Characterization and Transcript Expression Studies of Interferon Regulatory Factors in Atlantic cod *(Gadus morhua)*

by

©Sabrina M. Inkpen

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
Faculty of Science

Memorial University of Newfoundland

September 2014

St. John's Newfoundland and Labrador

Abstract

The interferon regulatory factor (IRF) family of genes encode a group of transcription factors which have important roles not only in regulating the expression of Type I interferons (IFNs) and other genes in the interferon pathway, but also in growth, development and regulation of oncogenesis. In this study, several IRF family members in Atlantic cod (*Gadus morhua*) were characterized at the cDNA and putative amino acid level, allowing for phylogenetic analysis of these genes in teleost fish, and the development of paralogue specific PCR primers which were used in semi-quantitative RT-PCR and Quantitative PCR (QPCR) analyses. Two Atlantic cod *Irf10* splice variants were identified and named *Irf10-v1* and *Irf10-v2*, and their presence was confirmed by sequencing of the *Irf10* genomic region. RT-PCR showed that *Irf7, Irf8* and both *Irf10* transcripts were detected in 15 cod tissues, while *Irf4a* and *Irf4b* appeared to be absent in some tissues. RT-PCR in embryo and larval samples showed unique transcript expression profiles of IRFs during development and indicated potential stage specific roles that will be investigated in future studies. QPCR analysis of spleen expression expanded upon previous studies, confirming that all transcripts were responsive to stimulation by the viral mimic poly(I:C) and showing that all except *Irf4a* were responsive to killed *Aeromonas salmonicida* (ASAL). Temperature was observed to affect the responsiveness of all except *Irf4a* to poly(I:C) and/or ASAL, supporting earlier studies. The effect of increased temperature on immune responsiveness to pathogens is of particular interest to Atlantic cod aquaculture in Newfoundland, where fish experience seasonal fluctuations in temperature.

Acknowledgements

I would first like to say a huge thank you to my supervisor Dr. Matthew Rise for the opportunity to carry out this research, and for his guidance and especially his patience while this thesis slowly came together.

Thank you as well to the members of my supervisory committee, Dr. Brian Staveley and Dr. Ross McGowan for all of their advice throughout this project.

Thanks to past and present members of the Rise lab and the OSC community for all of their help during the past two and a half years, especially to Jennifer Hall and Dr. Tiago Hori for training me in the techniques used throughout this research. Thanks as well to the staff of JBARB for their help with sample collection.

I would also like to thank the Memorial University of Newfoundland School of Graduate Studies for providing me with a Fellowship, to the Department of Biology for several Graduate Assistantships and the opportunity to gain experience as a Teaching Assistant, and to the National Sciences and Engineering Council of Canada for awarding a Discovery Grant to Dr. Matthew Rise, making this research possible.

Lastly, thank you to my friends and family, especially my parents for all of their support throughout my academic career and in everything.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AA – Amino acid
ANOVA – Analysis of variance
arp – actin-related protein 2/3 complex subunit 4
ASAL – Formalin-killed *Aeromonas salmonicida*
Bl -Blood
BLAST – Basic local alignment search tool
bp – Base pair
Br – Brain
°C – Degree centigrade
C – Control
CGP – Atlantic cod Genomics and Broodstock Development Project
C_T – Threshold cycle
cDNA – Complimentary DNA
DBD – DNA binding domain
DNA – Deoxyribonucleic acid
dpf – Days post-fertilization
EF1α – Elongation Factor 1 alpha
EST – Expressed sequence tag
EtBr –Ethidium bromide
Ey - Eye
Fig – Figure
GaP – Genomics and proteomics facility
Gi - Gill
GSP – Gene specific primer
g – Grams
HG - Hindgut
HK – Head kidney
HPI – Hours post-injection
Hr - Heart
H – Hour
IFN – Interferon
IRF – Interferon Regulatory Factor
IP – Intraperitoneal
ISG – Interferon stimulated gene
Ig – Immunoglobulin
LB – Luria-Bertani medium
LPS – Lipopolysaccharide
Li – Liver
L – Litre
MG - Midgut
Mu – Skeletal muscle
MH – Major histocompatibility gene
MMLV – Moloney murine leukemia virus
mg – Milligram

min – Minute
mL – Millilitre
mRNA – Messenger RNA
NCBI – National Center for Biotechnology Information
NTC – No template control
ng – Nanogram
nM – Nanomolar
nr – Non-redundant
OSC – Ocean Sciences Centre
ORF – Open reading frame
PBS – Phosphate buffered saline
PC – Pyloric caecum
PCR – Polymerase chain reaction
PK – Posterior kidney
Poly(I:C) – Polyriboinosinic polyribocytidylic acid
p – p-value
QC – Quality check
QPCR – Quantitative reverse transcription PCR
RACE – Rapid amplification of cDNA ends
RNA – Ribonucleic acid
RT – Reverse transcription
RQ – Relative quantity
SE – Standard error of the mean
s – Seconds
Sk - Skin
SNP – Single nucleotide polymorphism
Sp - Spleen
SSH – Suppression subtractive hybridization
St - Stomach
TLR – Toll-like receptor
TMS – Tricaine methane sulphonate
U - Units
UTR – Untranslated region
µg – Microgram
µL - Microlitre
µm – Micrometre

Co-Authorship Statement

For this thesis, experimental design was planned by myself and Dr. M.L. Rise, and I was primarily responsible for implementation of experiments, data analysis and manuscript preparation. Exceptions include: a) experimental design and sampling for the spleen transcript expression response experiment was carried out by Dr. Tiago Hori, Dr. A. Kurt Gamperl, Gordon Nash and Dr. Matthew L. Rise as part of a previous set of experiments; and b) sampling and RNA extraction for the juvenile cod tissue panel study was carried out in cooperation with Xi Xue (Ocean Sciences Centre). The contents of this thesis (excluding developmental series RT-PCR) will be submitted for publication in a manuscript with authors SM Inkpen, TS Hori, AK Gamperl, GW Nash, and ML Rise, in preparation for submission to Fish and Shellfish Immunology.

1. Introduction

1.1 Importance of immunological research in Atlantic cod

A thorough understanding of fish molecular immunology is of great importance to research in various areas, including comparative vertebrate immunology, fisheries and aquaculture. For example, the study of genes and pathways involved in innate and adaptive immune responses and stress responses of fishes should aid in the development of tools and methods (e.g. molecular tests, vaccines, therapeutics) to help reduce disease and stress in cultured fish (Booman and Rise, 2012). The identification of fish genes that are involved in defense responses could also lead to the development of molecular markers [e.g. single nucleotide polymorphisms (SNPs) in trait-relevant genes] for selection of aquaculture broodstock with desirable traits such as resistance to pathogens or environmental stress (Booman and Rise, 2012)]. With the depletion of some wild stocks of Atlantic cod (*Gadus morhua*), for example in Newfoundland (Marteinsdottir *et al*., 2005), such developments will be particularly valuable in creating a successful farming industry for the species. Although cod aquaculture has been of interest in several countries (e.g. Canada, Norway, and Iceland) for some time, the development of successful hatchery and culture methods has been slow (Brown *et al.,* 2003; Rosenlund and Hallorsson, 2007), and many challenges still exist. For example, normal aquaculture methods induce stress for fish, from routine handling (Brown *et al.*, 2003) to exposure to variable temperatures in sea cages (Gollock *et al.*, 2006). Recent research showing that Atlantic cod stress and immune responses are affected by increasing temperature (Perez-Casanova *et al.*, 2008; Hori *et al.*, 2012) suggests that fluctuating temperatures in sea

cages can impact cod immune system function and responses to pathogens and other stressors. Further study of the structure, regulation, and function of immune-relevant genes involved in these responses is required to overcome such challenges.

Genomics resources such as DNA microarrays and sequence databases for Atlantic cod have increased dramatically in recent years. Currently, there are 57,041 sequences in the non-redundant nucleotide (nt) database, 257,453 in the expressed sequence tag database (dbEST) and 2,896 in the protein database of GenBank for this species (NCBI, 2014). The construction and sequencing of multiple normalized and suppression subtractive hybridization (SSH) cDNA libraries representing various life stages, tissues and treatments (Bowman *et al.,* 2011), the development of microarray platforms [e.g. a 20,000 gene (20 K) oligonucleotide microarray (Booman *et al.,* 2011)] and the sequencing of the Atlantic cod genome (Star *et al.,* 2011) have allowed for a wide range of functional genomics research in this species. This growing genomic knowledge base makes Atlantic cod an excellent species in which to study the developing fish immune system at the genetic level. Furthermore, while Atlantic cod develop more slowly than zebrafish (*Danio rerio*, a common research model for developmental biology and genetics), cod have transparent embryos/larvae and are highly fecund, making them particularly suitable for developmental studies (Hall *et al.,* 2004). Several studies indicate the Atlantic cod immune system is unique among teleosts and among vertebrates in general, showing higher serum levels of immunoglobulin M than other teleosts, as well as a relatively low antibody response to pathogens (reviewed in Solem and Stenvik, 2006; Star *et al.,* 2011). Sequencing and analysis of the Atlantic cod genome indicated the

species has approximately 100 major histocompatibility (MH) class I loci, a much higher number than other teleosts [e.g. an estimated 14 in stickleback (*Gasterosteus aculeatus*)]. That study also provided evidence for the loss of several important immune-relevant genes [e.g. MH class II, invariant chain (Ii), and the MH II-interacting protein CD4], suggesting a loss of function of the classical pathway for adaptive immunity in Atlantic cod (Star *et al.,* 2011). These unusual characteristics make further study of the genes and molecular pathways involved in cod immune responses, and the evolution of immune-related gene families in cod of great interest to researchers in areas such as comparative immunology and evolutionary biology.

1.2 The interferon pathway and interferon regulatory factors

In fish, as well as in all other vertebrates, secreted proteins called interferons (IFNs) play important roles in the innate immune response to viral pathogens (Robertsen, 2006; Rise *et al.,* 2008). IFNs are divided into two families, Type I and Type II, based on structural properties and functions. As part of the cellular response to viral infection, Type I IFNs (IFNα and IFNβ) are secreted and bind to specific receptors on other cells, activating the JAK-STAT (Janus kinase-signal transducer and activator of transcription) signal transduction pathway and leading to the transcription of many downstream genes (Barnes *et al.,* 2002; Robertsen, 2006; Rise *et al.,* 2008). Currently, the genes and mechanisms involved in this IFN pathway are better understood in humans and other mammals than in fish, although our knowledge of the molecular basis of fish antiviral responses has been increasing since the identification of the first fish IFN genes in 2003 (Altmann *et al.,* 2003; Lutfalla *et al.,* 2003; Robertsen *et al.,* 2003). As both wild and

cultured fish are susceptible to viruses such as infectious salmon anemia virus (ISAV) and nodavirus (Lang *et al.,* 2009 and references therein), the study of fish antiviral responses, and in particular the genes involved in the IFN pathway, will be of value to both fisheries and aquaculture. While several groups have investigated fish gene and protein expression responses to viral infection, most of these studies have involved later life stage fish (Workenhe *et al.,* 2010; Verrier *et al.,* 2011), and less is known about how fish embryos/larvae defend themselves against viral infections. Recent work on early life stage Atlantic cod in the Rise lab has fully or partially characterized several virus-responsive transcripts and has shown that some of them [e.g. interferon regulatory factor (*Irf*)*1, Irf7*] have dynamic mRNA expression profiles during embryonic development (Rise *et al.,* 2008; Rise *et al.,* 2012). The study of other cod IRF genes, and the comparison of cod IRF gene structure and expression with orthologous genes in other teleost species, will be of interest to determine potential functions of these genes as well as to examine the expansion and diversification of the gene family through evolutionary history.

Genes in the IRF family encode transcription factors which either positively or negatively regulate the expression of IFN genes, and thus are vital to the cellular antiviral response. Nine IRF genes (*Irf1-Irf9*) have been described in most vertebrates, although a tenth (*Irf10*) is present in several avian and fish species, and another potential family member (*Irf11* or *Irf1b*) has been identified in zebrafish and other teleost fish (Stein *et al.,* 2007; Huang *et al.,* 2010). All IRF proteins share a conserved amino (N) terminus DNA-binding domain (DBD) of about 115 amino acids, containing five conserved tryptophan

(Trp) residues and forming a helix-loop-helix motif (Taniguchi *et al.*, 2001). The DBD recognizes the interferon stimulated response element (ISRE) DNA sequence, which has the consensus sequence A/GNGAAANNGAAACT (Darnell *et al.*, 1994), and is found in the promoters of Type I IFNs and many genes induced by Type I IFNs [e.g. interferon stimulated genes (ISGs)]. The carboxyl (C) terminus of each IRF family member contains one of two types of association modules, called IRF associated domain 1 (IAD1; in all IRFs except *Irf1*and *Irf2*), and IAD2 (found in *Irf1*and *Irf2*; Savitsky *et al.*, 2010). Outside the IAD, the C-terminus is not well conserved, and thus is the region that gives each IRF specific functions.

1.3 Recent progress in understanding interferon regulatory factors

The roles of proteins encoded by IRF family genes have been quite well-studied in mammals, and found to include not only regulation of IFN expression, but also various aspects of immune system regulation, growth, development, and regulation of oncogenesis (for reviews see Honda and Taniguchi, 2006; Ozato *et al.*, 2007; and Savitsky *et al.*, 2010). For example, IRF1, IRF3 and IRF7 are known to induce transcription of type I IFN genes in mice and in mammalian cell lines, whereas IRF2 is a negative regulator of the IFN response in mammals (Taniguchi *et al.*, 2001 and references therein). IRF9 acts as part of a transcriptional activator complex stimulated by type I IFN which activates several IFN pathway genes (Taniguchi *et al.*, 2001). While the role of IRF6 in immune regulation has not been determined, this gene has been shown to be important to development in several vertebrate species, as discussed below.

The majority of IRF research thus far has been focused on mammalian species, although investigation into this gene family in multiple teleost species has increased in recent years [e.g. in mandarin fish (*Siniperca chuasti*) (Sun *et al.*, 2007), rainbow trout (*Oncorhynchus mykiss*) (Holland *et al.*, 2008), Atlantic salmon (*Salmo salar*) (Bergan *et al.*, 2010) and rock bream (*Oplegnathus fasciatus*) (Bathige *et al.*, 2012]. Studies of IRF family genes involving zebrafish as a model fish species have so far included analysis of gene structure based on mining of public sequence databases (Nehyba *et al.*, 2009; Huang *et al.*, 2010), investigation of function in selected genes using morpholino-based targeted gene knockdown (Sabel *et al.*, 2009; Li *et al.*, 2011), and expression studies of selected paralogues (Ben *et al.*, 2005; Xiang *et al.*, 2010). Studies in various species show that, as expected, the IRF family members that are most closely related (based on sequence comparison) often share similar functions.

1.4 Interferon regulatory factor gene family sub-groups

Based on molecular phylogenetic analysis, the IRF gene family can be divided into four sub-groups: IRF1-G (*Irf1*, *Irf2*), IRF3-G (*Irf3*, *Irf7*), IRF4-G (*Irf4*, *Irf8*, *Irf9*, *Irf10*), and IRF5-G (*Irf5*, *Irf6*), reflecting expansion and diversification over evolutionary history (Nehyba *et al.*, 2002; 2009). As indicated in Figure 1, IRF1-G may also be referred to as IRF1 supergroup (SG) while all other IRFs are grouped as IRF4-SG, mainly based on the presence of the well-conserved IAD1 in the carboxyl terminus of the latter group.

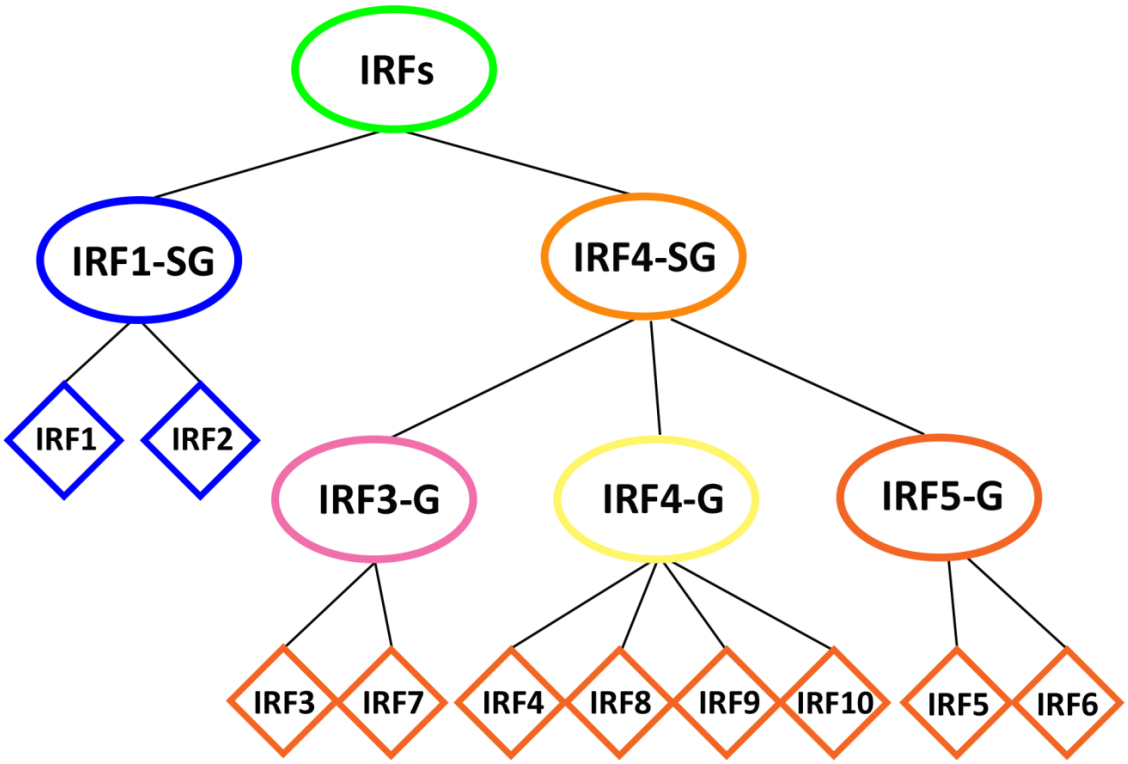


Figure 1: Summary of Interferon Regulatory Factor gene family organization. Schematic based on phylogenetic analysis by Nehyba *et al.*, (2002), in which IRF protein sequences from human, chicken, clawed toad, Japanese flounder, mouse, quail, rat, sheep, and fugu were aligned and used to construct a neighbour-joining tree. (See Fig, 2; Table 2, Nehyba *et al*., 2002).

