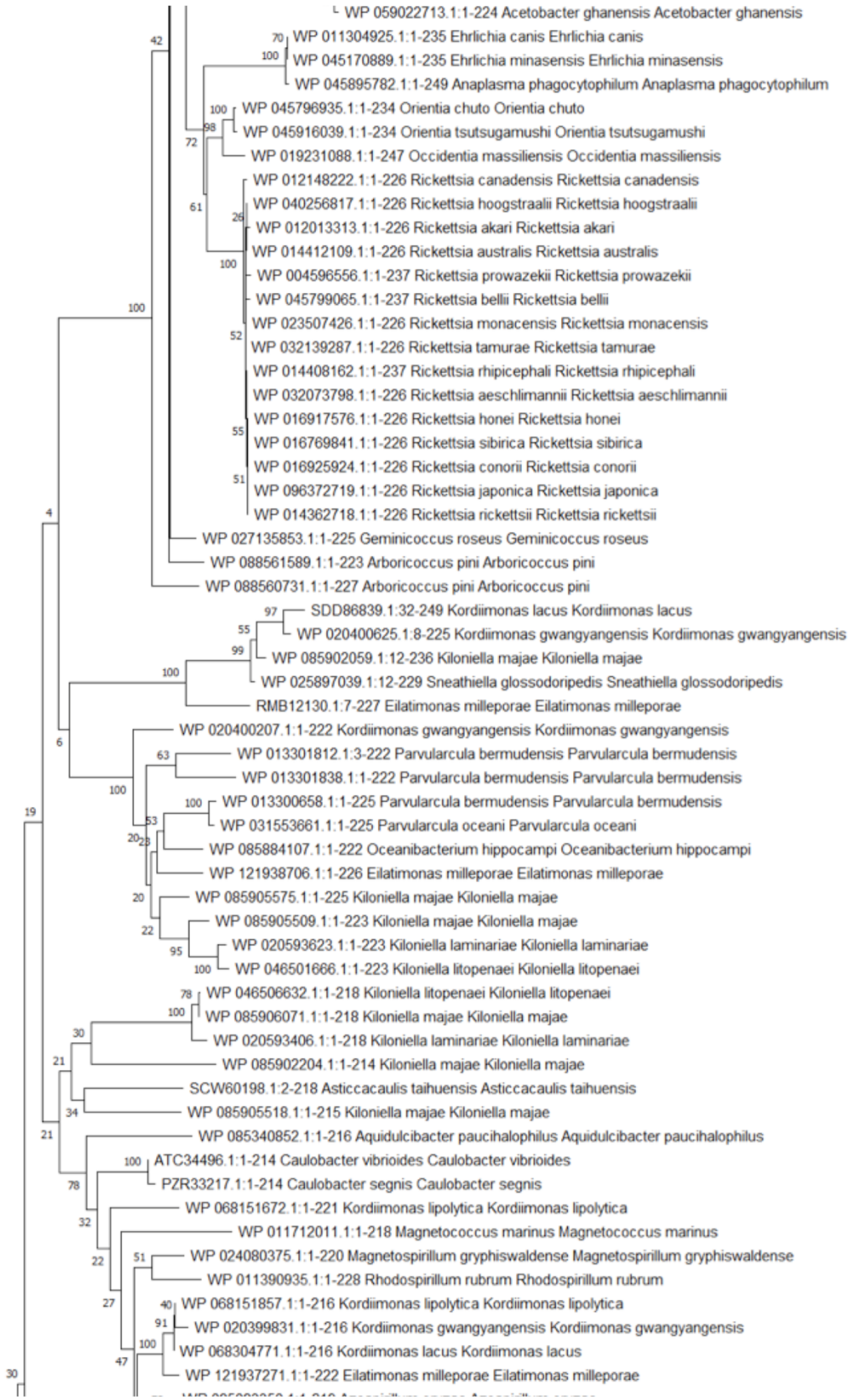

WP 038145212.1:1-237 Thioclava atlantica Thioclava atlantica
WP 028028601.1:1-231 Gemmobacter nectariphilus Gemmobacter nectariphilus
WP 076528764.1:1-231 Gemmobacter megaterium Gemmobacter megaterium
WP 022704497.1:1-237 Pseudorhodobacter ferrugineus Pseudorhodobacter ferrugineus
WP 050519127.1:1-237 Pseudorhodobacter antarcticus Pseudorhodobacter antarcticus
WP 054004572.1:1-224 Pseudorhodobacter psychrotolerans Pseudorhodobacter psychrotolerans
WP 103255894.1:1-224 Tabrizicola aquatica Tabrizicola aquatica
WP 091298249.1:1-237 Gemmobacter aquatilis Gemmobacter aquatilis
WP 107673853.1:1-224 Rhodobacter blasticus Rhodobacter blasticus
WP 112309466.1:1-237 Pseudogemmobacter bohemicus Pseudogemmobacter bohemicus
WP 021097010.1:1-225 Rubellimicrobium thermophilum Rubellimicrobium thermophilum
WP 037279073.1:1-238 Rubellimicrobium mesophilum Rubellimicrobium mesophilum
WP 039003605.1:1-237 Halocynthiibacter arcticus Halocynthiibacter arcticus
WP 091430055.1:1-237 Aliiroseovarius sediminilitoris Aliiroseovarius sediminilitoris
WP 108856390.1:1-237 Aliiroseovarius pelagivivens Aliiroseovarius pelagivivens
WP 035712854.1:1-224 Haematobacter massiliensis Haematobacter massiliensis
WP 035746022.1:1-224 Haematobacter missouriensis Haematobacter missouriensis
WP 055192113.1:1-237 Aliiroseovarius crassostreae Aliiroseovarius crassostreae
WP 058123869.1:1-237 Thalassobacter stenotrophicus Thalassobacter stenotrophicus
WP 055682445.1:1-234 Jannaschia rubra Jannaschia rubra
WP 043918128.1:1-225 Jannaschia aquimarina Jannaschia aquimarina
WP 092776251.1:1-234 Jannaschia pohangensis Jannaschia pohangensis
WP 092764616.1:1-234 Jannaschia helgolandensis Jannaschia helgolandensis
WP 092644234.1:1-225 Jannaschia faecimaris Jannaschia faecimaris
WP 055664570.1:1-234 Jannaschia seosinensis Jannaschia seosinensis
WP 055086227.1:1-233 Jannaschia donghaensis Jannaschia donghaensis
WP 084350182.1:1-237 Pseudooceanicola flagellatus Pseudooceanicola flagellatus
WP 132248181.1:1-238 Primorskyibacter sedentarius Primorskyibacter sedentarius
WP 073251175.1:1-224 Pseudopelagicola gijangensis Pseudopelagicola gijangensis
WP 085866621.1:1-225 Pseudoruegeria aquimaris Pseudoruegeria aquimaris
WP 050529307.1:1-237 Pseudaestuariivita atlantica Pseudaestuariivita atlantica
WP 068331173.1:1-238 Ruegeria profundi Ruegeria profundi
WP 085824113.1:1-238 Ruegeria meonggei Ruegeria meonggei
WP 058277902.1:1-238 Ruegeria atlantica Ruegeria atlantica
WP 010439997.1:1-238 Ruegeria conchae Ruegeria conchae
WP 074734646.1:1-238 Ruegeria halocynthiae Ruegeria halocynthiae
WP 093962018.1:1-238 Ruegeria arenilitoris Ruegeria arenilitoris
WP 100164762.1:1-238 Pseudooceanicola lipolyticus Pseudooceanicola lipolyticus
WP 027261367.1:1-238 Sedimentitalea nanhaiensis Sedimentitalea nanhaiensis
WP 050927779.1:1-238 Aestuariivita boseongensis Aestuariivita boseongensis
WP 068344186.1:1-238 Ruegeria marisrubri Ruegeria marisrubri
WP 092361598.1:1-243 Cribrihabitans marinus Cribrihabitans marinus
WP 113821694.1:1-238 Phaeobacter gallaeciensis Phaeobacter gallaeciensis
WP 011047414.1:1-238 Ruegeria pomeroyi Ruegeria pomeroyi
WP 088648076.1:1-238 Marinibacterium profundimaris Marinibacterium profundimaris
WP 121064998.1:1-229 Chachezhania antarctica Chachezhania antarctica
WP 104006534.1:1-237 Jhaorihella thermophila Jhaorihella thermophila
WP 134013282.1:1-224 Litoreibacter halocynthiae Litoreibacter halocynthiae