1.4.1 IRF1 sub-group

IRF1 (named because it was the first of the family to be identified) is a transcriptional activator of IFNα/β expressed in most cell types and tissues, whose expression can be induced by IFNs and many other cytokines, or by viral infection (reviewed in Taniguchi *et al.,* 2001). In addition to its role in the innate immune response, IRF1 is required for DNA damage-induced apoptosis, and is thus known as a tumor suppressor (Tanaka *et al.,* 1996). IRF2 can be said to act opposite to IRF1, negatively regulating type I IFN responses (Honda and Taniguchi 2006) and has been shown to have pro-oncogenic activity (reviewed in Yanai *et al.,* 2012), indicating an opposing role to IRF1 in oncogenesis as well.

*Irf1*and *Irf2* cDNA sequences have been partially or fully characterized in several fish species, including Atlantic salmon (Bergan *et al.,* 2010) and the paddlefish *Polyodon spathula* (Xiaoni *et al.,* 2011), and were upregulated in each of these species by stimulation with polyriboinosinic polyribocytidylic acid [poly(I:C)], a synthetic double-stranded RNA which mimics a viral infection. *Irf1* is the only IRF gene in Atlantic cod that was fully characterized at the cDNA level (Feng *et al.,* 2009) prior to the current study, and spleen transcript expression was previously found to be upregulated by both the viral mimic poly(I:C) and bacterial antigens (formalin-killed *Aeromonas salmonicida*) (Rise *et al.,* 2008; Feng *et al.,* 2009). Table 1 and Table 2 summarize current knowledge of expression of *Irf1* and *Irf 2* (and all other family members) expression in mammalian species and fish species, respectively.

Table 1: Studies of Interferon Regulatory Factor protein expression and function in mammalian species

Paralogue	Expression	Roles in innate immunity
IRF1	Human: constitutive in many cell types; upregulated by viral infection or IFN stimulation (Taniguchi *et al.*, 2001, Savitsky *et al.*, 2010).	Mouse: inhibits immunosuppressive features of dendritic cells (Gabriele and Ozato, 2007). Activates transcription of type I IFNs (Taniguchi *et al.*, 2001).
IRF2	Human: constitutive in many cell types; upregulated by IFN stimulation (Taniguchi *et al.*, Savitsky *et al.*, 2010).	Human: attenuates type I IFN responses by antagonizing IRF1and IRF9 (Savitsky *et al.*, 2010).
IRF3	Human: constitutively expressed in all tissues (Au *et al.*, 1995).	Human: activates transcription of type I IFNs and other cytokines (Savitsky *et al.*, 2010). Mouse: triggers necrotic cell death of macrophages in response to infection (Di Paolo *et al.*, 2013).
IRF4	Mouse: constitutive only in lymphoid cells (Matsuyama *et al.*, 1995). Human: constitutive in lymphocytes (Taniguchi *et al.*, 2001).	Mouse: regulates myeloid/lymphoid cell differentiation (Gabriele and Ozato, 2007); negatively regulates Toll-like receptor (TLR) signalling (Negishi *et al.*, 2005); required for B cell differentiation into plasma cells (Sciammas *et al*., 2006).
IRF5	Human: constitutive in B-cells and dendritic cells; inducible in other lymphoid cells by IFN (Barnes *et al.*, 2002).	Human: activates transcription of type I IFNs and other cytokines (Takaoka *et al.*, 2005). Mouse: important to B-cell differentiation and maturation (Lien *et al.*, 2010).
IRF6	Human: constitutively expressed in skin (Savitsky *et al.*, 2010).	Human: important to development of the lip and palate; involved in development of skin and external genitalia (Kondo *et al.*, 2002).
IRF7	Human: ubiquitous but predominantly in lymphoid cells; dependant on IFN signaling (Taniguchi *et al.*, 2001, Barnes *et al.*, 2002).	Human: activates transcription of type I IFNs and other cytokines (Taniguchi *et al.*, 2001). Mouse: main regulator of IFN production in plasmacytoid dendritic cells (Honda *et al.*,, 2005); required for differentiation of medullary thymic epithelial cells (Otero *et al.*, 2013)
IRF8	Human: lymphoid and myeloid cell lineages (Taniguchi *et al.*, 2001). Mouse: constitutively expressed in B cells (Nelson *et al.*, 1996).	Mouse: regulates myeloid cell differentiation (Tamura and Ozato, 2002); contributes to high IFN induction in dendritic cells (Gabriele and Ozato, 2007); functions in microglia development in the CNS (Minten *et al*., 2012).
IRF9	Human: constitutive in many cell types (Taniguchi *et al.*, 2001, Savitsky *et al.*, 2010).	Human: activated by type I IFN signaling; part of ISGF3 complex (Savitsky *et al.*, 2010).
IRF10	*not found in mammalian species	

Table 2: Studies of interferon regulatory factor transcript expression and response to immune stimulation in fish species

Paralogue	Constitutive Transcript Expression	Effect of Poly(I:C) / other treatments in fish on transcript expression
Irf1	Paddlefish: constitutively expressed in various tissues (Xiaoni *et al.*, 2012). Yellow croaker (*Pseudosciaena crocea*): constitutively expressed in various tissues; highly expressed in gill and spleen (Yao *et al.*, 2010). Mandarin fish: constitutively expressed in various tissues (Sun *et al.*, 2007). Atlantic cod: expressed throughout development with peak in early segmentation (Rise *et al.*, 2012).	Paddlefish: upregulated by poly(I:C) in gill, head kidney, trunk kidney, liver and spleen (Xiaoni *et al.*, 2011). Yellow croaker: upregulated by poly(I:C) and lipopolysaccharide (LPS) in blood, spleen and liver (Yao *et al.*, 2010). Atlantic cod: upregulated by poly(I:C) and killed *A. salmonicida* (ASAL) in spleen (Rise *et al.*, 2008; Feng *et al.*, 2009); response is affected by elevated temperature (Hori *et al.*, 2012). Atlantic salmon: upregulated by poly(I:C) in head kidney cells (Bergan *et al.*, 2010).
Irf2	Paddlefish: constitutively expressed in various tissues (Xiaoni *et al.*, 2012).	Paddlefish: upregulated by poly(I:C) in gill, head kidney, trunk kidney, liver and spleen (Xiaoni *et al.*, 2011). Atlantic salmon: upregulated by poly(I:C) in head kidney cells (Bergan *et al.*, 2010).
Irf3	Turbot (*Scophthalmus maximus*); Japanese flounder (*Paralichthys olivaceus*): constitutively expressed in various tissues; highly expressed in spleen and head kidney (Hu *et al.*, 2011a;b).	Carp: upregulated by poly(I:C) and IFN inducersin cell lines (Sun *et al.*, 2010). Turbot: upregulated by poly(I:C) and turbot reddish body iridovirus (TRBIV) in spleen, head kidney and gills (Hu *et al.*, 2011a). Japanese flounder: upregulated by poly(I:C) in head kidney and gill (Hu *et al.*, 2011b). Trout: upregulated by poly(I:C) in cell lines (Holland *et al.*, 2008). Atlantic salmon: upregulated by poly(I:C) in head kidney cells (Bergan *et al.*, 2010).
Irf4	Trout: highest expression in spleen, head kidney, gills (Holland *et al.*, 2010). Rock bream: constitutive expression in various tissues; highest in blood and spleen (Bathige *et al.*, 2012).	Rock bream: upregulated by *Edwardsiella. tarda* (Gram negative bacterium) but downregulated by LPS in head kidney and spleen (Bathige *et al.*, 2012). Trout: downregulated by LPS; no response to poly(I:C) in splenocytes (Holland *et al.*, 2010).

Irf5	Grass carp (*Ctenopharyngodon idellus*); paddlefish: constitutively expressed in various tissues (Xu *et al.*, 2010; Xiaoni *et al.*, 2012).	Turbot: not upregulated by poly(I:C); upregulated by turbot reddish body iridovirus in gill, head kidney, spleen and muscle (Xia *et al.*, 2012). Paddlefish: not upregulated by poly(I:C) in gill, head kidney, liver, or spleen (Xiaoni *et al.*, 2012). Grass carp: induced by grass carp reovirus in spleen and head kidney (Xu *et al.*, 2010).
Irf6	Zebrafish: maternal transcript in egg; epithelial cells of endoderm derived tissues in larvae (Ben *et al.*, 2005).	*no data available
Irf7	Orange spotted grouper (*Epinephelus coioides)*; turbot: constitutively expressed in various tissues (highly in spleen and kidney) (Cui *et al.*, 2011; Hu *et al.*, 2011c). Japanese flounder: constitutively expressed in various tissues (Hu *et al.*, 2010). Mandarin fish: constitutively expressed in various tissues (Sun *et al.*, 2007). Atlantic cod: expressed in unfertilized eggs and throughout development with peak in early segmentation (Rise *et al.*, 2012).	Orange-spotted grouper: upregulated by *Vibrio. vulnificus* and Singapore grouper iridovirus (SGIV) in spleen (Cui *et al.*, 2011). Turbot: upregulated by TRBIV in head kidney (Hu *et al.*, 2011c). Japanese flounder: upregulated by poly(I:C) in head kidney and gill (Hu *et al.*, 2010). Trout: upregulated by poly(I:C) in cell lines (Holland *et al.*, 2008). Atlantic cod: upregulated by poly(I:C) in spleen (Rise *et al.*, 2008); response is affected by elevated temperature (Hori *et al.*, 2012); upregulated by nervous necrosis virus in brain (Krasnov *et al.*, 2013). Atlantic salmon: upregulated by poly(I:C) in head kidney cells (Bergan *et al.*, 2010).
Irf8	Trout: highest expression in spleen, head kidney, and gills (Holland *et al.*, 2010). Rock bream: constitutively expressed in various tissues (Bathige *et al.*, 2012). Japanese flounder: constitutively expressed in various tissues (Hu *et al.*, 2013).	Trout: upregulated by poly(I:C) in splenocytes (Holland *et al.*, 2010). Rock bream: upregulated by poly(I:C) and bacterial infection in head kidney and spleen (Bathige *et al.*, 2012). Japanese flounder: upregulated by poly(I:C) and lymphocystis disease virus in spleen (Hu *et al.*, 2013).
Irf9	Crucian carp (*Carassius auratus*): expressed in blastulae embryonic cells (Shi *et al.*, 2012).	*no data available
Irf10	Japanese flounder: constitutively expressed in various tissues (Suzuki *et al.*, 2011).	Japanese flounder: upregulated by LPS, poly (I:C), and several pathogens in peripheral blood lymphocytes (Suzuki *et al.*, 2011) Atlantic cod: upregulated by poly(I:C) in spleen (Rise *et al.*, 2008); response is affected by elevated temperature (Hori *et al.*, 2012).

1.4.2 IRF3 sub-group

IRF3 and IRF7 are both important regulators of type I IFN antiviral response, and can act individually or as part of a heterodimer or homodimer with each other, with differing effects (reviewed in Honda and Taniguchi, 2006). IRF7 is known as a master regulator of the IFN response, and is essential for the induction of IFN α/β genes (Honda *et al.*, 2005). It also plays a role in the regulation of oncogenesis, acting to prevent metastasis, while IRF3 is thought to have a role in mediating virus-induced apoptosis (Yanai *et al.*, 2012). *Irf3* and *Irf7* cDNA sequences have been characterized in several fish species, including rainbow trout (Holland *et al.*, 2008), Atlantic salmon (Bergan *et al.*, 2010), Japanese flounder (Hu *et al.*, 2010; 2011a), and turbot (Hu *et al.*, 2011b); and transcript expression was observed to be upregulated in response to poly(I:C) stimulation in several tissues in these species, as described in Table 2.

1.4.3 IRF4 sub-group

In mammals, IRF4 (also called multiple myeloma oncogene 1, MUM1) and IRF8 (also called interferon consensus sequence binding protein, ICSBP) have been shown to have important roles in the differentiation and development of dendritic cells (Gabriele and Ozato, 2007). While several mammalian IRFs are constitutively expressed in all cell types (see Table 1), the IRF4 protein in mammals only appears to be expressed in lymphocytes, playing an important role in development and function of those cells (reviewed in Taniguchi *et al.*, 2001), and the murine IRF8 protein is expressed only in myeloid and lymphoid cell lineages (Nelson *et al.*, 1996). The roles of these genes appear to be similar in fish; for example, *Irf8* has been shown to regulate myeloid lineage

differentiation during zebrafish development (Li *et al.*, 2011). *Irf4* and *Irf8* have been characterized at the cDNA level in several teleosts including rock bream (Bathige *et al.*, 2012) and rainbow trout (Holland *et al.*, 2010), and mRNA expression was seen to be upregulated in response to viral and bacterial stimulation in some species (as summarized in Table 2).

Irf10, also closely related to *Irf4/Irf8*, has not been found in mammals and is thus less well-studied than the other family members. This gene was first identified in chicken, where transcript expression was observed to be highest in cells of hematopoietic origin based on Northern blot analysis (Nehyba *et al.*, 2002). *Irf10* has been identified in several fish species, including zebrafish, stickleback, pufferfish and Atlantic cod (Stein *et al.*, 2007; Rise *et al.*, 2008; Huang *et al.*, 2010); but to our knowledge the complete cDNA has only been characterized in the Japanese flounder, *Paralichthys olivaceus* (Suzuki *et al.*, 2011), where *Irf10* mRNA expression was found to be upregulated in peripheral blood lymphocytes in response to both bacterial and viral stimulation.

1.4.4 IRF5 sub-group

In many species, the IRF6 protein is known to play a crucial role in the differentiation of epithelia. Mutations in human *Irf6* leads to Van der Woude syndrome, or cleft palate (Kondo *et al.*, 2002), and in zebrafish and the frog *Xenopus laevis Irf6* has been shown to be a maternal transcript necessary for epithelial differentiation (Ben *et al.*, 2005; Sabel *et al.*, 2009). This gene has been shown in humans to have a potential role in tumor suppression (Restivo *et al.*, 2011), but is the only IRF family member without a known role in innate immunity.

In mammals, IRF5 is known to function in Toll-like receptor (TLR) signalling, acting downstream of TLR stimulation as an inducer of pro-inflammatory cytokines (Takaoka *et al.*, 2005), and also plays an important role in B-cell differentiation (Lien *et al.*, 2010). Genetic variations (e.g. SNPs) in human *Irf5* have also been associated with the pathogenesis of systemic lupus erythematosus (SLE), a complex autoimmune disease (Cham *et al.*, 2012). The *Irf5* cDNA sequence has been characterized in several fish species, including turbot (Xia *et al.*, 2012) and Japanese flounder (Hu *et al.*, 2012), where its transcript expression was upregulated in response to viral stimulation, as described in Table 2.

1.5 Research Objectives

Knowledge of several IRF family genes in Atlantic cod has been increasing, particularly in terms of their response to immune stimulation (Rise *et al.*, 2008; Hori *et al.*, 2012), but these genes were still largely uncharacterized prior to the current research. cDNA libraries generated as part of the Atlantic Cod Genomics and Broodstock Development Project (CGP, *http://codgene.ca*; Bowman *et al.*, 2011) provided EST evidence for cod orthologues of *Irf1, Irf4, Irf7, Irf8* and *Irf10*, but as previously mentioned, only *Irf1* had been characterized at the cDNA and hypothetical amino acid level in this species prior to the current study (Feng *et al.*, 2009). *Irf1*, *Irf7*, and *Irf10* had been shown to respond to stimulation with viral mimic poly(I:C) with increased transcription (Rise *et al.*, 2008), and interestingly this response was seen to be modulated by temperature change (Hori *et al.*, 2012). *Irf1, Irf4* and *Irf7* have also been shown to respond to nervous necrosis virus infection in the brain, based on microarray analysis

(Krasnov *et al.*, 2013). Investigation of developmental transcript expression of *Irf1* and *Irf7* has also indicated a possible stage-specific function for these genes during embryogenesis (Rise *et al.*, 2012).

To further our knowledge of the molecular immunology of teleost fish, the goals of this research have been to characterize several Atlantic cod IRF genes (specifically *Irf4, Irf7, Irf8, and Irf10*) at the cDNA and hypothetical amino acid levels, and investigate the mRNA expression of these genes throughout embryonic development, in adult tissues, and in response to viral and bacterial stimulation and temperature change. A better understanding of how these genes are expressed should help in the determination of possible novel roles of IRF family members, for example during early development. Bioinformatics analysis and molecular techniques such as rapid amplification of cDNA ends (RACE), reverse transcription - polymerase chain reaction (RT-PCR) and quantitative real-time PCR (QPCR) were used to carry out these objectives, while phylogenetic analyses were also used to compare the evolutionary history of this gene family in Atlantic cod and other teleost fish species.

2. Methods

2.1 cDNA characterization of selected cod IRF paralogues

2.1.1 Database mining and RACE

A simplified schematic outlining the steps taken for cDNA characterization is shown in Figure 2. Briefly, bioinformatics tools and genomics resources [BLASTn and tBLASTn searches of dbEST using *Danio rerio* IRF (protein and cDNA) sequences; collection of predicted Atlantic cod cDNA sequences from Ensembl database (*www.ensembl.org*); search of the CGP database (*www.codgene.ca*) for Atlantic cod IRF-like sequences] were used to compile partial nucleotide sequences for all cod IRF paralogues. EST evidence for *Irf4*, *Irf7*, *Irf8* and *Irf10* was used to design paralogue-specific RACE primers. Since cod *Irf4, Irf7* and *Irf10* had previously been subjected to transcript expression analyses (Rise *et al.,* 2008; 2012; Hori *et al.,* 2012; Krasnov *et al.,* 2013), and *Irf8* is part of the same sub-family as *Irf4/10* (IRF4-G) and is known to have important roles in other species (see Table 1), these four paralogues were chosen for the main focus of this research. Partial predicted sequences were also available in the Ensembl database (*www.ensembl.org*) for cod *Irf2*, *Irf3*, *Irf5*, *Irf6*, and *Irf9*, although EST evidence for these genes was not found in dbEST. In continuation of the current research, these predicted sequences may be used to carry out RACE and TA cloning/sequencing of the remaining potential Atlantic cod IRF paralogues.