WP 090216974.1:1-224 Litoreibacter janthinus Litoreibacter janthinus
WP 073141260.1:1-224 Litoreibacter ascidiaceicola Litoreibacter ascidiaceicola
WP 089946732.1:1-224 Litoreibacter albidus Litoreibacter albidus
WP 121024941.1:1-224 Litoreibacter meonggei Litoreibacter meonggei
WP 021100478.1:1-224 Litoreibacter arenae Litoreibacter arenae
WP 107844064.1:1-228 Litoreibacter ponti Litoreibacter ponti
WP 074255502.1:1-238 Oceanicola litoreus Oceanicola litoreus
WP 132599588.1:1-225 Rhodovulum adriaticum Rhodovulum adriaticum
WP 132950324.1:1-238 Rhodovulum bhavnagarense Rhodovulum bhavnagarense
WP 132695342.1:1-239 Rhodovulum steppense Rhodovulum steppense
WP 132462218.1:1-225 Rhodovulum marinum Rhodovulum marinum

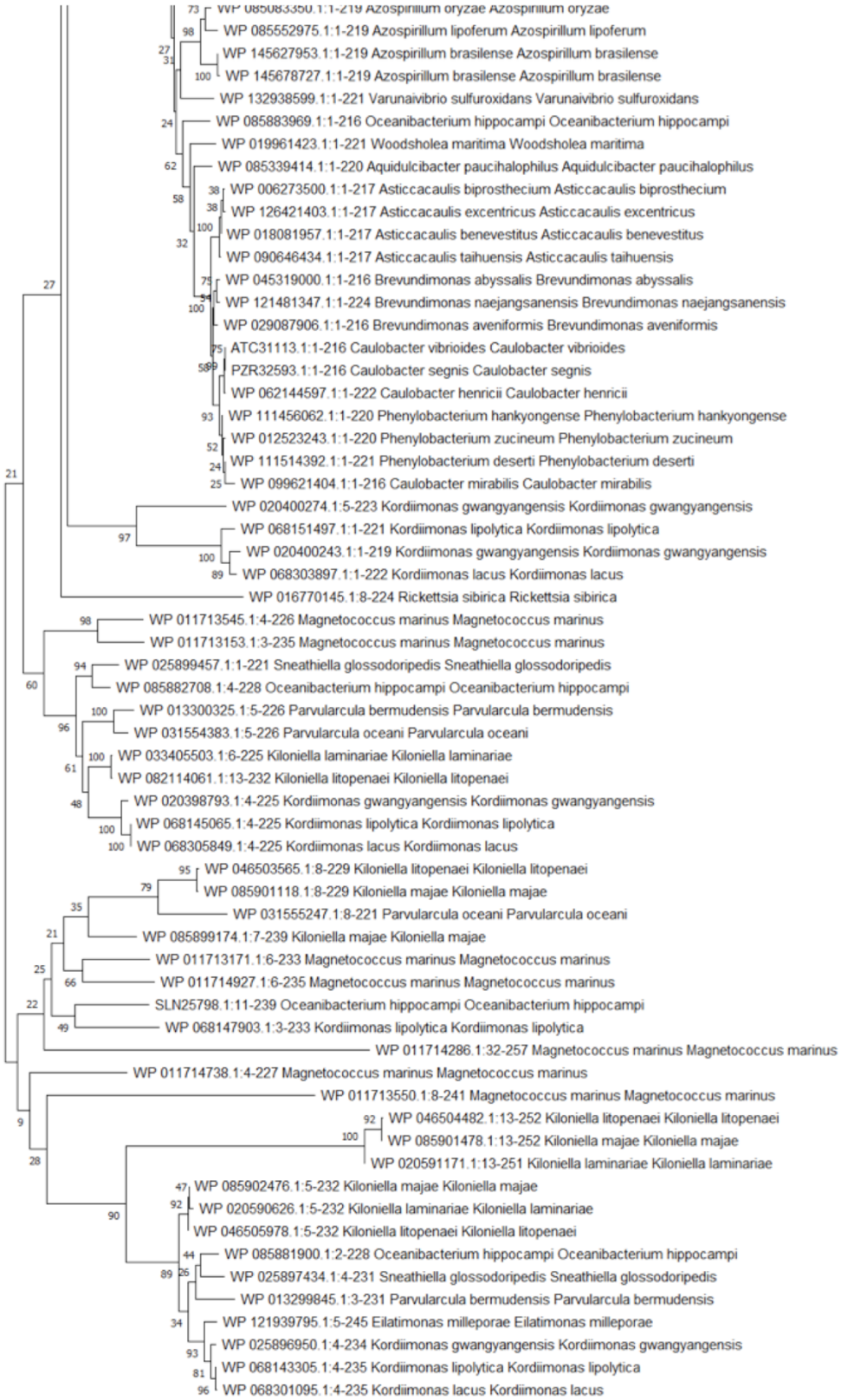

WP 132546140.1:1-239 Rhodovulum euryhalinum Rhodovulum euryhalinum
WP 042462016.1:1-225 Rhodovulum sulfidophilum Rhodovulum sulfidophilum
WP 106263926.1:1-224 Donghicola tyrosinivorans Donghicola tyrosinivorans
WP 072789228.1:1-225 Cognatishimia maritima Cognatishimia maritima
WP 012178180.1:1-225 Dinoroseobacter shibae Dinoroseobacter shibae
WP 107891575.1:1-225 Rhodovulum imhoffii Rhodovulum imhoffii
WP 125403245.1:1-237 Rhodovulum robiginosum Rhodovulum robiginosum
WP 093149574.1:1-237 Pseudoruegeria lutimaris Pseudoruegeria lutimaris
WP 106203391.1:1-237 Pseudoruegeria haliotis Pseudoruegeria haliotis
WP 068116873.1:1-237 Pseudoruegeria marinistellae Pseudoruegeria marinistellae
WP 089232964.1:1-237 Tropicimonas sediminicola Tropicimonas sediminicola
WP 086452442.1:1-237 Marivivens niveibacter Marivivens niveibacter
WP 092061002.1:1-225 Poseidonocella pacifica Poseidonocella pacifica
WP 092077779.1:1-237 Poseidonocella sedimentorum Poseidonocella sedimentorum
WP 035247497.1:1-239 Actibacterium atlanticum Actibacterium atlanticum
WP 095589635.1:1-237 Actibacterium ureilyticum Actibacterium ureilyticum
WP 095596369.1:1-224 Actibacterium pelagium Actibacterium pelagium
All87069.1:1-237 Planktomarina temperata RCA23 Planktomarina temperata RCA23
WP 085791763.1:1-225 Roseivivax jejudonensis Roseivivax jejudonensis
WP 043842422.1:1-226 Roseivivax marinus Roseivivax marinus
WP 043768217.1:1-234 Roseivivax isoporae Roseivivax isoporae
WP 076446140.1:1-238 Roseivivax lentus Roseivivax lentus
WP 103908572.1:1-224 Thalassococcus halodurans Thalassococcus halodurans
WP 093010767.1:1-224 Roseivivax halotolerans Roseivivax halotolerans
WP 115979220.1:1-223 Rhodosalinus sediminis Rhodosalinus sediminis
WP 072777022.1:1-236 Marivita hallyeonensis Marivita hallyeonensis
WP 085635873.1:1-236 Marivita geojedonensis Marivita geojedonensis
WP 085309206.1:1-224 Planktotalea arctica Planktotalea arctica
WP 108886665.1:1-238 Primorskyibacter insulae Primorskyibacter insulae
WP 099912367.1:1-238 Puniceibacterium antarcticum Puniceibacterium antarcticum
WP 110813191.1:1-226 Pseudoroseicyclus aestuarii Pseudoroseicyclus aestuarii
WP 058860874.1:1-225 Ponticoccus marisrubri Ponticoccus marisrubri
WP 089879424.1:1-233 Citreimonas salinaria Citreimonas salinaria
WP 097802679.1:1-225 Pelagimonas varians Pelagimonas varians
WP 099245838.1:1-225 Tropicibacter phthalicicus Tropicibacter phthalicicus
WP 094035779.1:1-225 Antarctobacter heliothermus Antarctobacter heliothermus
WP 043139223.1:1-225 Mameliella alba Mameliella alba
CUH80454.1:8-232 Tropicibacter naphthalenivorans Tropicibacter naphthalenivorans
WP 005857700.1:1-225 Sagittula stellata Sagittula stellata
WP 094019692.1:1-234 Maliponia aquimaris Maliponia aquimaris