To obtain cDNA to be used in RACE, column-purified RNA was pooled using 5 µg from each of 10 spleen samples from fish injected with poly(I:C) [sampled at 24 hours

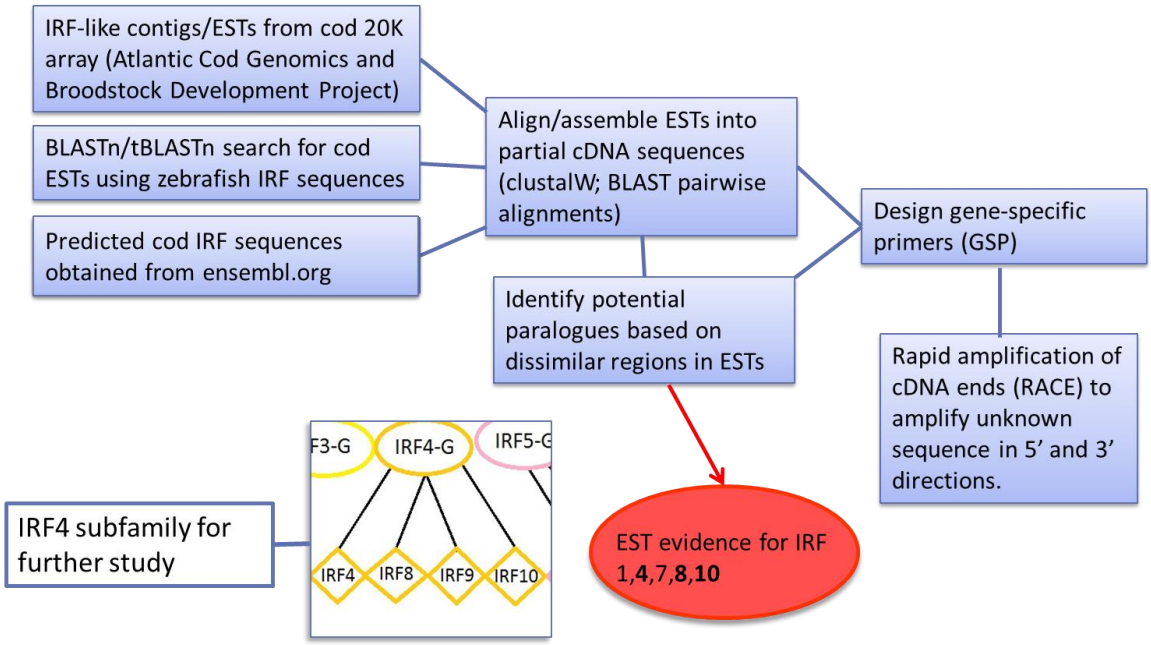


Figure 2: Steps taken to identify cod IRF paralogues and choose targets for cDNA characterization.

post injection (HPI), 5 at 10°C and 5 at 16°C]. Experimental setup and sampling procedure (Hori *et al.,* 2012), and RNA preparation are described in section 2.3.1. Five µg of pooled RNA was used to prepare RACE-ready cDNA using the GeneRacer Kit (Invitrogen, Burlington ON), according to the manufacturer's instructions. PCR amplification of cDNA ends was carried out in 50 µL reactions containing 1 µL (1 U/µL) Dynazyme polymerase (Thermo Scientific, Ottawa, ON), Dynazyme EXT buffer (1X final concentration, and either reverse gene specific primer (GSP) and GeneRacer 5' primer or forward GSP and GeneRacer 3' primer for 5' RACE and 3'RACE respectively. Primers used for RACE are listed in Table 3. Touchdown PCR was carried out using an initial denaturation at 94°C for 2 min followed by 40 cycles of [30s at 94°C; 30 s at 70°C →60°C, decreasing 0.3°C per cycle; 2 min at 72°C] and a final extension of 8 min at 72°C. Approximate size of PCR products was verified by electrophoresis on 1% agarose/tris acetate-EDTA (TAE) buffer gels stained with ethidium bromide, and DNA bands were excised under UV transillumination using a sterile scalpel blade and purified using QIAquick Gel Extraction Kit (QIAGEN, Mississauga, ON) according to manufacturer's instructions.

2.1.2 TA cloning and sequencing

RACE products were ligated into pGEM-T Easy vector (Promega, Madison, WI, USA) in 10 µL reactions containing 5 µL ligation buffer, 50 ng insert DNA, 1 µL vector and 1 µL ligase (3 U/µL), with incubation at 4°C overnight. Two µL of the ligation reaction was added to 50 µL Subcloning Efficiency DH5α chemically competent cells (Invitrogen, Burlington, ON.) and transformations were carried out according to

Table 3: Primers used for cDNA characterization of cod IRF genes

Primer Name	Sequence (5'-3')	Application	Predicted amplicon size
IRF4-gsp-fwd*	GATGGGTCACGACGGCCTGTAT	3'RACE	N/A
IRF4-gsp-rev*	ACACATGCAGGCGAAGGTCAGAA	5'RACE	
IRF7-gsp-fwd	CGGAATATGTCGTCAACATGTGCT	3'RACE	N/A
IRF7-gsp-rev	CGTGGCCTCGTTGCCGTAGTG	5'RACE	
IRF8-gsp-fwd	CATGACCTCGGCAACGCCAAGA	3'RACE	N/A
IRF8-gsp-rev	CTGCATGGTGTCGGAGCTGTAG	5'RACE	
IRF10-gsp-fwd	CCGCACACCGAGAAGCCCAATA	3'RACE	N/A
IRF10-gsp-rev	GCACGCAGCCCTGCAGGATGA	5'RACE	
IRF4a-gsp-fwd	TCCATCCTACCCTGCCCTTCAC	3'RACE	N/A
IRF4a-gsp-rev	AGGAAGGCCTGCTCCGGGTAG	5'RACE	
IRF4b-gsp-fwd	GGCTTTCGTCATGAGAAGACACA	3'RACE	N/A
IRF4b-gsp-rev	GTATGTGTGCGTACGTGTGAGTG	5'RACE	
IRF10b-gsp-fwd	CGAGTCTGACCAGAGAGCAGGT	3'RACE	N/A
IRF10b-gsp-rev	CGTCTGATCAGACTCTGAGGAAG	5'RACE	
IRF4b-orf-fwd	TGACGGACAGATGAACCTCGAA	ORF-PCR	1441 bp
IRF4b-orf-rev	AGCTCAACCAATCGGGATTTCA	ORF-PCR	
IRF4a-orf-fwd	ACTTTGCCCAATCTCGTGGTGT	ORF-PCR	628 bp
IRF4a-orf-rev	GTGTGTGAACGCCTTGGAAAGA	ORF-PCR	
IRF7-orf-fwd	GGGACGACACAACGAGGTACAC	ORF-PCR	1577 bp
IRF7-orf-rev	AAAACCACGTCCCCACTACCAA	ORF-PCR	
IRF8-orf-rev	GAGCTTAAAGCCCGGAGCTCAT	ORF-PCR	1287 bp
IRF8-orf-fwd	AAGATGTCGAACACGGGAGGAC	ORF-PCR	
IRF10a-orf-fwd**	CATGAGGCGGCCTATTTGAAAG	ORF-PCR	1423 bp
IRF10a-orf-rev**	CACAGAACTGTCAACTGCCAAG	ORF-PCR	
IRF10b-orf-fwd**	TGCGCTGATGTTATGGACCTTG	ORF-PCR	651 bp
IRF10b-orf-rev**	GAGACTGTGGGAGACTGGCGTA	ORF-PCR	

*RACE products from this primer set were not used in final sequence assemblies, based on evidence of two paralogues; RACE was repeated using "*IRF4a*" and "*IRF4b*" primer sets.

***Irf10a* and *Irf10b* were renamed *Irf10-v1* (splice variant 1) and *Irf10-v2* (splice variant 2) respectively later in the study.

ORF = open reading frame; gsp = gene-specific primer.

manufacturer's instructions. Colonies containing inserts were obtained by blue/white selection on LB agar/carbenicillin (50 μg/mL) plates containing 40 μL of 40 mg/mL X-gal (Sigma, Oakville, ON), and then grown overnight at 37°C in liquid LB media containing 50 μL/mL carbenicillin. The presence of inserts was confirmed by digestion with *Eco*RI (Invitrogen) followed by electrophoresis on a 1% agarose gel, and DNA was then isolated using the QIAprep Spin Miniprep Kit (QIAGEN), following the manufacturer's instructions. For each RACE product, DNA from four colonies was sequenced in both directions using M13F and M13R primers. Sequencing was carried out by staff at the GaP (Genomics and Proteomics) facility, CREAIT network, Memorial University. Briefly, insert DNA was amplified and purified using the BigDye Terminator v3.1 Cycle Sequencing Kit and BigDye XTerminator Purification Kit (Applied Biosystems), following the manufacturer's instructions. Sequencing reactions were processed by capillary electrophoresis using the Applied Biosystems 3730 DNA Analyzer. Sequence data was compiled and analyzed using Lasergene SeqMan Pro software V. 7.1.0 (DNASTAR, Inc., Madison, WI). Amino acid sequences for each paralogue were predicted based on cDNA sequence using the ExPASy Translate tool (see Web References).

2.1.3 Paralogue and splice variant discovery

Since assembly of *Irf4* RACE sequences indicated three different contiguous sequences (contigs), further analysis of all *Irf4*-like ESTs was carried out. Based on BLAST analysis, one set of ESTs (GenBank accession numbers ES784419 and ES785894) was found to be more similar to *Irf10*, and was named *Irf10b* (with *Irf10*

above renamed as *Irf10a*). The remaining *Irf4*-like ESTs were predicted, based on nucleotide sequence comparison, to represent two paralogues, which were named *Irf4a* and *Irf4b* (Appendix 1). New GSPs were designed based on the aligned ESTs, in regions of relatively low similarity between the two paralogues. New primers were also designed to isolate *Irf10b*, in a region with relatively low similarity to *Irf10a*, and RACE, TA-cloning, and sequencing were carried out as above. Although the sequences named *Irf10a* and *Irf10b* were initially thought to be paralogues, they were later determined to be splice variants and re-named (see below).

As the 5' and 3' RACE products for each IRF paralogue had very little overlap, PCR amplification, cloning and sequencing of the open reading frames (ORFs) of all 6 paralogues were carried out, with paralogue-specific PCR primers placed 20 to 50 bp before the start codon and after the stop codon. PCR was carried out using cDNA corresponding to 25 ng or 50 ng input RNA in 50 µL reactions containing primers at a final concentration of 2.5 µM. Cycling conditions were a 3 min denaturation step at 94°C followed by 30 cycles of [30s at 94°; 30s at 60°C; 2 min at 72°C] and 10 min at 72°C. All cloning and sequencing steps were carried out as above, except that insert DNA from only one colony was sequenced 6x for each gene. Sequences were assembled using Lasergene SeqMan Pro software (DNASTAR, Inc.), and consensus sequences were used to search the NCBI non redundant (nr) protein database (BLASTx search), to confirm similarity to putative orthologous IRF sequences in other species.

2.1.4 *Irf10* genomic DNA sequencing

Based on sequence assembly and mapping to the predicted cod *Irf10* genomic region (available online from the Ensembl database), the *Irf10a* and *Irf10b* sequences were predicted to be a short and long splice variant of the same gene. To confirm this, the complete *Irf10* genomic region was cloned and sequenced. Genomic DNA was extracted from one spleen and one head kidney sample [fish injected with phosphate buffered saline (PBS) as part of the immune stimulation experiment described below (section 2.3.1)], using the DNeasy blood and tissue kit (QIAGEN) according to the manufacturer's instructions. Primers were designed in the 5' untranslated region (UTR) ("*IRF10*-genomic-fwd1") and 3' UTR ("*IRF10*-genomic-rev1") of *IRF10a* to cover most of the predicted genomic region, and PCR was carried out using the Advantage 2 Polymerase kit (Clontech) using approximately 100 ng genomic DNA per reaction, following the manufacturer's instructions. The PCR program consisted of an initial denaturation at 94°C for 1 min followed by 35 cycles of (30 s at 94°C; 4 min at 68°C) and a final extension of 4 min at 68°C. The product was electrophoretically separated on a 1% agarose/TAE gel, to confirm the presence of a product approximately 4 kb in size. Additional primers were designed to amplify and sequence the complete region in 5 parts of 800 to 1000 bp each (Table 4). PCR was carried out for each part as above, with an extension time of 1 min instead of 4 min. The PCR products were purified using QIAGEN MinElute PCR purification kit following the manufacturer's protocol, and sequenced using the same primers (carried out by staff at GaP facility, Memorial University, as above). Products were also electrophoretically separated on a 1%

Table 4: Primers used in cod *Irf10* genomic region sequencing

Primer name	**Sequence 5'-3'**
IRF10-genomic-fwd1	ACCTGAACCAGCTGGACATC
IRF10-genomic-rev1	TCGCTGGCATGTAGAAAGTC
IRF10_p1_fwd	TGCGCTGATGTTATGGACCTTG
IRF10_p1_rev	CGAGATGTCCAGCTGGTTCAG
IRF10_p2_fwd	ACAAGGTGGGCAGCGACAAGGA
IRF10_p2_rev1	TTTGTGGTTCGGCCTCTGTGTAT
IRF10_p3_fwd	GCATCGAGATCCATTTCCTCTAC
IRF10_p3_rev	GTGTCACCTGATATGGCCAGAGT
IRF10_p4_fwd	GTGCTGCTTCAGGTGCTTTGTG
IRF10_p4_rev	CTCTCCAGTTTATTGGGCTTCTC
IRF10_p5_fwd	CAGCCCAGAAGGGTGCTTCATC
IRF10_p5_rev1	CGTACTTGATTATTGTTCCAGTGC

agarose/TAE gel alongside 1 kb Plus DNA Ladder (Invitrogen) to confirm the correct approximate size.

2.1.5 Phylogenetic analysis

Homologous IRF protein sequences from other teleost species (zebrafish, Atlantic salmon, Japanese flounder, grass carp, rock bream) were collected from the NCBI non redundant (nr) protein database using the BLASTx alignment search tool and Atlantic cod *Irf* transcripts as queries. Predicted IRF amino acid sequences were aligned with the ClustalW function of MEGA5 software (Tamura *et al*., 2011). Based on the multiple sequence alignment, a phylogenetic tree was constructed using the neighbour-joining method in MEGA5, where the bootstrap consensus tree was constructed from 5000 replicates. A second multiple sequence alignment and phylogenetic tree were constructed with all sequences trimmed to the length of the shortest orthologue [Atlantic cod IRF4a (144 AA)] to remove technical bias.

2.2 RT-PCR expression analysis in juvenile cod tissues

2.2.1 Sampling and RNA extraction

All procedures involving sampling of embryonic, larval or juvenile cod were conducted with approval of Memorial University's Institutional Animal Care Committee, following the guidelines of the Canadian Council on Animal Care (protocol no. 10-02-MR). In this experiment, two juvenile Atlantic cod were individually removed from a 10°C tank and quickly euthanized by a lethal dose of tricaine methanesulphonate (TMS; 400 mg/L; Syndel Laboratories, Qualicum Beach, BC). Tissues (blood, brain, eye, gill, head kidney, heart, hindgut, liver, midgut, posterior kidney, pyloric caecum, skeletal muscle, skin, spleen, stomach) were collected by team dissection, placed in 1.5 mL microcentrifuge tubes and immediately flash frozen in liquid nitrogen before storage at -80°C. Separate instruments were used to collect each tissue, and all instruments were cleaned with RNAse Away (Sigma) between dissections.

To extract total RNA, each frozen sample was transferred to a 2 mL tube containing a 5 mm stainless steel bead and 400 μL TRIzol (Invitrogen) and homogenized by high speed agitation (TissueLyser II, QIAGEN). Further homogenization using QIAshredder (QIAGEN) columns and TRIzol extraction of RNA were performed following manufacturers' methods. RNA was treated with DNaseI (RNase Free DNase Set, QIAGEN) and column-purified using the RNeasy MinElute Cleanup Kit (Invitrogen) following the manufacturer's instructions. RNA quality (A260/280 and A260/230) and concentration were assessed by Nanodrop (ThermoFisher, Mississauga, Ont.) spectrophotometry for both crude and purified RNA, and RNA integrity was assessed by

agarose gel electrophoresis. Samples with A260/280 or A260/230 ratios of less than 1.8 were re-cleaned or omitted. One μg of each clean RNA sample was used for cDNA synthesis in 20 μL reactions containing M-MLV (Moloney Murine Leukemia Virus) Reverse Transcriptase (200 U, Invitrogen), random primers (250 ng, Invitrogen) with the manufacturer's first strand buffer and DTT (10 mM final concentration), carried out at 37°C for 50 min. cDNA was diluted 10x to 200 μL with nuclease free water (Life Technologies) and stored at -20°C.

2.2.2 RT-PCR

PCR was carried out using TopTaq DNA polymerase (QIAGEN) in 25 μL reactions containing 2 μL cDNA (corresponding to 10 ng input RNA). The same paralogue-specific primers designed for QPCR (see below) were used for RT-PCR, at a final concentration of 2.5 μM. Table 5 lists primer sequences and amplicon sizes. For each primer set, a no-template control containing all reaction components except cDNA was also run. Cycling conditions were a 3 min denaturation step at 94°C followed by 30 cycles of [30 s at 94°C; 30 s at 60°C; 1 min at 72°C] and 10 min at 72°C. PCR products were electrophoretically separated on 1.7% agarose/TAE gels (stained with ethidium bromide) alongside 1 Kb Plus DNA Ladder (Invitrogen) for 75 min at 95 V, after testing several combinations of gel percentage, running time and voltage to produce optimal resolution. *EF1-α* (elongation factor 1 α) was used as a control, as it showed similar transcript expression in all tissues studied.