WP 093027004.1:1-224 Ruegeria marina Ruegeria marina
WP 020039956.1:1-225 Salipiger mucosus Salipiger mucosus
WP 107975610.1:1-224 Roseivivax pacificus Roseivivax pacificus
WP 076703734.1:1-225 Pelagibaca abyssi Pelagibaca abyssi
WP 111551056.1:1-225 Salipiger aestuarii Salipiger aestuarii
WP 092420774.1:1-225 Yangia pacifica Yangia pacifica
WP 089851756.1:1-225 Salipiger marinus Salipiger marinus
WP 093732192.1:1-224 Sulfitobacter litoralis Sulfitobacter litoralis
WP 093914977.1:1-238 Sulfitobacter marinus Sulfitobacter marinus
WP 025060768.1:1-230 Sulfitobacter donghicola Sulfitobacter donghicola
WP 025062013.1:1-235 Sulfitobacter guttiformis Sulfitobacter guttiformis
WP 011568769.1:1-239 Roseobacter denitrificans Roseobacter denitrificans
WP 013961871.1:1-239 Roseobacter litoralis Roseobacter litoralis
WP 108828388.1:1-239 Ascidiaceihabitans donghaensis Ascidiaceihabitans donghaensis
WP 076627493.1:1-239 Tateyamaria omphalii Tateyamaria omphalii
WP 007120435.1:1-230 Sulfitobacter indolifex Sulfitobacter indolifex
WP 093926395.1:1-230 Sulfitobacter dubius Sulfitobacter dubius
WP 093922531.1:1-225 Sulfitobacter brevis Sulfitobacter brevis


WP 093742256.1:1-230 Sulfitobacter delicatus Sulfitobacter delicatus
WP 025044629.1:1-230 Sulfitobacter geojensis Sulfitobacter geojensis
WP 025057497.1:1-225 Sulfitobacter noctilucicola Sulfitobacter noctilucicola
WP 025052854.1:1-225 Sulfitobacter noctilucae Sulfitobacter noctilucae
WP 025048148.1:1-230 Sulfitobacter mediterraneus Sulfitobacter mediterraneus
WP 093320659.1:1-225 Shimia haliotis Shimia haliotis
WP 058240289.1:1-225 Shimia marina Shimia marina
WP 132859326.1:1-225 Shimia isoporae Shimia isoporae
WP 093115686.1:1-223 Salinihabitans flavidus Salinihabitans flavidus
WP 093094050.1:1-229 Shimia aestuarii Shimia aestuarii
PSL17389.1:1-230 Shimia abyssi Shimia abyssi
WP 111361948.1:1-224 Aquicoccus porphyridii Aquicoccus porphyridii
WP 028092150.1:1-237 Pseudodonghicola xiamenensis Pseudodonghicola xiamenensis
WP 024090282.1:1-239 Leisingera methylohalidivorans Leisingera methylohalidivorans
WP 141890204.1:1-239 Leisingera aquaemixtae Leisingera aquaemixtae
WP 027258794.1:1-239 Leisingera aquimarina Leisingera aquimarina
WP 005611702.1:1-239 Epibacterium mobile Epibacterium mobile
WP 106164032.1:1-239 Epibacterium scottomollicae Epibacterium scottomollicae
WP 058291744.1:1-239 Epibacterium multivorans Epibacterium multivorans
WP 072504480.1:1-239 Phaeobacter porticola Phaeobacter porticola
WP 008561700.1:1-239 Phaeobacter italicus Phaeobacter italicus
WP 136461289.1:1-238 Aliishimia ponticola Aliishimia ponticola
WP 009804620.1:1-238 Pseudooceanicola batsensis Pseudooceanicola batsensis
WP 093448250.1:1-238 Pseudooceanicola nitratireducens Pseudooceanicola nitratireducens
WP 076978727.1:1-224 Brevirhabdus pacifica Brevirhabdus pacifica
EAP76201.1:1-233 Roseovarius nubinhibens ISM Roseovarius nubinhibens ISM
HAR52580.1:1-238 Roseovarius nubinhibens Roseovarius nubinhibens
WP 071690322.1:1-224 Nioella sediminis Nioella sediminis
WP 071675710.1:1-224 Nioella nitratireducens Nioella nitratireducens
WP 085805211.1:1-239 Roseovarius albus Roseovarius albus
WP 073032032.1:1-238 Roseovarius pacificus Roseovarius pacificus
WP 085816707.1:1-238 Roseovarius halotolerans Roseovarius halotolerans
WP 074784387.1:1-229 Roseovarius tolerans Roseovarius tolerans
CUH83110.1:1-226 Thalassobius mediterraneus Thalassobius mediterraneus
WP 058260947.1:1-238 Thalassobius gelatinovorus Thalassobius gelatinovorus