2.3 QPCR expression analysis: response to immune stimulation and increased temperature in spleen

2.3.1 Experimental setup and sampling

Atlantic cod spleen samples used in this experiment were collected as part of a previous study, as described in Hori *et al.,* (2012; 2013). Briefly, Atlantic cod from 10 different families belonging to the Atlantic Cod Genomics and Broodstock Development Project (CGP) year class 3 (~60 g) were kept in 500 L tanks, four of which were held at 10°C and four of which were gradually increased over 1 month to 16°C. After

Table 5: Paralogue-specific primers used in RT-PCR and QPCR experiments

Primer name	Sequence 5'-3'	Amplicon size[1]	% Efficiency[2]
cod-ef1a-fwd	CCCTCCAGGACGTCTACAAG	150 bp	89.91
cod-ef1a-rev	GAGACTCGTGGTGCATCTCA		
arp-1-fwd	TCTGAAGCTAAGGCCCTCAA	141 bp	N/A (only used in RT-PCR)
arp-1-rev	ATCGTCGTGGAGGATCAGAG		
IRF4a-qpcr-fwd	TGTACCGTATCATCCCAGAGG	111 bp	100.58
IRF4a-qpcr-rev	AGTGGGGTATCTGGCTGTGA		
IRF4b-qpcr-fwd	TGGACATCACCGAACCCTAC	106 bp	92.25
IRF4b-qpcr-rev	CATGACGAAAGCCATCTGAA		
IRF10a-qpcr-fwd	CCGAGAAGCCCAATAAACTG	143 bp	97.74
IRF10a-qpcr-rev	ATACTCCTCGCCAAAGCAGA		
IRF10b-qpcr-fwd	GGTCCAACGCAGTAACGATT	134 bp	98.62
IRF10b-qpcr-rev	ACTGTGGGAGACTGGCGTAT		
IRF7-qpcr-fwd	CATGTGCTTTGGGGAGAAGT	152 bp	93.51
IRF7-qpcr-rev	TCTGTAGGCTGACGTTGGTG		
IRF8-qpcr-fwd	TCGGGGAGGAACTACATGAC	158 bp	91.83
IRF8-qpcr-rev	GGCCATCTCGTCTGACATCT		

[1]Forward and reverse primers were placed in adjacent predicted exons so that the amplicon would span an intron, allowing for detection of genomic DNA contamination.

[2] Percent amplification efficiency (as in Pfaffl, 2001) calculated as the average of two standard curves (see section 2.3.2 for detailed primer quality testing methods).

acclimation for 1 week, fish were intraperitoneally (IP) injected with one of the following: poly(I:C) (Sigma Co, St. Louis, MO) in sterile phosphate-buffered saline (PBS); formalin-killed, typical *A. salmonicida* (ASAL) in PBS; or PBS alone (see Hori *et al.*, 2012; 2013 for further details). As stated in Hori *et al.*, (2013), ASAL (Furogen dip vaccine, Novartis, PE) was pelleted by centrifugation (2000x *g* for 10 min at 4 °C) and washed with ice-cold, 0.2 μm filtered PBS three times; following the third wash, the pelleted cell debris was resuspended in ice-cold PBS to an optical density of 1.0 at 600 nm wavelength (OD600). Fish were injected with 4 μL of ASAL solution per gram of wet mass solution. Poly(I:C) injections contained 2 μg of poly(I:C) g^{-1} wet mass, 0.5 μg μL^{-1} in ice-cold 0.2 μm-filtered PBS (Hori *et al*., 2012). Sampling was carried out at 6 and 24 hours post-injection (HPI), using aseptic techniques as described above, and samples were stored at -80°C. Figure 3 (modified from Hori *et al.*, 2012) shows the experimental design used. For the current research, previously extracted total RNA was treated with DNaseI (RNase Free DNase Set, QIAGEN) and column-purified using the RNeasy MinElute Cleanup Kit (Invitrogen) following the manufacturer's instructions. RNA quality was determined by agarose gel electrophoresis and Nanodrop spectrophotometry, and cDNA was prepared using M-MLV reverse transcriptase as above.

2.3.2 Primer quality testing

Paralogue-specific primers (Table 4) were designed using Primer3 software (see Web References), with forward and reverse primers placed in adjacent predicted exons. The amplicon produced from each primer set would therefore include the position of an

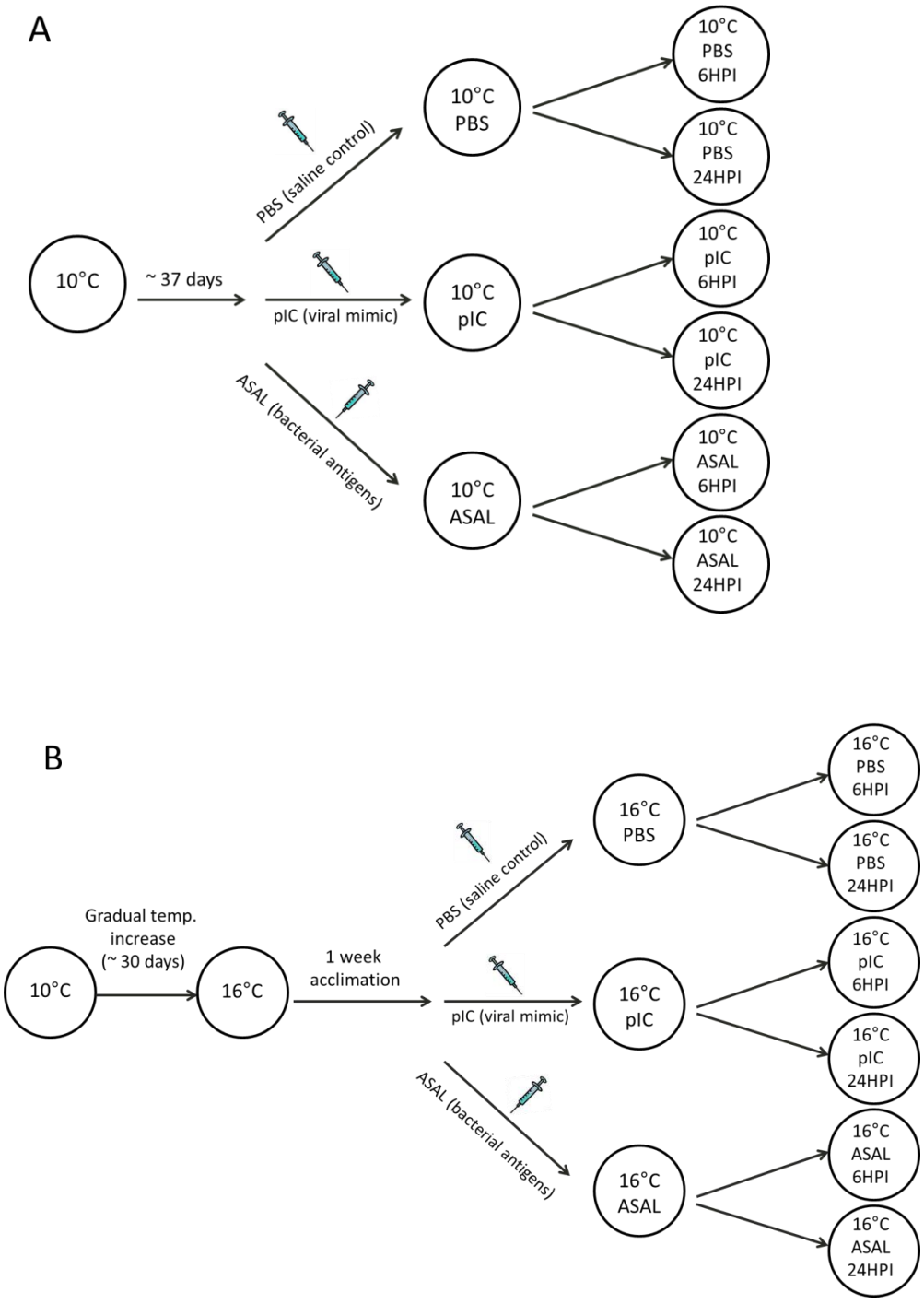


Figure 3: Overview of immune response and temperature increase experimental setup. (Modified from Hori *et al.*, 2012.) Polyriboinosinic polyribocytidylic acid, elsewhere abbreviated as poly(I:C), is abbreviated as pIC for space in this figure.

intron, allowing for the detection of genomic DNA contamination. All primer pairs were quality tested using pooled cDNA from both the 10°C, 24 h post-injection poly(I:C) and PBS sampling groups. Where possible, a 5-point, 5-fold dilution standard curve (starting with cDNA corresponding to 10 ng input RNA) was used to calculate amplification efficiency as described in Pfaffl (2001) in both poly(I:C) and PBS pools, with final amplification efficiency reported as the average of the two. However, due to low expression of several transcripts, 4-fold (*Irf4a*, *Irf8*) or 3- fold (*Irf4b*, *Irf10-v2*) 5-point dilution series had to be used for those standard curves. Triplicate reactions were carried out for all standard curves, controls and experimental samples. Melt curve analysis was carried out to ensure that a single product was amplified and that no primer-dimers were present. *EF1α* was confirmed as a suitable normalizer by testing in approximately one third of the experimental samples, including all time points and treatments. The range of threshold cycle (C_T) values for *EF1α* was 1.7 cycles, indicating a similar level of expression in the included samples.

2.3.3 QPCR analysis

All QPCR analyses were performed using SYBR Green chemistry and the ViiA7 Real Time PCR System (Applied Biosystems, Foster City, CA, USA). PCR amplification was carried out in 13 μL reactions containing 6.5 μL Power SYBR Green master mix (Applied Biosystems), 0.52 μL each of forward and reverse primers (50 nM final concentration), 3.46 μL nuclease-free water and 2 μL cDNA (corresponding to 10 ng input RNA). All samples were run as triplicate technical replicates, and no-template controls were included for each primer set in each plate. A linker sample of cDNA made

from pooled PBS-injected (10°C, 24 HPI) samples was run on each plate; all linker C_T values were within 1 cycle. To confirm the absence of any genomic DNA, a no reverse transcription (no-RT) control was also included in which a cDNA synthesis reaction using the linker RNA pool was carried out with all components except reverse transcriptase. The reaction product was run in triplicate (2 μL as with cDNA samples), and no amplification was observed in the no-RT control.

Gene of interest expression was normalized to *EF1α* expression, and relative quantities (RQ) were calculated with the Applied Biosystems ViiA7 Software Relative Quantification Study Application using the Pfaffl method (Pfaffl, 2001) and automatic thresholds, incorporating calculated amplification efficiencies. The lowest expressing sample for each gene of interest was set as the calibrator (RQ set as 1.0) for analysis of that gene. RQ values were analyzed statistically and plotted using Prism v5.0 (GraphPad Software Inc., La Jolla, CA, USA). A two-way ANOVA with treatment and temperature as factors was carried out for each time point. If the effect of one factor was statistically significant ($p<0.05$), t-tests were performed to compare groups, as described in Hori *et al.*, (2012).

2.4 RT-PCR expression analysis: developmental expression

2.4.1 Experimental setup and sampling

Adult (broodstock) Atlantic cod involved in this study were handled by the staff of the Dr. Joe Brown Aquatic Research Building (JBARB) at the Ocean Sciences Centre of Memorial University. Broodstock were wild fish caught in Smith Sound,

Newfoundland. After communal spawning, fertilized eggs were collected in 3 batches and ozonated at 1.5-2 ppm for 1.5 min and placed in three 250 L incubators with air stones. Temperature was recorded daily and maintained at 5-7 °C for the duration of sampling, and non-buoyant dead embryos and/or shells from hatched larvae were removed daily by draining from the bottom of each incubator before sampling.

Sampling was carried out from 0 to 17 days post-fertilization (dpf). Each day, the air stone was removed to allow embryos to float to the top of the incubator, and a mesh screen was used to collect a small number of embryos. For each incubator, ~250 µL of embryos were placed in a 1.5 mL RNase-free microcentrifuge tube containing 1 mL RNA Later (Life Technologies) and stored at 4°C overnight. Samples were divided into groups of 30 embryos the following day using a sterile spatula (after removing liquid) and then stored at -80°C. Each day, additional samples of ~250 µL embryos were collected from each incubator, flash frozen in liquid nitrogen and stored at -80°C for use in future work. Embryos were also observed under a light microscope to estimate developmental stage, and pictures were taken of representative samples for each day.

2.4.2 RT-PCR

RNA extraction of two complete sets of samples (0 dpf to 17 dpf, from two different incubators) was carried out by homogenization in ~600 µL TRIzol (Invitrogen) using a motorized Kontes RNase-Free Pellet Pestle Grinder (Kimble Chase, Vineland, NJ) and sterile plastic pestles. Samples were immediately transferred to QIAshredders and RNA extraction, cleaning, quality checking, and cDNA synthesis were carried out as described above for the tissue panel RT-PCR (section 2.2). For PCR, an acidic ribosomal

protein (*arp*) transcript was used as a control / housekeeping gene instead of *ef1α* based on its evaluation in a previous study (Lanes *et al.,* 2012) and on preliminary QPCR data (not shown) which suggested it was more stable than *ef1α* in the included embryonic/larval samples. QPCR was not completed due to very low constitutive expression of IRF transcripts in the early life stage samples. Instead, RT-PCR only was carried out, using TopTaq DNA polymerase kit (QIAGEN) as in the tissue panel study above (using the same primers, cDNA quantity, etc.), and 12.5 µL of each reaction was electrophoretically separated on a 1.7% agarose/TAE gel alongside 1 Kb Plus DNA Ladder (Invitrogen).

3. Results

3.1 Characterization of *Irf4a, Irf4b, Irf7, Irf8, Irf10-v1* and *Irf10-v2* cDNA sequences

Primers were designed based on RACE sequence assemblies to amplify the ORF of each paralogue (from 20 to 100 bp before the start codon to 20 to 100 bp after the stop codon) to confirm assemblies were correct and to ensure all assemblies contained 6x coverage of every base. Agarose gel electrophoresis (Figure 4) of the PCR products shows that bands of the approximate predicted sizes (listed in Table 3) were obtained for each of the six IRF transcripts.

Assembly of *Irf4a* sequencing reads (RACE sequences as well as additional ORF sequencing reads to confirm overlapping region; Appendix 2) produced a 796 bp cDNA sequence (excluding poly-A tail) (Figure 5). The sequence consists of a 435 bp (144 AA)

Figure 4: Agarose gel image of PCR amplified IRF open reading frames. Composite of two 1% agarose gels, each using 1 kb plus ladder (Invitrogen) to determine approximate band size. Two reactions were run for each gene, starting with 5 µL and 10 µL of cDNA (corresponding to 25 ng and 50 ng input RNA, respectively) in 50 µL reactions (45 µL of each reaction was run on the gel). Primer sequences and expected band sizes are indicated in Table 3. Note that amplicons are longer than the ORF for each gene (spanning from before the start codon to after the stop codon). Bands matching predicted approximate sizes for each amplicon are indicated in red, and were excised for TA-cloning and sequencing. The gel section showing 1 kb plus ladder is replicated for easier estimation of band sizes.

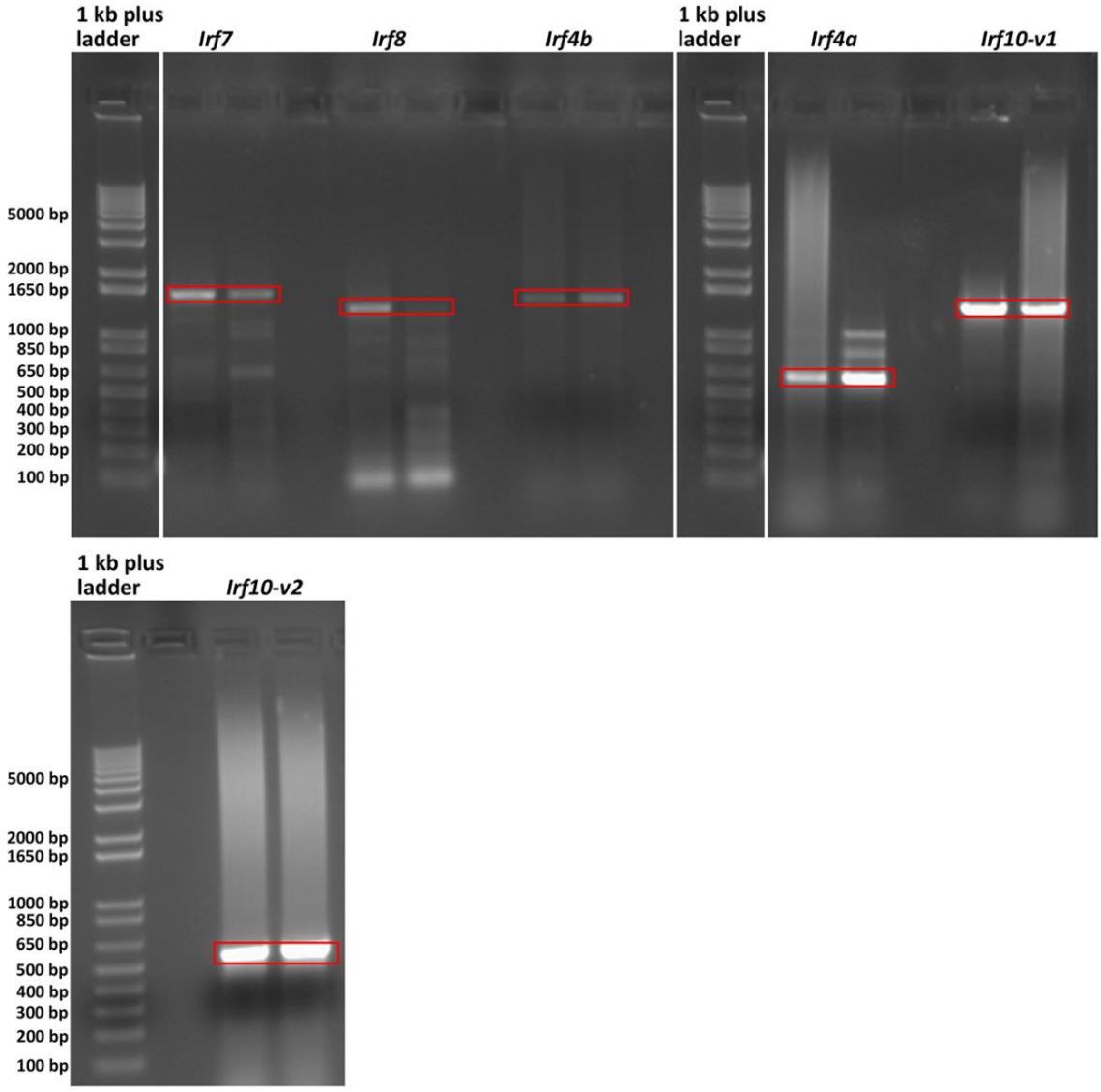

1 kb plus ladder
Irf7
Irf8
Irf4b
1 kb plus ladder
Irf4a
Irf10-v1
5000 bp
2000 bp
1650 bp
1000 bp
850 bp
650 bp
500 bp
400 bp
300 bp
200 bp
100 bp
1 kb plus ladder
Irf10-v2
5000 bp
2000 bp
1650 bp
1000 bp
850 bp
650 bp
500 bp
400 bp
300 bp
200 bp
100 bp

Figure 5: Nucleotide sequence of Atlantic cod *Irf4a* cDNA and inferred amino acid translation. Nucleotide sequence is numbered on the left, while the predicted amino acid sequence is numbered on the right. The open reading frame is shown in upper case letters while 5' and 3' untranslated regions are in lower case letters. Nucleotide sequence of the DNA binding domain is shaded in grey. Locations of predicted introns are indicated based on Ensembl predicted transcript ENSGMOT00000005509. The stop codon is marked with an asterisk (*). Arrows indicate position of gene-specific primers used in QPCR. A possible polyadenylation signal (GTTAAA; MacDonald and Redondo, 2002) is bolded.


1 atcctggttgcgaacttgattgagaataagttgaaaatgcttgagggtctttgattattttcgaggtcaa
71 atccgttccataacattttctgttaagctgatgtaaaacttctgatactttcatcttactttgcccaatc
141 tcgtggtgtttgaacagagATGCATTTCGAGGAGGACGTCAATCTGTCAGTCAGTTGCGGCAACGGGAAG
M H F E E D V N L S V S C G N G K 17
211 CTTAGACAGTGGCTGATCGATCAGATTGACAGCAAGAGCTACCTGGGCTTGGTTTGGGAGAATGTGGAGA
L R Q W L I D Q I D S K S Y L G L V W E N V E 40
281 AATCCATTTTCAGGATACCGTGGAAGCATGCGGGCAAACAAGATTACAACAGAGATGAGGATGCTGCGCT
K S I F R I P W K H A G K Q D Y N R D E D A A L 64
predicted intron 1
351 TTTCAAGGCCTGGGCACTTTTCAAGGACAAATACAAAGAAGGTGTGGACAAACCGGACCCCCCCACATGG
F K A W A L F K D K Y K E G V D K P D P P T W 87
421 AAAACCCGTCTACGGTGTGCTCTGAACAAAAGCAACGACTTTGACGAGCTGGTGGACAGAAGCCAGCTGG
K T R L R C A L N K S N D F D E L V D R S Q L 110
predicted intron 2
491 ACATCACCGAACCCTACAAAGTCTACAGAATCATCCCCGAGGGGGTCAAAAGAGGCAAGCCCATCAATAA
D I T E P Y K V Y R I I P E G V K R G K P I N K 134
561 AGTGTCTGCAATATTCAGATGGCTTTCGTCATGAgaagacacatttattgtacagatgtgcagacttccc
V S A I F R W L S S * 144
631 tgattgcgtgcagttacacacatactcacacactcacacgtacgcacacatacccacacactgcagcgtg
701 acaaagcggggcaactctgtggtcatggttaaatctttccaaggcgttcacacactgacgtgaaaacacc
771 catagagacacacgcgaacactttaca(n)

ORF, a 159 bp 5' untranslated region (UTR), and a 202 bp 3'-UTR. The most common polyadenylation signal (AAUAAA, located 10 to 30 nt upstream of the polyadenylation site; see Colgan and Manley, 1997 for review) is not present in the 3'-UTR; instead the sequence GUUAAA may act as the polyadenylation signal for this transcript. This hexamer has previously been identified as a potential polyadenylation signal in mouse germ cells (MacDonald and Redondo, 2002). Assembly of sequencing reads for the longer *Irf4* paralogue (*Irf4b*) produced a 1,685 bp cDNA sequence (excluding poly-A tail) (Figure 6; Appendix 3). The cDNA consists of a 1,347 bp (448 AA) ORF, a 171 bp 5'-UTR, and a 167 bp 3'-UTR). A possible polyadenylation signal (ACUAAA) was identified 25 nt upstream of the poly-A tail.

Irf7 sequencing reads were assembled to produce a 2,037 bp cDNA sequence consisting of a 1,326 bp (441 AA) ORF, a 36 bp 5'-UTR and a 675 bp 3'-UTR containing an AUUAAA polyadenylation signal (Figure 7; Appendix 4). Assembly of *Irf8* sequencing reads produced a 1,827 bp cDNA sequence consisting of a 1,266 bp (421 AA) ORF, a 99 bp 5'-UTR, and a 461 bp 3'-UTR containing the polyadenylation signal AAUAAA (Figure 8; Appendix 5).

Irf10-v1 (splice variant 1) RACE and ORF PCR sequencing reads were assembled to produce a 1,721 bp cDNA sequence consisting of a 1,191 bp (396 AA) ORF, a 106 bp 5'-UTR and a 417 bp 3'-UTR containing a possible AGUAAA polyadenylation signal (Figure 9; Appendix 6). The *Irf10-v2* cDNA is much shorter (1,171 bp), with an ORF of only 381 bp (126 AA), a 128 bp 5'-UTR, and a 663 bp 3'-UTR containing a possible AAUAAA polyadenylation signal [although it should be noted that this hexamer is

Figure 6: Nucleotide sequence of Atlantic cod *Irf4b* cDNA and inferred amino acid translation. Nucleotide sequence is numbered on the left, while the predicted amino acid sequence is numbered on the right. The open reading frame is shown in upper case letters while 5' and 3' untranslated regions are in lower case letters. Nucleotide sequence of the DNA binding domain is shaded in grey. Locations of predicted introns are indicated based on Ensembl predicted transcript ENSGMOT00000018695. The stop codon is indicated by an asterisk (*). Arrows indicate position of gene-specific primers used in QPCR. A possible polyadenylation signal (ACTAAA; MacDonald and Redondo, 2002) is indicated in bold.

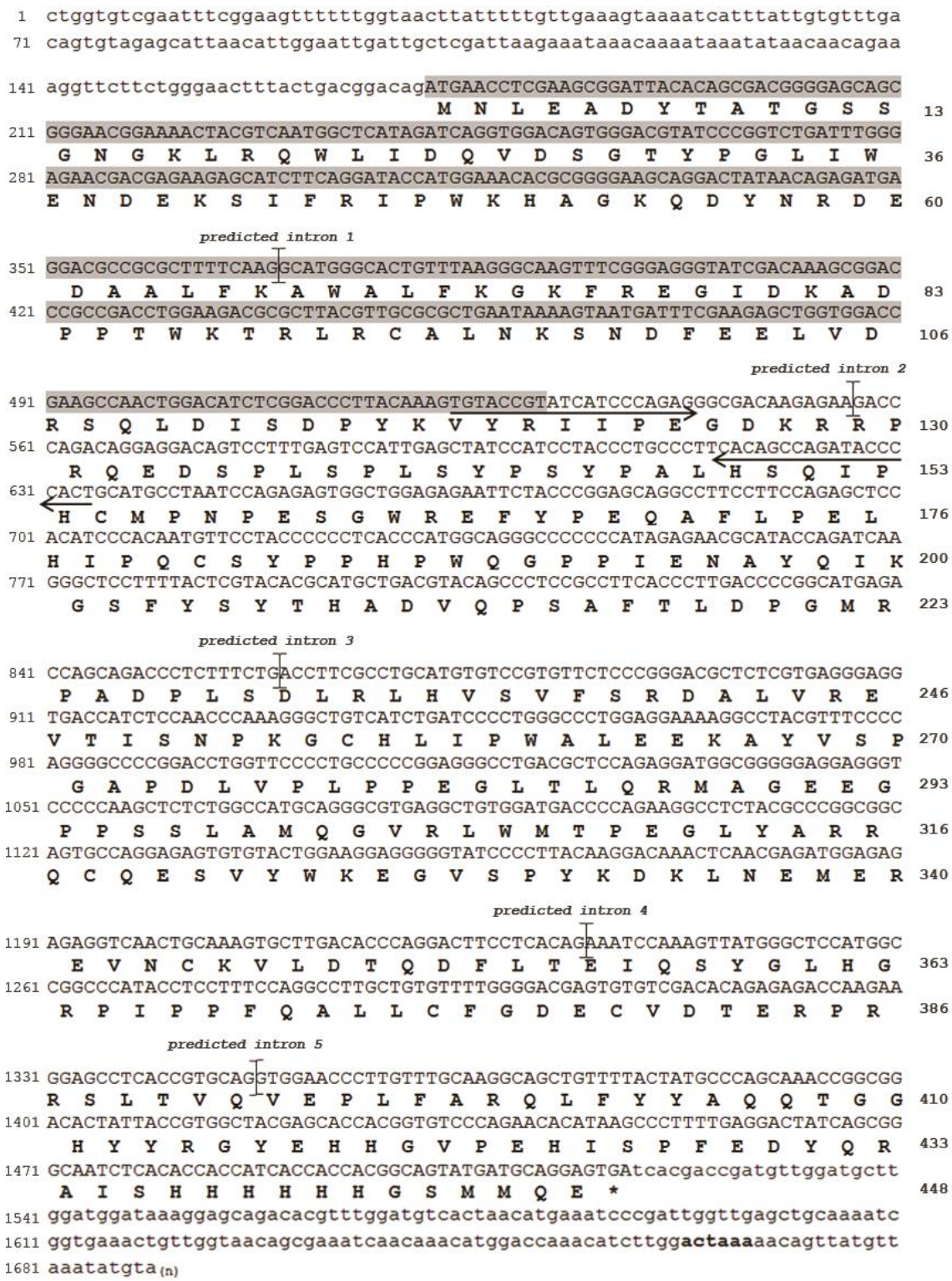

predicted intron 1
predicted intron 2
predicted intron 3
predicted intron 4
predicted intron 5

Figure 7: Nucleotide sequence of Atlantic cod *Irf7* cDNA and inferred amino acid translation. Nucleotide sequence is numbered on the left, while the predicted amino acid sequence is numbered on the right. The open reading frame is shown in upper case letters while 5' and 3' untranslated regions are in lower case letters. Nucleotide sequence of the DNA binding domain is shaded in grey. Locations of predicted introns are indicated based on Ensembl predicted transcript ENSGMOT00000010511. The stop codon is indicated by an asterisk (*). Arrows indicate position of gene-specific primers used in QPCR. The polyadenylation signal is indicated in bold.


predicted intron 1
predicted intron 2
predicted intron 3
predicted intron 4
predicted intron 5
predicted intron 6

Figure 8: Nucleotide sequence of Atlantic cod *Irf8* cDNA and inferred amino acid translation. Nucleotide sequence is numbered on the left, while the predicted amino acid sequence is numbered on the right. The open reading frame is shown in upper case letters while 5' and 3' untranslated regions are in lower case letters. Nucleotide sequence of the DNA binding domain is shaded in grey. Locations of predicted introns are indicated based on Ensembl predicted transcript ENSGMOT00000004315. The stop codon is indicated by an asterisk (*). Arrows indicate position of gene-specific primers used in QPCR. The polyadenylation signal is indicated in bold.

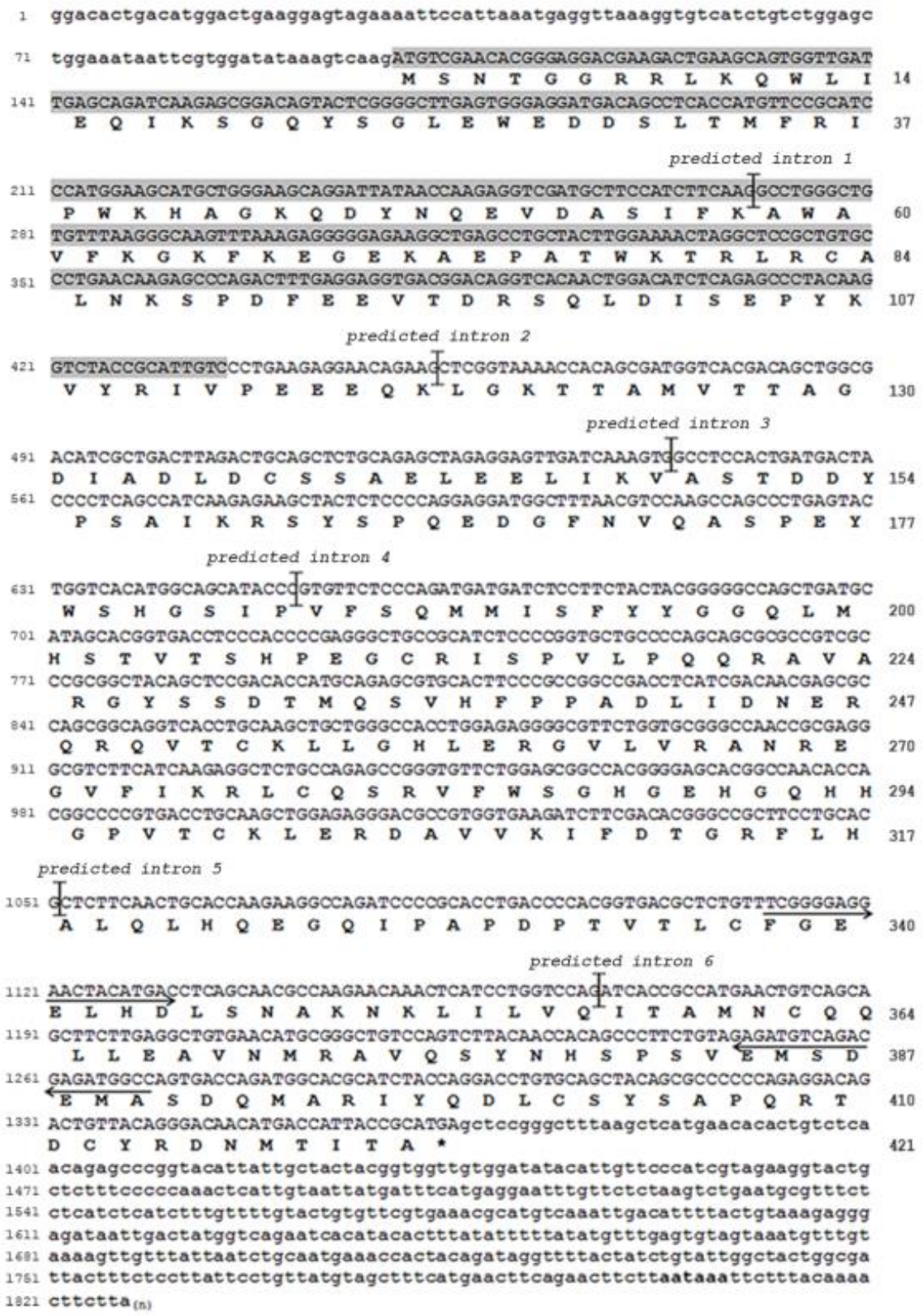
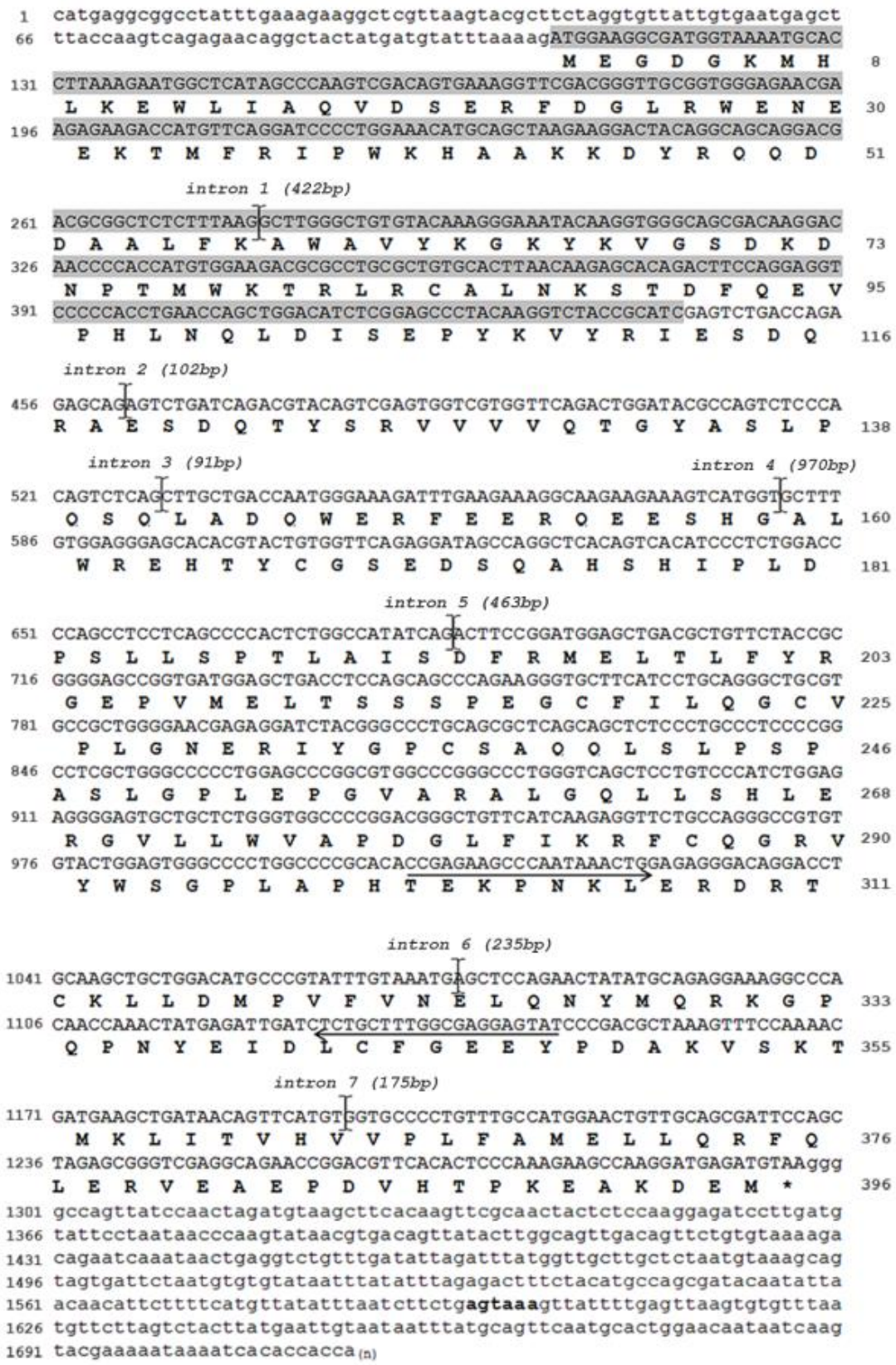

predicted intron 1
predicted intron 2
predicted intron 3
predicted intron 4
predicted intron 5
predicted intron 6

Figure 9: Nucleotide sequence of Atlantic cod *Irf10-v1* cDNA and inferred amino acid translation. Nucleotide sequence is numbered on the left, while the predicted amino acid sequence is numbered on the right. The open reading frame is shown in upper case letters while 5' and 3' untranslated regions are in lower case letters. Nucleotide sequence of the DNA binding domain is shaded in grey. The stop codon is indicated by an asterisk (*). Arrows indicate position of gene-specific primers used in QPCR. A possible polyadenylation signal (AGTAAA, MacDonald and Redondo, 2002) is indicated in bold. Position of introns is indicated based on mapping to genomic sequence.