WP 090025710.1:1-241 Lutimaribacter saemankumensis Lutimaribacter saemankumensis
WP 104020499.1:1-225 Roseovarius nitratireducens Roseovarius nitratireducens
MTJ05252.1:1-240 Sediminimonas qiaohouensis Sediminimonas qiaohouensis
WP 136338864.1:1-238 Youngimonas vesicularis Youngimonas vesicularis
WP 048597595.1:1-238 Nereida ignava Nereida ignava
WP 089891328.1:1-230 Lentibacter algarum Lentibacter algarum
WP 138016316.1:1-234 Pelagicola litoralis Pelagicola litoralis
WP 057789204.1:1-240 Roseovarius atlanticus Roseovarius atlanticus
WP 103761276.1:1-240 Roseovarius confluentis Roseovarius confluentis
WP 135504783.1:1-238 Roseovarius aestuariivivens Roseovarius aestuariivivens
WP 092833556.1:1-238 Roseovarius lutimaris Roseovarius lutimaris
WP 073194030.1:1-238 Roseovarius marisflavi Roseovarius marisflavi
WP 085825165.1:1-238 Roseovarius gaetbuli Roseovarius gaetbuli
WP 076535169.1:1-238 Roseovarius nanhaiticus Roseovarius nanhaiticus
WP 113910890.1:1-238 Pelagivirga dicentrarchi Pelagivirga dicentrarchi
WP 085799074.1:1-238 Roseovarius aestuarii Roseovarius aestuarii
WP 028285667.1:1-238 Pseudooceanicola nanhaiensis Pseudooceanicola nanhaiensis
WP 057815505.1:1-238 Roseovarius indicus Roseovarius indicus
WP 085890628.1:1-238 Pelagicola litorisediminis Pelagicola litorisediminis
WP 085850486.1:1-237 Pacificibacter marinus Pacificibacter marinus
WP 123791690.1:1-237 Pacificibacter maritimus Pacificibacter maritimus
WP 093038763.1:1-238 Roseovarius azorensis Roseovarius azorensis
WP 085836407.1:1-237 Aquimixticola soesokkakensis Aquimixticola soesokkakensis


WP 043871698.1:1-239 Celeribacter indicus Celeribacter indicus
WP 066599969.1:1-238 Celeribacter halophilus Celeribacter halophilus
WP 066709692.1:1-239 Celeribacter ethanolicus Celeribacter ethanolicus
WP 107817639.1:1-239 Celeribacter persicus Celeribacter persicus
WP 062215404.1:1-237 Celeribacter marinus Celeribacter marinus
WP 009571891.1:1-239 Celeribacter baekdonensis Celeribacter baekdonensis
WP 090057720.1:1-239 Celeribacter neptunius Celeribacter neptunius
WP 086466013.1:1-225 Oceanibaculum nanhaiense Oceanibaculum nanhaiense
WP 008945675.1:1-231 Oceanibaculum indicum Oceanibaculum indicum
WP 067559112.1:1-227 Oceanibaculum pacificum Oceanibaculum pacificum
WP 127767748.1:1-226 Hwanghaeella grinnelliae Hwanghaeella grinnelliae
WP 115934713.1:1-230 Aestuariispira insulae Aestuariispira insulae
WP 085552657.1:1-227 Azospirillum lipoferum Azospirillum lipoferum
WP 044437413.1:1-237 Skermanella aerolata Skermanella aerolata
WP 029007126.1:1-226 Azospirillum halopraeferens Azospirillum halopraeferens
WP 045776988.1:1-223 Elstera litoralis Elstera litoralis
WP 094407314.1:1-223 Elstera cyanobacteriorum Elstera cyanobacteriorum
WP 029013737.1:1-225 Niveispirillum irakense Niveispirillum irakense
WP 094456117.1:1-225 Niveispirillum lacus Niveispirillum lacus
WP 133612373.1:1-229 Dongia mobilis Dongia mobilis
PZP85741.1:1-225 Azospirillum brasilense Azospirillum brasilense
WP 114580140.1:1-224 Ferruginivarius sediminum Ferruginivarius sediminum
WP 090018482.1:1-228 Limimonas halophila Limimonas halophila
WP 027289079.1:1-225 Rhodovibrio salinarum Rhodovibrio salinarum
WP 020592224.1:1-228 Kiloniella laminariae Kiloniella laminariae
WP 085900876.1:1-237 Kiloniella majae Kiloniella majae
WP 022727365.1:1-237 Fodinicurvata sediminis Fodinicurvata sediminis
WP 026987209.1:1-237 Fodinicurvata fenggangensis Fodinicurvata fenggangensis
WP 014744746.1:1-227 Tistrella mobilis Tistrella mobilis