1 catgaggcggcctatttgaaagaaggctcgttaagtacgcttctaggtgttattgtgaatgagct
66 ttaccaagtcagagaacaggctactatgatgtatttaaaagATGGAAGGCGATGGTAAAATGCAC
M E G D G K M H 8
131 CTTAAAGAATGGCTCATAGCCCAAGTCGACAGTGAAAGGTTCGACGGGTTGCGGTGGGAGAACGA
L K E W L I A Q V D S E R F D G L R W E N E 30
196 AGAGAAGACCATGTTCAGGATCCCCTGGAAACATGCAGCTAAGAAGGACTACAGGCAGCAGGACG
E K T M F R I P W K H A A K K D Y R Q Q D 51
intron 1 (422bp)
261 ACGCGGCTCTCTTTAAGGCTTGGGCTGTGTACAAAGGGAAATACAAGGTGGGCAGCGACAAGGAC
D A A L F K A W A V Y K G K Y K V G S D K D 73
326 AACCCCACCATGTGGAAGACGCGCCTGCGCTGTGCACTTAACAAGAGCACAGACTTCCAGGAGGT
N P T M W K T R L R C A L N K S T D F Q E V 95
391 CCCCCACCTGAACCAGCTGGACATCTCGGAGCCCTACAAGGTCTACCGCATCGAGTCTGACCAGA
P H L N Q L D I S E P Y K V Y R I E S D Q 116
intron 2 (102bp)
456 GAGCAGAGTCTGATCAGACGTACAGTCGAGTGGTCGTGGTTCAGACTGGATACGCCAGTCTCCCA
R A E S D Q T Y S R V V V V Q T G Y A S L P 138
intron 3 (91bp)
intron 4 (970bp)
521 CAGTCTCAGCTTGCTGACCAATGGGAAAGATTTGAAGAAAGGCAAGAAGAAAGTCATGGTGCTTT
Q S Q L A D Q W E R F E E R Q E E S H G A L 160
586 GTGGAGGGAGCACACGTACTGTGGTTCAGAGGATAGCCAGGCTCACAGTCACATCCCTCTGGACC
W R E H T Y C G S E D S Q A H S H I P L D 181
intron 5 (463bp)
651 CCAGCCTCCTCAGCCCCACTCTGGCCATATCAGACTTCCGGATGGAGCTGACGCTGTTCTACCGC
P S L L S P T L A I S D F R M E L T L F Y R 203
716 GGGGAGCCGGTGATGGAGCTGACCTCCAGCAGCCCAGAAGGGTGCTTCATCCTGCAGGGCTGCGT
G E P V M E L T S S S P E G C F I L Q G C V 225
781 GCCGCTGGGGAACGAGAGGATCTACGGGCCCTGCAGCGCTCAGCAGCTCTCCCTGCCCTCCCCGG
P L G N E R I Y G P C S A Q Q L S L P S P 246
846 CCTCGCTGGGCCCCCTGGAGCCCGGCGTGGCCCGGGCCCTGGGTCAGCTCCTGTCCCATCTGGAG
A S L G P L E P G V A R A L G Q L L S H L E 268
911 AGGGGAGTGCTGCTCTGGGTGGCCCCGGACGGGCTGTTCATCAAGAGGTTCTGCCAGGGCCGTGT
R G V L L W V A P D G L F I K R F C Q G R V 290
976 GTACTGGAGTGGGCCCCTGGCCCCGCACACCGAGAAGCCCAATAAACTGGAGAGGGACAGGACCT
Y W S G P L A P H T E K P N K L E R D R T 311
intron 6 (235bp)
1041 GCAAGCTGCTGGACATGCCCGTATTTGTAAATGAGCTCCAGAACTATATGCAGAGGAAAGGCCCA
C K L L D M P V F V N E L Q N Y M Q R K G P 333
1106 CAACCAAACTATGAGATTGATCTCTGCTTTGGCGAGGAGTATCCCGACGCTAAAGTTTCCAAAAC
Q P N Y E I D L C F G E E Y P D A K V S K T 355
intron 7 (175bp)
1171 GATGAAGCTGATAACAGTTCATGTGGTGCCCCTGTTTGCCATGGAACTGTTGCAGCGATTCCAGC
M K L I T V H V V P L F A M E L L Q R F Q 376
1236 TAGAGCGGGTCGAGGCAGAACCGGACGTTCACACTCCCAAAGAAGCCAAGGATGAGATGTAAggg
L E R V E A E P D V H T P K E A K D E M * 396
1301 gccagttatccaactagatgtaagcttcacaagttcgcaactactctccaaggagatccttgatg
1366 tattcctaataacccaagtataacgtgacagttatacttggcagttgacagttctgtgtaaaaga
1431 cagaatcaaataactgaggtctgtttgatattagatttatggttgcttgctctaatgtaaagcag
1496 tagtgattctaatgtgtgtataatttatatttagagactttctacatgccagcgatacaatatta
1561 acaacattcttttcatgttatatttaatcttctgagtaaagttattttgagttaagtgtgtttaa
1626 tgttcttagtctacttatgaattgtaataatttatgcagttcaatgcactggaacaataatcaag
1691 tacgaaaaataaaatcacaccacca(n)

further upstream (110 bp from the poly-A tail) than a usual polyadenylation signal] (Figure 10; Appendix 7). Alignment of both *Irf10* sequences and comparison with the predicted cod *Irf10* genomic region obtained from the Ensembl database indicated they were likely alternate splice variants rather than different paralogues.

Sequencing and assembly of the cod *Irf10* genomic region produced a 3,828 bp consensus sequence which was aligned with *Irf10-v1* and *Irf10-v2* transcripts to determine intron positions. The positions of 7 introns, ranging from 91 bp to 970 bp in length, are indicated for *Irf10-v1* (Figure 9), dividing the transcript into 8 exons. *Irf10-v2*, while identical to *Irf10-v1* up to the end of exon 2, appears to retain intron 2 producing a premature stop codon (Figure 10). The 3'-UTR of the *Irf10-v2* transcript appears to contain exon 3, intron 3, exon 4 and part of intron 4. The putative intron/exon structure of the cod *Irf10* gene (and the difference between splice variants) based on these sequences is shown in Figure 11.

For *Irf4a*, *Irf4b*, *Irf7* and *Irf8* the location and size on introns were estimated based on comparison to predicted sequences obtained from the Ensembl database (Figure 12). However, because these genomic sequences are not complete, some intron placements and sizes are still uncertain. It is therefore of interest to sequence the complete genomic region for each of these paralogues in the future. Interestingly, the structure of *Irf4a* appears to be similar to the shorter *Irf10* splice variant (*Irf10-v2*) and both are of similar length, encoding putative proteins of 144 and 126 AA respectively. While *Irf4a* and *Irf4b* are paralogues and not splice variants (having approximately 74% identiy at the amino acid level; see Appendix 8), it is possible that a longer splice variant

Figure 10: Nucleotide sequence of Atlantic cod *Irf10-v2* cDNA and inferred amino acid translation. Nucleotide sequence is numbered on the left, while the predicted amino acid sequence is numbered on the right. The open reading frame is shown in upper case letters while 5' and 3' untranslated regions are in lower case letters. Nucleotide sequence of the DNA binding domain is shaded in grey. The stop codon is indicated by an asterisk (*). Arrows indicate position of gene-specific primers used in QPCR. A possible polyadenylation signal is indicated in bold. Position of introns is indicated based on mapping to *Irf10-v1* transcript and genomic sequence.

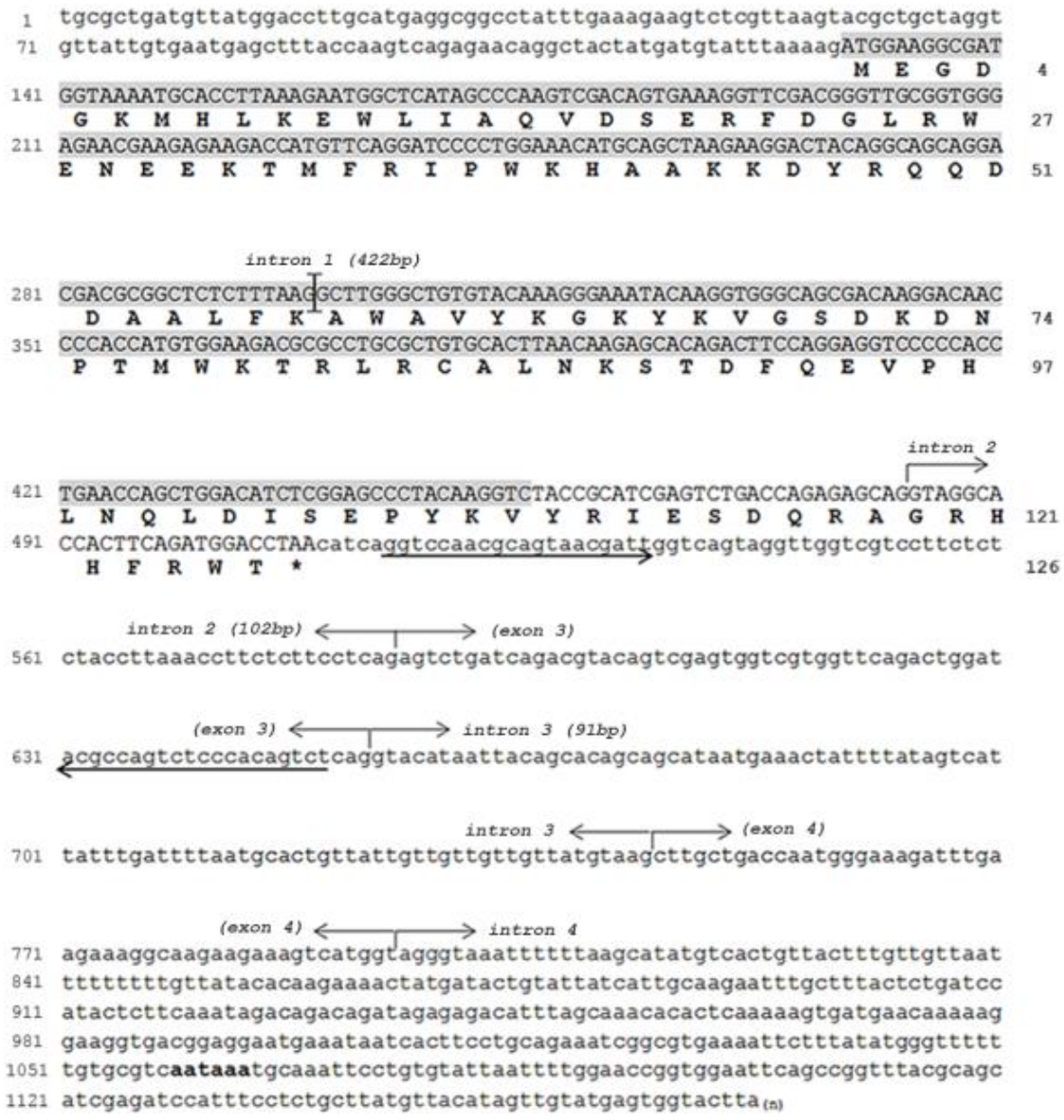

1 tgcgctgatgttatggaccttgcatgaggcggcctatttgaaagaagtctcgttaagtacgctgctaggt
71 gttattgtgaatgagctttaccaagtcagagaacaggctactatgatgtatttaaaagATGGAAGGCGAT
M E G D 4
141 GGTAAAATGCACCTTAAAGAATGGCTCATAGCCCAAGTCGACAGTGAAAGGTTCGACGGGTTGCGGTGGG
G K M H L K E W L I A Q V D S E R F D G L R W 27
211 AGAACGAAGAGAAGACCATGTTCAGGATCCCCTGGAAACATGCAGCTAAGAAGGACTACAGGCAGCAGGA
E N E E K T M F R I P W K H A A K K D Y R Q Q D 51
intron 1 (422bp)
281 CGACGCGGCTCTCTTTAAGGCTTGGGCTGTGTACAAAGGGAAATACAAGGTGGGCAGCGACAAGGACAAC
D A A L F K A W A V Y K G K Y K V G S D K D N 74
351 CCCACCATGTGGAAGACGCGCCTGCGCTGTGCACTTAACAAGAGCACAGACTTCCAGGAGGTCCCCCACC
P T M W K T R L R C A L N K S T D F Q E V P H 97
intron 2
421 TGAACCAGCTGGACATCTCGGAGCCCTACAAGGTCTACCGCATCGAGTCTGACCAGAGAGCAGGTAGGCA
L N Q L D I S E P Y K V Y R I E S D Q R A G R H 121
491 CCACTTCAGATGGACCTAAcatcaggtccaacgcagtaacgattggtcagtaggttggtcgtccttctct
H F R W T * 126
intron 2 (102bp)
(exon 3)
561 ctaccttaaaccttctcttcctcagagtctgatcagacgtacagtcgagtggtcgtggttcagactggat
(exon 3)
intron 3 (91bp)
631 acgccagtctcccacagtctcaggtacataattacagcacagcagcataatgaaactattttatagtcat
intron 3
(exon 4)
701 tatttgattttaatgcactgttattgttgttgttgttatgtaagcttgctgaccaatgggaaagatttga
(exon 4)
intron 4
771 agaaaggcaagaagaaagtcatggtagggtaaatttttttaagcatatgtcactgttactttgttgttaat
841 ttttttttgttatacacaagaaaactatgatactgtattatcattgcaagaatttgctttactctgatcc
911 atactcttcaaatagacagacagatagagagacatttagcaaacacactcaaaaagtgatgaacaaaaag
981 gaaggtgacggaggaatgaaataatcacttcctgcagaaatcggcgtgaaaattctttatatgggttttt
1051 tgtgcgtcaataaatgcaaattcctgtgtattaattttggaaccggtggaattcagccggtttacgcagc
1121 atcgagatccatttcctctgcttatgttacatagttgtatgagtggtactta(n)

Figure 11: Schematic representation of predicted intron/exon organization of Atlantic cod *Irf10*. Exons are shown as black boxes, with length above (in bp), while introns are shown as horizontal lines with lengths below (in bp). Noncoding regions of exons are shown as grey shaded boxes. Drawings are to scale, except where long introns are depicted as bent lines. The structure of the Irf10-v2 transcript is depicted below, where the portion of intron 2 which is included in the ORF shown as a white box.

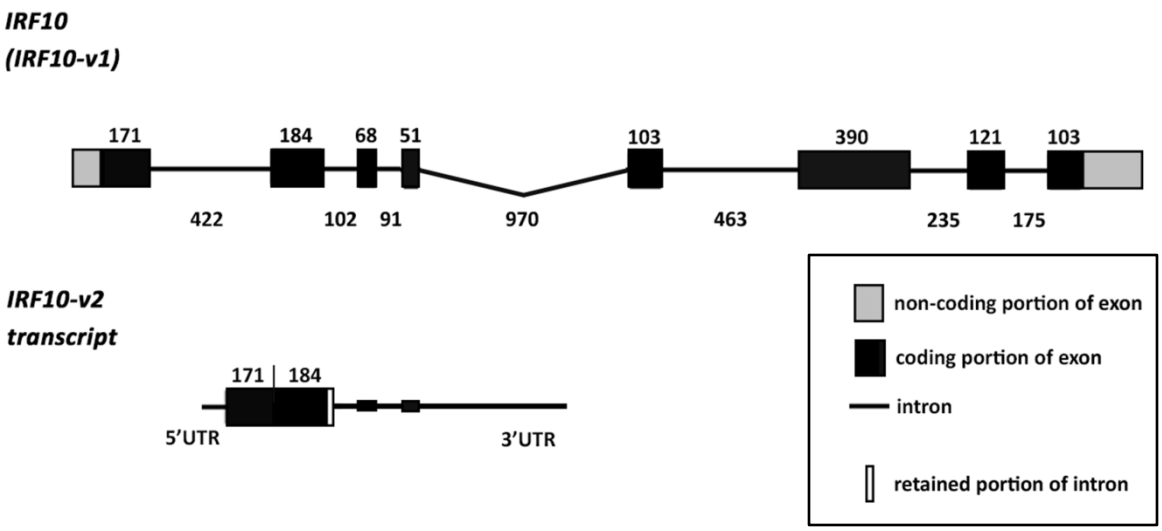
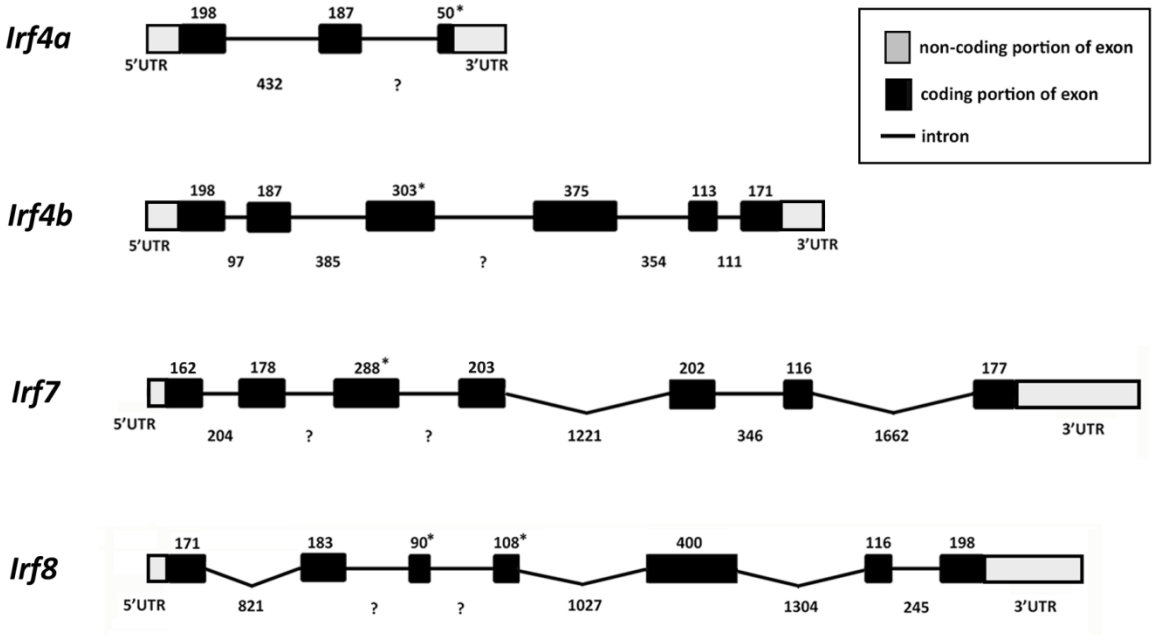

IRF10
(IRF10-v1)
171
184
68
51
103
390
121
103
422
102
91
970
463
235
175
IRF10-v2
transcript
171
184
5'UTR
3'UTR
non-coding portion of exon
coding portion of exon
intron
retained portion of intron

Figure 12: Schematic representation of predicted intron/exon organization of Atlantic cod *Irf4a, Irf4b, Irf7, Irf8*. Exons are shown as black boxes, with length above (in bp), while introns are shown as horizontal lines with lengths below (in bp). Non-coding regions of exons (5' and 3' UTRs) are shown as grey boxes. Drawings are to scale, except where long introns are depicted as bent lines. Introns whose positions do not match Ensembl predicted genome sequences and whose length could therefore not be estimated are marked by "?". Exons that differ from Ensembl predicted sequences are marked by an asterisk. Note that Exons 1 and 2 of *Irf4a* and *Irf4b*, while identical in length, are not identical in sequence between the two paralogues (see alignment in Figure 13).


Irf4a
198
187
50*
5'UTR
3'UTR
432
?
Irf4b
198
187
303*
375
113
171
5'UTR
3'UTR
97
385
?
354
111
Irf7
162
178
288*
203
202
116
177
5'UTR
3'UTR
204
?
?
1221
346
1662
Irf8
171
183
90*
108*
400
116
198
5'UTR
3'UTR
821
?
?
1027
1304
245
non-coding portion of exon
coding portion of exon
intron

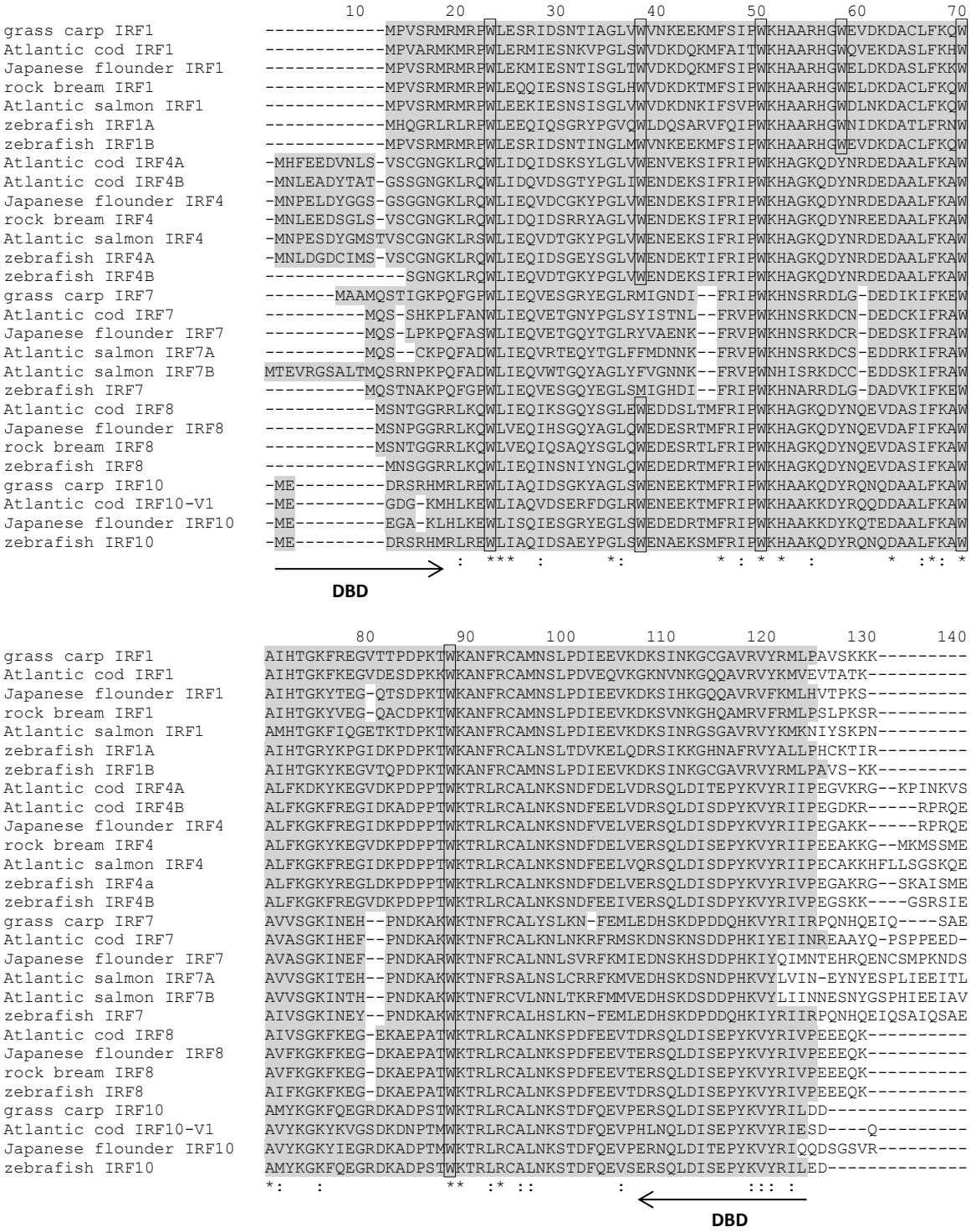
of *Irf4a* (i.e. more similar in length to *Irf4b*) is present in Atlantic cod but was not found in the current study.

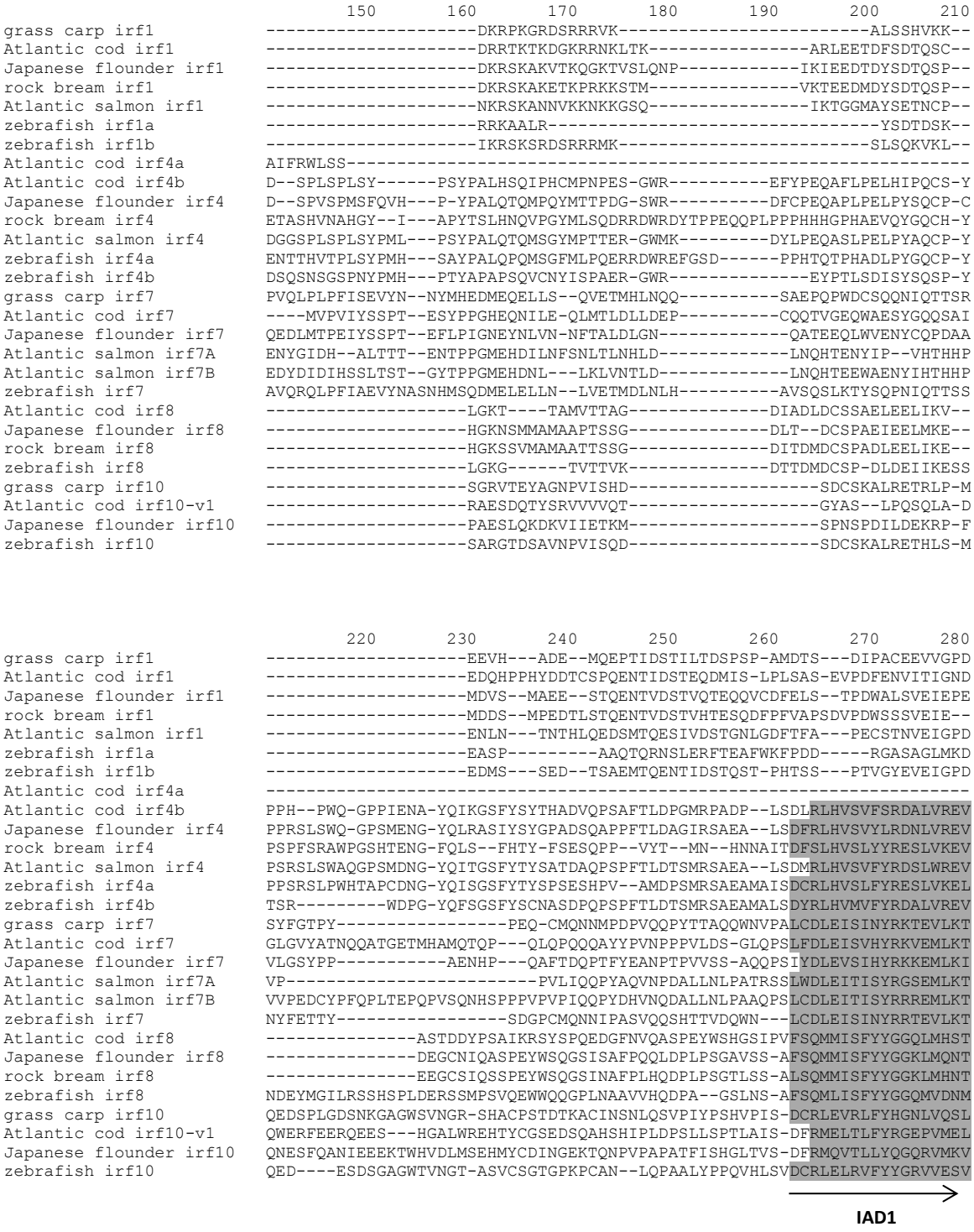
Phylogenetic analysis of IRF proteins from several teleost species (including Atlantic cod IRF amino acid sequences shown in Figures 5-10 and the previously sequenced cod IRF1) indicates that the cod IRF paralogues sequenced in this study are orthologous to IRFs from other fish species, and also fit into the sub-groups depicted in Figure 1. Multiple sequence alignment shows that the DNA binding domain (first 110-120 AA) of all sequences included are quite similar, with several conserved amino acids including the multiple Trp residues found in all IRFs (Figure 13). A high degree of similarity can also be seen in IRF4, 7, 8, and 10 sequences in the IRF association domain (IAD), which is not shared by IRF1 sequences. In a phylogenetic tree based on the multiple sequence alignment (Figure 14), all IRF4, IRF8, and IRF10 sequences group together (IRF4-G sub-group), while the IRF1 and IRF7 proteins form separate branches (representing IRF1-G and IRF3-G sub-groups, respectively). The teleost fish species used for comparison belong to several different superorders; Atlantic cod (superorder Paracanthopterygii) IRF proteins appear to be more similar in sequence overall to those of flounder (superorder Acanthopterygii) than to zebrafish or carp (superorder Ostariophysi).

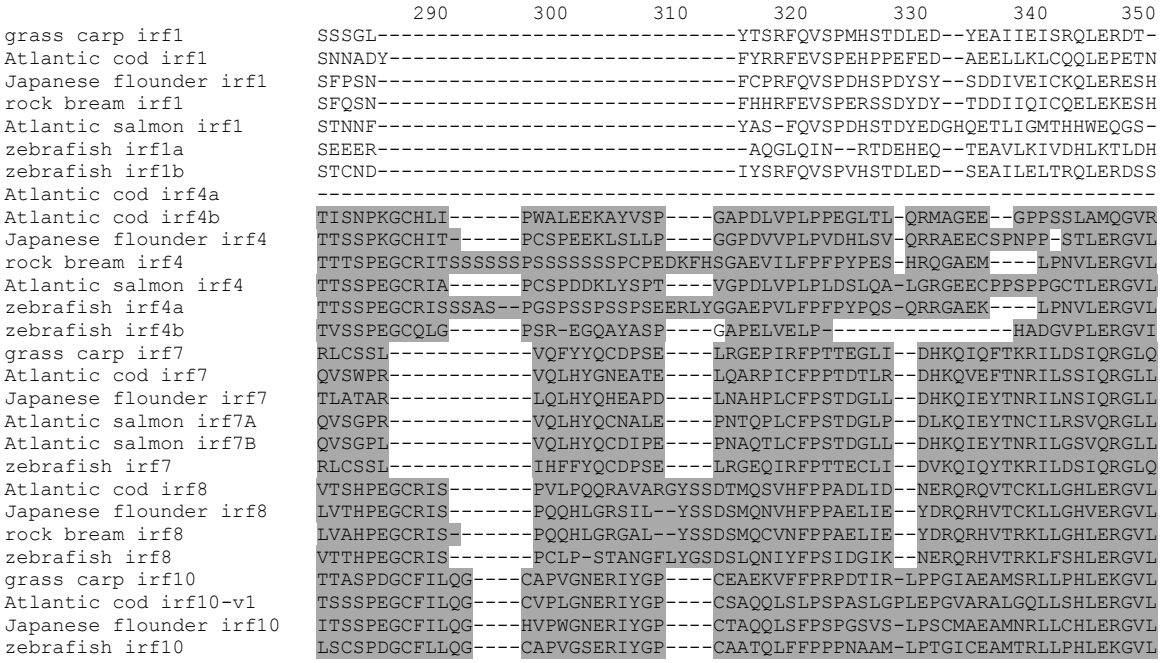
3.2 RT-PCR expression analysis in juvenile Atlantic cod tissues

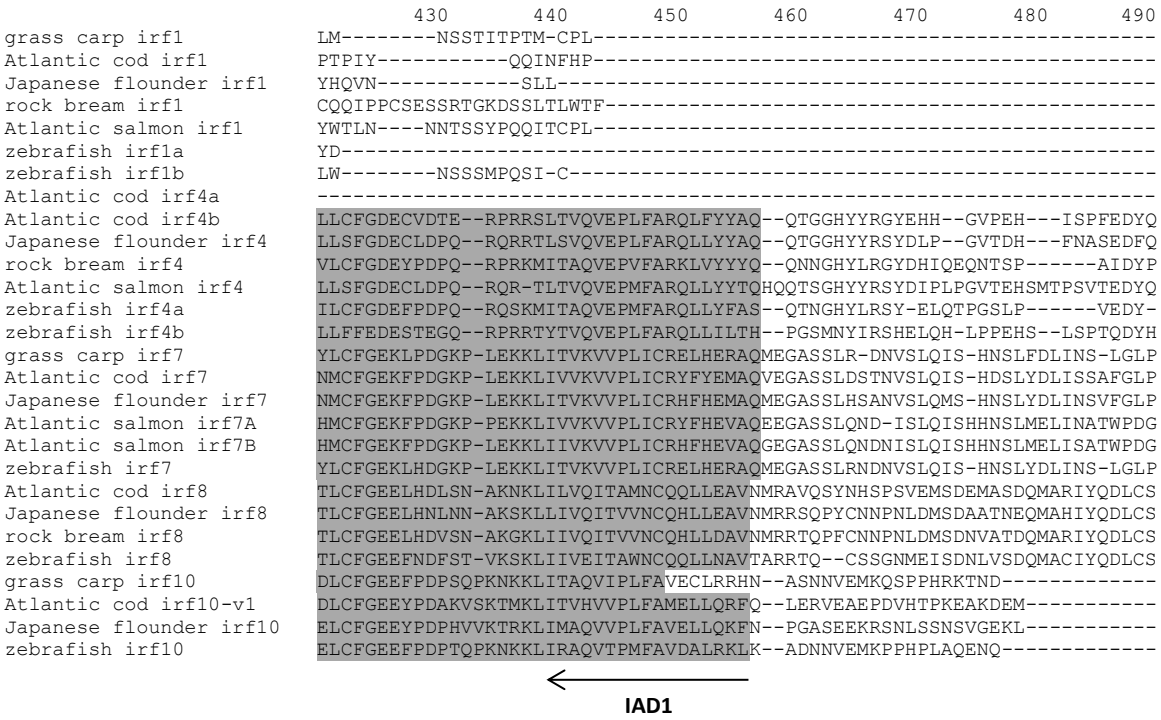
RT-PCR and agarose gel electrophoresis analysis was used to investigate constitutive transcript expression of cod IRF paralogues in 15 different tissues of juvenile Atlantic cod. While *Irf7*, *Irf8*, and *Irf10-v1* all appeared to be expressed at a moderate

Figure 13: Multiple sequence alignment of Atlantic cod IRF1, IRF4, IRF7, IRF8, and IRF10 protein sequences with homologous sequences from other teleost fish species. Sequences were retrieved from the NCBI non-redundant protein database (see Table 6, below). Alignment was carried out using the ClustalW algorithm in MEGA5 software (Tamura *et al.*, 2011). Identical amino acids are indicated by asterisks (*); conservative substitutions are indicated by colons (:). DNA binding domain and IRF-associated domain are shaded in grey and marked "DBD" and "IAD1" respectively; conserved tryptophan residues are boxed. The translation of the shorter IRF10 splice variant (*Irf10-v2*) was not included in the alignment.