WP 028879163.1:1-224 Terasakiella pusilla Terasakiella pusilla
WP 062959876.1:1-225 Thalassospira xiamenensis Thalassospira xiamenensis
WP 085620984.1:1-225 Thalassospira alkalitolerans Thalassospira alkalitolerans
WP 085580703.1:1-225 Thalassospira mesophila Thalassospira mesophila
WP 101263889.1:1-225 Thalassospira marina Thalassospira marina
WP 069957282.1:1-226 Magnetovibrio blakemorei Magnetovibrio blakemorei
WP 132938453.1:1-226 Varunaivibrio sulfuroxidans Varunaivibrio sulfuroxidans
WP 024082208.1:1-225 Magnetospirillum gryphiswaldense Magnetospirillum gryphiswaldense
WP 068498488.1:1-225 Magnetospirillum moscoviense Magnetospirillum moscoviense
WP 068492688.1:1-225 Magnetospirillum marisnigri Magnetospirillum marisnigri
WP 008616069.1:1-225 Magnetospirillum caucaseum Magnetospirillum caucaseum
WP 041041883.1:1-225 Magnetospirillum magnetotacticum Magnetospirillum magnetotacticum
SEH51586.1:14-238 Phaeospirillum fulvum Phaeospirillum fulvum
SOD96028.1:1-227 Caenispirillum bisanense Caenispirillum bisanense
WP 019644211.1:1-226 Novispirillum itersonii Novispirillum itersonii
WP 066136033.1:1-224 Haematospirillum jordaniae Haematospirillum jordaniae
WP 076401106.1:1-228 Insolitispirillum peregrinum Insolitispirillum peregrinum
WP 092781039.1:1-224 Rhodospira trueperi Rhodospira trueperi
WP 011388283.1:1-225 Rhodospirillum rubrum Rhodospirillum rubrum
WP 041793636.1:1-225 Pararhodospirillum photometricum Pararhodospirillum photometricum
WP 094473039.1:1-226 Sandarakinorhabdus cyanobacteriorum Sandarakinorhabdus cyanobacteriorum
WP 159348062.1:1-223 Roseomonas deserti Roseomonas deserti
WP 043835504.1:1-223 Roseomonas aerilata Roseomonas aerilata
WP 073130964.1:1-223 Roseomonas rosea Roseomonas rosea
PZP46705.1:1-223 Azospirillum brasilense Azospirillum brasilense
WP 145135330.1:1-223 Roseomonas gilardii Roseomonas gilardii
WP 111400555.1:1-223 Humitalea rosea Humitalea rosea
WP 076959643.1:1-223 Roseomonas deserti Roseomonas deserti
WP 099095590.1:1-223 Roseomonas rhizosphaerae Roseomonas rhizosphaerae
WP 109516546.1:1-223 Roseomonas aestuarii Roseomonas aestuarii


WP 135465881.1:1-225 Crenalkalicoccus roseus Crenalkalicoccus roseus
WP 133223299.1:1-223 Paracraurococcus ruber Paracraurococcus ruber
WP 133286809.1:1-224 Dankookia rubra Dankookia rubra
WP 043338659.1:1-224 Belnapia moabensis Belnapia moabensis
WP 090560554.1:1-224 Belnapia rosea Belnapia rosea
WP 164693641.1:1-224 Roseomonas stagni Roseomonas stagni
WP 048862760.1:1-231 Acidisphaera rubrifaciens Acidisphaera rubrifaciens
WP 104520457.1:1-228 Rhodopila globiformis Rhodopila globiformis
WP 026299661.1:1-222 Elioraea tepidiphila Elioraea tepidiphila
WP 027283480.1:1-222 Rubritepida flocculans Rubritepida flocculans
WP 078524975.1:1-222 Komagataeibacter nataicola Komagataeibacter nataicola
WP 110569016.1:1-222 Komagataeibacter sucrofermentans Komagataeibacter sucrofermentans
WP 014105286.1:1-222 Komagataeibacter medellinensis Komagataeibacter medellinensis
WP 102325505.1:1-222 Komagataeibacter saccharivorans Komagataeibacter saccharivorans
WP 025812717.1:1-224 Komagataeibacter kakiaceti Komagataeibacter kakiaceti
WP 048883519.1:1-222 Komagataeibacter intermedius Komagataeibacter intermedius
WP 110507166.1:1-222 Komagataeibacter oboediens Komagataeibacter oboediens
WP 019090956.1:1-222 Komagataeibacter europaeus Komagataeibacter europaeus
WP 110557183.1:1-222 Komagataeibacter swingsii Komagataeibacter swingsii
WP 117125908.1:1-222 Komagataeibacter xylinus Komagataeibacter xylinus
WP 003622130.1:1-222 Komagataeibacter hansenii Komagataeibacter hansenii
WP 110094386.1:1-222 Komagataeibacter maltaceti Komagataeibacter maltaceti
WP 110913904.1:1-222 Gluconacetobacter entanii Gluconacetobacter entanii
WP 012553641.1:1-221 Gluconacetobacter diazotrophicus Gluconacetobacter diazotrophicus
WP 114726701.1:1-227 Gluconacetobacter liquefaciens Gluconacetobacter liquefaciens
WP 026396969.1:1-225 Acetobacter nitrogenifigens Acetobacter nitrogenifigens
WP 010667799.1:1-225 Acetobacter aceti Acetobacter aceti
WP 166314779.1:1-230 Acetobacter estunensis Acetobacter estunensis
WP 062016303.1:1-222 Gluconobacter japonicus Gluconobacter japonicus
WP 099210672.1:1-222 Gluconobacter cerinus Gluconobacter cerinus
WP 008850534.1:1-228 Gluconobacter morbifer Gluconobacter morbifer
WP 011253215.1:1-228 Gluconobacter oxydans Gluconobacter oxydans
WP 062029140.1:1-224 Gluconobacter albidus Gluconobacter albidus
WP 099285455.1:1-224 Gluconobacter kondonii Gluconobacter kondonii
WP 068170479.1:1-222 Neokomagataea thailandica Neokomagataea thailandica
WP 048848830.1:1-226 Tanticharoenia sakaeratensis Tanticharoenia sakaeratensis
WP 070404552.1:1-225 Kozakia baliensis Kozakia baliensis
WP 042055455.1:1-222 Acidomonas methanolica Acidomonas methanolica
WP 077806366.1:1-222 Neoasaia chiangmaiensis Neoasaia chiangmaiensis
WP 018980877.1:1-226 Saccharibacter floricola Saccharibacter floricola
WP 025886312.1:1-222 Asaia prunellae Asaia prunellae
WP 122050258.1:1-222 Asaia bogorensis Asaia bogorensis
WP 045541769.1:1-211 Acetobacter tropicalis Acetobacter tropicalis
WP 035382586.1:1-223 Acetobacter tropicalis Acetobacter tropicalis
WP 048838918.1:1-223 Acetobacter cibinongensis Acetobacter cibinongensis
WP 086552089.1:1-223 Acetobacter orientalis Acetobacter orientalis
WP 025829232.1:1-223 Acetobacter persici Acetobacter persici
WP 048836254.1:1-223 Acetobacter orleanensis Acetobacter orleanensis
WP 062248623.1:1-224 Acetobacter cerevisiae Acetobacter cerevisiae
WP 086653646.1:1-225 Acetobacter malorum Acetobacter malorum
WP 166120266.1:1-223 Acetobacter farinalis Acetobacter farinalis
WP 042786441.1:1-223 Acetobacter pomorum Acetobacter pomorum
WP 124296244.1:1-223 Acetobacter pasteurianus Acetobacter pasteurianus
WP 166115653.1:1-223 Acetobacter lovaniensis Acetobacter lovaniensis
WP 048855043.1:1-236 Acetobacter syzygii Acetobacter syzygii
WP 086639793.1:1-223 Acetobacter okinawensis Acetobacter okinawensis
WP 095349471.1:1-223 Acetobacter fabarum Acetobacter fabarum
WP 165991212.1:1-223 Acetobacter lambici Acetobacter lambici


WP 059022713.1:1-224 Acetobacter ghanensis Acetobacter ghanensis
WP 011304925.1:1-235 Ehrlichia canis Ehrlichia canis
WP 045170889.1:1-235 Ehrlichia minasensis Ehrlichia minasensis
WP 045895782.1:1-249 Anaplasma phagocytophilum Anaplasma phagocytophilum
WP 045796935.1:1-234 Orientia chuto Orientia chuto
WP 045916039.1:1-234 Orientia tsutsugamushi Orientia tsutsugamushi
WP 019231088.1:1-247 Occidentia massiliensis Occidentia massiliensis
WP 012148222.1:1-226 Rickettsia canadensis Rickettsia canadensis
WP 040256817.1:1-226 Rickettsia hoogstraalii Rickettsia hoogstraalii
WP 012013313.1:1-226 Rickettsia akari Rickettsia akari
WP 014412109.1:1-226 Rickettsia australis Rickettsia australis
WP 004596556.1:1-237 Rickettsia prowazekii Rickettsia prowazekii
WP 045799065.1:1-237 Rickettsia bellii Rickettsia bellii
WP 023507426.1:1-226 Rickettsia monacensis Rickettsia monacensis
WP 032139287.1:1-226 Rickettsia tamurae Rickettsia tamurae
WP 014408162.1:1-237 Rickettsia rhipicephali Rickettsia rhipicephali
WP 032073798.1:1-226 Rickettsia aeschlimannii Rickettsia aeschlimannii
WP 016917576.1:1-226 Rickettsia honei Rickettsia honei
WP 016769841.1:1-226 Rickettsia sibirica Rickettsia sibirica
WP 016925924.1:1-226 Rickettsia conorii Rickettsia conorii
WP 096372719.1:1-226 Rickettsia japonica Rickettsia japonica
WP 014362718.1:1-226 Rickettsia rickettsii Rickettsia rickettsii
WP 027135853.1:1-225 Geminicoccus roseus Geminicoccus roseus
WP 088561589.1:1-223 Arboricoccus pini Arboricoccus pini
WP 088560731.1:1-227 Arboricoccus pini Arboricoccus pini
SDD86839.1:32-249 Kordiimonas lacus Kordiimonas lacus
WP 020400625.1:8-225 Kordiimonas gwangyangensis Kordiimonas gwangyangensis
WP 085902059.1:12-236 Kiloniella majae Kiloniella majae
WP 025897039.1:12-229 Sneathiella glossodoripedis Sneathiella glossodoripedis
RMB12130.1:7-227 Eilatimonas milleporae Eilatimonas milleporae
WP 020400207.1:1-222 Kordiimonas gwangyangensis Kordiimonas gwangyangensis
WP 013301812.1:3-222 Parvularcula bermudensis Parvularcula bermudensis
WP 013301838.1:1-222 Parvularcula bermudensis Parvularcula bermudensis
WP 013300658.1:1-225 Parvularcula bermudensis Parvularcula bermudensis
WP 031553661.1:1-225 Parvularcula oceani Parvularcula oceani
WP 085884107.1:1-222 Oceanibacterium hippocampi Oceanibacterium hippocampi
WP 121938706.1:1-226 Eilatimonas milleporae Eilatimonas milleporae
WP 085905575.1:1-225 Kiloniella majae Kiloniella majae
WP 085905509.1:1-223 Kiloniella majae Kiloniella majae
WP 020593623.1:1-223 Kiloniella laminariae Kiloniella laminariae
WP 046501666.1:1-223 Kiloniella litopenaei Kiloniella litopenaei
WP 046506632.1:1-218 Kiloniella litopenaei Kiloniella litopenaei
WP 085906071.1:1-218 Kiloniella majae Kiloniella majae
WP 020593406.1:1-218 Kiloniella laminariae Kiloniella laminariae
WP 085902204.1:1-214 Kiloniella majae Kiloniella majae
SCW60198.1:2-218 Asticcacaulis taihuensis Asticcacaulis taihuensis
WP 085905518.1:1-215 Kiloniella majae Kiloniella majae
WP 085340852.1:1-216 Aquidulcibacter paucihalophilus Aquidulcibacter paucihalophilus
ATC34496.1:1-214 Caulobacter vibrioides Caulobacter vibrioides
PZR33217.1:1-214 Caulobacter segnis Caulobacter segnis
WP 068151672.1:1-221 Kordiimonas lipolytica Kordiimonas lipolytica
WP 011712011.1:1-218 Magnetococcus marinus Magnetococcus marinus
WP 024080375.1:1-220 Magnetospirillum gryphiswaldense Magnetospirillum gryphiswaldense
WP 011390935.1:1-228 Rhodospirillum rubrum Rhodospirillum rubrum
WP 068151857.1:1-216 Kordiimonas lipolytica Kordiimonas lipolytica
WP 020399831.1:1-216 Kordiimonas gwangyangensis Kordiimonas gwangyangensis
WP 068304771.1:1-216 Kordiimonas lacus Kordiimonas lacus
WP 121937271.1:1-222 Eilatimonas milleporae Eilatimonas milleporae


WP 085552975.1:1-219 Azospirillum lipoferum Azospirillum lipoferum
WP 145627953.1:1-219 Azospirillum brasilense Azospirillum brasilense
WP 145678727.1:1-219 Azospirillum brasilense Azospirillum brasilense
WP 132938599.1:1-221 Varunaivibrio sulfuroxidans Varunaivibrio sulfuroxidans
WP 085883969.1:1-216 Oceanibacterium hippocampi Oceanibacterium hippocampi
WP 019961423.1:1-221 Woodsholea maritima Woodsholea maritima
WP 085339414.1:1-220 Aquidulcibacter paucihalophilus Aquidulcibacter paucihalophilus
WP 006273500.1:1-217 Asticcacaulis biprosthecium Asticcacaulis biprosthecium
WP 126421403.1:1-217 Asticcacaulis excentricus Asticcacaulis excentricus
WP 018081957.1:1-217 Asticcacaulis benevestitus Asticcacaulis benevestitus
WP 090646434.1:1-217 Asticcacaulis taihuensis Asticcacaulis taihuensis
WP 045319000.1:1-216 Brevundimonas abyssalis Brevundimonas abyssalis
WP 121481347.1:1-224 Brevundimonas naejangsanensis Brevundimonas naejangsanensis
WP 029087906.1:1-216 Brevundimonas aveniformis Brevundimonas aveniformis
ATC31113.1:1-216 Caulobacter vibrioides Caulobacter vibrioides
PZR32593.1:1-216 Caulobacter segnis Caulobacter segnis
WP 062144597.1:1-222 Caulobacter henricii Caulobacter henricii
WP 111456062.1:1-220 Phenylobacterium hankyongense Phenylobacterium hankyongense
WP 012523243.1:1-220 Phenylobacterium zucineum Phenylobacterium zucineum
WP 111514392.1:1-221 Phenylobacterium deserti Phenylobacterium deserti
WP 099621404.1:1-216 Caulobacter mirabilis Caulobacter mirabilis
WP 020400274.1:5-223 Kordiimonas gwangyangensis Kordiimonas gwangyangensis
WP 068151497.1:1-221 Kordiimonas lipolytica Kordiimonas lipolytica
WP 020400243.1:1-219 Kordiimonas gwangyangensis Kordiimonas gwangyangensis
WP 068303897.1:1-222 Kordiimonas lacus Kordiimonas lacus
WP 016770145.1:8-224 Rickettsia sibirica Rickettsia sibirica
WP 011713545.1:4-226 Magnetococcus marinus Magnetococcus marinus
WP 011713153.1:3-235 Magnetococcus marinus Magnetococcus marinus
WP 025899457.1:1-221 Sneathiella glossodoripedis Sneathiella glossodoripedis
WP 085882708.1:4-228 Oceanibacterium hippocampi Oceanibacterium hippocampi
WP 013300325.1:5-226 Parvularcula bermudensis Parvularcula bermudensis
WP 031554383.1:5-226 Parvularcula oceani Parvularcula oceani
WP 033405503.1:6-225 Kiloniella laminariae Kiloniella laminariae
WP 082114061.1:13-232 Kiloniella litopenaei Kiloniella litopenaei
WP 020398793.1:4-225 Kordiimonas gwangyangensis Kordiimonas gwangyangensis
WP 068145065.1:4-225 Kordiimonas lipolytica Kordiimonas lipolytica
WP 068305849.1:4-225 Kordiimonas lacus Kordiimonas lacus
WP 046503565.1:8-229 Kiloniella litopenaei Kiloniella litopenaei
WP 085901118.1:8-229 Kiloniella majae Kiloniella majae
WP 031555247.1:8-221 Parvularcula oceani Parvularcula oceani
WP 085899174.1:7-239 Kiloniella majae Kiloniella majae
WP 011713171.1:6-233 Magnetococcus marinus Magnetococcus marinus
WP 011714927.1:6-235 Magnetococcus marinus Magnetococcus marinus
SLN25798.1:11-239 Oceanibacterium hippocampi Oceanibacterium hippocampi
WP 068147903.1:3-233 Kordiimonas lipolytica Kordiimonas lipolytica
WP 011714286.1:32-257 Magnetococcus marinus Magnetococcus marinus
WP 011714738.1:4-227 Magnetococcus marinus Magnetococcus marinus
WP 011713550.1:8-241 Magnetococcus marinus Magnetococcus marinus
WP 046504482.1:13-252 Kiloniella litopenaei Kiloniella litopenaei
WP 085901478.1:13-252 Kiloniella majae Kiloniella majae
WP 020591171.1:13-251 Kiloniella laminariae Kiloniella laminariae
WP 085902476.1:5-232 Kiloniella majae Kiloniella majae
WP 020590626.1:5-232 Kiloniella laminariae Kiloniella laminariae
WP 046505978.1:5-232 Kiloniella litopenaei Kiloniella litopenaei
WP 085881900.1:2-228 Oceanibacterium hippocampi Oceanibacterium hippocampi
WP 025897434.1:4-231 Sneathiella glossodoripedis Sneathiella glossodoripedis
WP 013299845.1:3-231 Parvularcula bermudensis Parvularcula bermudensis
WP 121939795.1:5-245 Eilatimonas milleporae Eilatimonas milleporae
WP 025896950.1:4-234 Kordiimonas gwangyangensis Kordiimonas gwangyangensis
WP 068143305.1:4-235 Kordiimonas lipolytica Kordiimonas lipolytica
WP 068301095.1:4-235 Kordiimonas lacus Kordiimonas lacus