10 20 30 40 50 60 70
grass carp IRF1 ----------MPVSRMRMRPWLESRIDSNTIAGLVWVNKEEKMFSIPWKHAARHGWEVDKDACLFKQW
Atlantic cod IRF1 ----------MPVARMKMRPWLERMIESNKVPGLSWVDKDQKMFAITWKHAARHGWQVEKDASLFKHW
Japanese flounder IRF1 ----------MPVSRMRMRPWLEKMIESNTISGLTWVDKDQKMFSIPWKHAARHGWELDKDASLFKKW
rock bream IRF1 ----------MPVSRMRMRPWLEQQIESNSISGLHWVDKDKTMFSIPWKHAARHGWELDKDACLFKQW
Atlantic salmon IRF1 ----------MPVSRMRMRPWLEEKIESNSISGLVWVDKDNKIFSVPWKHAARHGWDLNKDACLFKQW
zebrafish IRF1A ----------MHQGRLRLRPWLEEQIQSGRYPGVQWLDQSARVFQIPWKHAARHGWNIDKDATLFRNW
zebrafish IRF1B ----------MPVSRMRMRPWLESRIDSNTINGLMWVNKEEKMFSIPWKHAARHGWEVDKDACLFKQW
Atlantic cod IRF4A -MHFEEDVNLS-VSCGNGKLRQWLIDQIDSKSYLGLVWENVEKSIFRIPWKHAGKQDYNRDEDAALFKAW
Atlantic cod IRF4B -MNLEADYTAT-GSSGNGKLRQWLIDQVDSGTYPGLIWENDEKSIFRIPWKHAGKQDYNRDEDAALFKAW
Japanese flounder IRF4 -MNPELDYGGS-GSGGNGKLRQWLIEQVDCGKYPGLVWENDEKSIFRIPWKHAGKQDYNRDEDAALFKAW
rock bream IRF4 -MNLEEDSGLS-VSCGNGKLRQWLIDQIDSRRYAGLVWENDEKSIFRIPWKHAGKQDYNREEDAALFKAW
Atlantic salmon IRF4 -MNPESDYGMSTVSCGNGKLRSWLIEQVDTGKYPGLVWENEEKSIFRIPWKHAGKQDYNRDEDAALFKAW
zebrafish IRF4A -MNLDGDCIMS-VSCGNGKLRQWLIEQIDSGEYSGLVWENDEKTIFRIPWKHAGKQDYNRDEDAALFKAW
zebrafish IRF4B --------------SGNGKLRQWLIEQVDTGKYPGLVWENDEKSIFRIPWKHAGKQDYNRDEDAALFKAW
grass carp IRF7 -------MAAMQSTIGKPQFGPWLIEQVESGRYEGLRMIGNDI--FRIPWKHNSRRDLG-DEDIKIFKEW
Atlantic cod IRF7 ----------MQS-SHKPLFANWLIEQVETGNYPGLSYISTNL--FRVPWKHNSRKDCN-DEDCKIFRAW
Japanese flounder IRF7 ----------MQS-LPKPQFASWLIEQVETGQYTGLRYVAENK--FRVPWKHNSRKDCR-DEDSKIFRAW
Atlantic salmon IRF7A ----------MQS--CKPQFADWLIEQVRTEQYTGLFFMDNNK--FRVPWKHNSRKDCS-EDDRKIFRAW
Atlantic salmon IRF7B MTEVRGSALTMQSRNPKPQFADWLIEQVWTGQYAGLYFVGNNK--FRVPWNHISRKDCC-EDDSKIFRAW
zebrafish IRF7 ----------MQSTNAKPQFGPWLIEQVESGQYEGLSMIGHDI--FRIPWKHNARRDLG-DADVKIFKEW
Atlantic cod IRF8 ----------MSNTGGRRLKQWLIEQIKSGQYSGLEWEDDSLTMFRIPWKHAGKQDYNQEVDASIFKAW
Japanese flounder IRF8 ----------MSNPGGRRLKQWLVEQIHSGQYAGLQWEDESRTMFRIPWKHAGKQDYNQEVDAFIFKAW
rock bream IRF8 ----------MSNTGGRRLKQWLVEQIQSAQYSGLQWEDESRTLFRIPWKHAGKQDYNQEVDASIFKAW
zebrafish IRF8 ----------MNSGGRRLKQWLIEQINSNIYNGLQWEDEDRTMFRIPWKHAGKQDYNQEVDASIFKAW
grass carp IRF10 -ME---------DRSRHMRLREWLIAQIDSGKYAGLSWENEEKTMFRIPWKHAAKQDYRQNQDAALFKAW
Atlantic cod IRF10-V1 -ME---------GDG-KMHLKEWLIAQVDSERFDGLRWENEEKTMFRIPWKHAAKKDYRQQDDAALFKAW
Japanese flounder IRF10 -ME---------EGA-KLHLKEWLISQIESGRYEGLSWEDEDRTMFRIPWKHAAKKDYKQTEDAALFKAW
zebrafish IRF10 -ME---------DRSRHMRLREWLIAQIDSAEYPGLSWENAEKSMFRIPWKHAAKQDYRQNQDAALFKAW
: *** : *: * : * * : * :*: *
DBD
80 90 100 110 120 130 140
grass carp IRF1 AIHTGKFREGVTTPDPKTWKANFRCAMNSLPDIEEVKDKSINKGCGAVRVYRMLPAVSKKK---------
Atlantic cod IRF1 AIHTGKFKEGVDESDPKKWKANFRCAMNSLPDVEQVKGKNVNKGQQAVRVYKMVEVTATK----------
Japanese flounder IRF1 AIHTGKYTEG-QTSDPKTWKANFRCAMNSLPDIEEVKDKSIHKGQQAVRVFKMLHVTPKS----------
rock bream IRF1 AIHTGKYVEG-QACDPKTWKANFRCAMNSLPDIEEVKDKSVNKGHQAMRVFRMLPSLPKSR---------
Atlantic salmon IRF1 AMHTGKFIQGETKTDPKTWKANFRCAMNSLPDIEEVKDKSINRGSGAVRVYKMKNIYSKPN---------
zebrafish IRF1A AIHTGRYKPGIDKPDPKTWKANFRCALNSLTDVKELQDRSIKKGHNAFRVYALLPHCKTIR---------
zebrafish IRF1B AIHTGKYKEGVTQPDPKTWKANFRCAMNSLPDIEEVKDKSINKGCGAVRVYRMLPAVS-KK---------
Atlantic cod IRF4A ALFKDKYKEGVDKPDPPTWKTRLRCALNKSNDFDELVDRSQLDITEPYKVYRIIPEGVKRG--KPINKVS
Atlantic cod IRF4B ALFKGKFREGIDKADPPTWKTRLRCALNKSNDFEELVDRSQLDISDPYKVYRIIPEGDKR-----RPRQE
Japanese flounder IRF4 ALFKGKFREGIDKPDPPTWKTRLRCALNKSNDFVELVERSQLDISDPYKVYRIIPEGAKK-----RPRQE
rock bream IRF4 ALFKGKYKEGVDKPDPPTWKTRLRCALNKSNDFDELVERSQLDISEPYKVYRIIPEEAKKG--MKMSSME
Atlantic salmon IRF4 ALFKGKFREGIDKPDPPTWKTRLRCALNKSNDFEELVQRSQLDISDPYKVYRIIPECAKKHFLLSGSKQE
zebrafish IRF4a ALFKGKYREGLDKPDPPTWKTRLRCALNKSNDFDELVERSQLDISDPYKVYRIVPEGAKRG--SKAISME
zebrafish IRF4B ALFKGKFREGVDKPDPPTWKTRLRCALNKSNDFEEIVERSQLDISDPYKVYRIVPEGSKK----GSRSIE
grass carp IRF7 AVVSGKINEH--PNDKAKWKTNFRCALYSLKN-FEMLEDHSKDPDDQHKVYRIIRPQNHQEIQ----SAE
Atlantic cod IRF7 AVASGKIHEF--PNDKAKWKTNFRCALKNLNKRFRMSKDNSKNSDDPHKIYEIINREAAYQ-PSPPEED-
Japanese flounder IRF7 AVASGKINEF--PNDKARWKTNFRCALNNLSVRFKMIEDNSKHSDDPHKIYQIMNTEHRQENCSMPKNDS
Atlantic salmon IRF7A AVVSGKITEH--PNDKAKWKTNFRSALNSLCRRFKMVEDHSKDSNDPHKVYLVIN-EYNYESPLIEEITL
Atlantic salmon IRF7B AVVSGKINTH--PNDKAKWKTNFRCVLNNLTKRFMMVEDHSKDSDDPHKVYLIINNESNYGSPHIEEIAV
zebrafish IRF7 AIVSGKINEY--PNDKAKWKTNFRCALHSLKN-FEMLEDHSKDPDDQHKIYRIIRPQNHQEIQSAIQSAE
Atlantic cod IRF8 AIVSGKFKEG-EKAEPATWKTRLRCALNKSPDFEEVTDRSQLDISEPYKVYRIVPEEEQK----------
Japanese flounder IRF8 AVFKGKFKEG-DKAEPATWKTRLRCALNKSPDFEEVTERSQLDISEPYKVYRIVPEEEQK----------
rock bream IRF8 AVFKGKFKEG-DKAEPATWKTRLRCALNKSPDFEEVTERSQLDISEPYKVYRIVPEEEQK----------
zebrafish IRF8 AIFKGKFKEG-DKAEPATWKTRLRCALNKSPDFEEVTDRSQLDISEPYKVYRIVPEEEQK----------
grass carp IRF10 AMYKGKFQEGRDKADPSTWKTRLRCALNKSTDFQEVPERSQLDISEPYKVYRILDD--------------
Atlantic cod IRF10-V1 AVYKGKYKVGSDKDNPTMWKTRLRCALNKSTDFQEVPHLNQLDISEPYKVYRIESD----Q----------
Japanese flounder IRF10 AVYKGKYIEGRDKADPTMWKTRLRCALNKSTDFQEVPERNQLDITEPYKVYRIQQDSGSVR----------
zebrafish IRF10 AMYKGKFQEGRDKADPSTWKTRLRCALNKSTDFQEVSERSQLDISEPYKVYRILED--------------
*: : ** :* :: : ::: :
DBD


150 160 170 180 190 200 210
grass carp irf1
Atlantic cod irf1
Japanese flounder irf1
rock bream irf1
Atlantic salmon irf1
zebrafish irf1a
zebrafish irf1b
Atlantic cod irf4a
Atlantic cod irf4b
Japanese flounder irf4
rock bream irf4
Atlantic salmon irf4
zebrafish irf4a
zebrafish irf4b
grass carp irf7
Atlantic cod irf7
Japanese flounder irf7
Atlantic salmon irf7A
Atlantic salmon irf7B
zebrafish irf7
Atlantic cod irf8
Japanese flounder irf8
rock bream irf8
zebrafish irf8
grass carp irf10
Atlantic cod irf10-v1
Japanese flounder irf10
zebrafish irf10
220 230 240 250 260 270 280
grass carp irf1
Atlantic cod irf1
Japanese flounder irf1
rock bream irf1
Atlantic salmon irf1
zebrafish irf1a
zebrafish irf1b
Atlantic cod irf4a
Atlantic cod irf4b
Japanese flounder irf4
rock bream irf4
Atlantic salmon irf4
zebrafish irf4a
zebrafish irf4b
grass carp irf7
Atlantic cod irf7
Japanese flounder irf7
Atlantic salmon irf7A
Atlantic salmon irf7B
zebrafish irf7
Atlantic cod irf8
Japanese flounder irf8
rock bream irf8
zebrafish irf8
grass carp irf10
Atlantic cod irf10-v1
Japanese flounder irf10
zebrafish irf10
IAD1


290 300 310 320 330 340 350
grass carp irf1 SSSGL----------YTSRFQVSPMHSTDLED--YEAIIEISRQLERDT-
Atlantic cod irf1 SNNADY----------FYRRFEVSPEHPPEFED--AEELLKLCQQLEPETN
Japanese flounder irf1 SFPSN----------FCPRFQVSPDHSPDYSY--SDDIVEICKQLERESH
rock bream irf1 SFQSN----------FHHRFEVSPERSSDYDY--TDDIIQICQELEKESH
Atlantic salmon irf1 STNNF----------YAS-FQVSPDHSTDYEDGHQETLIGMTHHWEQGS-
zebrafish irf1a SEEER----------AQGLQIN--RTDEHEQ--TEAVLKIVDHLKTLDH
zebrafish irf1b STCND----------IYSRFQVSPVHSTDLED--SEAILELTRQLERDSS
Atlantic cod irf4a ----------
Atlantic cod irf4b TISNPKGCHLI------PWALEEKAYVSP----GAPDLVPLPPEGLTL-QRMAGEE--GPPSSLAMQGVR
Japanese flounder irf4 TTSSPKGCHIT-----PCSPEEKLSLLP----GGPDVVPLPVDHLSV-QRRAEECSPNPP-STLERGVL
rock bream irf4 TTTSPEGCRITSSSSSSPSSSSSSSPCPEDKFHSGAEVILFPFPYPES-HRQGAEM----LPNVLERGVL
Atlantic salmon irf4 TTSSPEGCRIA------PCSPDDKLYSPT----VGPDLVPLPLDSLQA-LGRGEECPPSPPGCTLERGVL
zebrafish irf4a TTSSPEGCRISSSAS--PGSPSSPSSPSEERLYGGAEPVLFPFPYPQS-QRRGAEK----LPNVLERGVL
zebrafish irf4b TVSSPEGCQLG------PSR-EGQAYASP----GAPELVELP----------HADGVPLERGVI
grass carp irf7 RLCSSL----------VQFYYQCDPSE----LRGEPIRFPTTEGLI--DHKQIQFTKRILDSIQRGLQ
Atlantic cod irf7 QVSWPR----------VQLHYGNEATE----LQARPICFPPTDTLR--DHKQVEFTNRILSSIQRGLL
Japanese flounder irf7 TLATAR----------LQLHYQHEAPD----LNAHPLCFPSTDGLL--DHKQIEYTNRILNSIQRGLL
Atlantic salmon irf7A QVSGPR----------VQLHYQCNALE----PNTQPLCFPSTDGLP--DLKQIEYTNCILRSVQRGLL
Atlantic salmon irf7B QVSGPL----------VQLHYQCDIPE----PNAQTLCFPSTDGLL--DHKQIEYTNRILGSVQRGLL
zebrafish irf7 RLCSSL----------IHFFYQCDPSE----LRGEQIRFPTTECLI--DVKQIQYTKRILDSIQRGLQ
Atlantic cod irf8 VTSHPEGCRIS-------PVLPQQRAVARGYSSDTMQSVHFPPADLID--NERQRQVTCKLLGHLERGVL
Japanese flounder irf8 LVTHPEGCRIS-------PQQHLGRSIL--YSSDSMQNVHFPPAELIE--YDRQRHVTCKLLGHVERGVL
rock bream irf8 LVAHPEGCRIS-------PQQHLGRGAL--YSSDSMQCVNFPPAELIE--YDRQRHVTRKLLGHLERGVL
zebrafish irf8 VTTHPEGCRIS-------PCLP-STANGFLYGSDSLQNIYFPSIDGIK--NERQRHVTRKLFSHLERGVL
grass carp irf10 TTASPDGCFILQG----CAPVGNERIYGP----CEAEKVFFPRPDTIR-LPPGIAEAMSRLLPHLEKGVL
Atlantic cod irf10-v1 TSSSPEGCFILQG----CVPLGNERIYGP----CSAQQLSLPSPASLGPLEPGVARALGQLLSHLERGVL
Japanese flounder irf10 ITSSPEGCFILQG----HVPWGNERIYGP----CTAQQLSFPSPGSVS-LPSCMAEAMNRLLCHLERGVL
zebrafish irf10 LSCSPDGCFLLQG----CAPVGSERIYGP----CAATQLFFPPPNAAM-LPTGICEAMTRLLPHLEKGVL
360 370 380 390 400 410 420
grass carp irf1 LLLQNGA-FPKGFLANEVGTSESL-------SPQSHWSVSS-GEELE-FRLYTELS--PEE--SICTYTE
Atlantic cod irf1 WMQSSSDDRLSSGLHSDSNY-----------SPHSQWSDTSSGEDLD-MRLYTDLSTGTECYSPETWNMF
Japanese flounder irf1 FMPSSLD--VMGFLNNEPCT-----------SPGSHWSDSSSADELDELPHYTNLSSETA--TDALWNGL
rock bream irf1 WMTSSLD--GNGFLSNEACT-----------SPGSAWSESS-SDELEDMPQYTTLGSDLTNPTDDLWNSF
Atlantic salmon irf1 --VND-----KGFQSNEVGTAESFDTAESYHSQESQWSDNS-ETEIE-LRLYTELSSGLPIIDDILSYTD
zebrafish irf1a WASSYDG--ERGWR-----------------PNSTWTGCL-GETVD----FPAFSFQTDCNLHTISPAQ
zebrafish irf1b QWLQN---FGKGFLANEVCTTESL-------SPESQWSVSS-GEELE-LRLYTELT--PDLRTDSYTYTE
Atlantic cod irf4a ----------
Atlantic cod irf4b LWMTPEGLYARRQCQESVYWKEGVSP---YKDKLNEMEREVNCKVLDTQDFLTEIQSYGLHGRPIPPFQA
Japanese flounder irf4 LWMGADGLYACRLCQSRVYWQGGPSP---YGDKLNKLERDVTCKLLHSQDYLTELQSFGLHGRPLPRLQV
rock bream irf4 LWMMSDGLYAKRLCQGRVYWEGPLAP---YMDKPNKLEKEQPCKLFDTQQFLIELQDFAHNGRHLPRLQV
Atlantic salmon irf4 LWMAPDGLYARRLCQERVFWEGGLSS---YADKPNKLEREHTCKLLHTQDYLTELQGYALHCRPPPRLQV
zebrafish irf4a LWLSPDGLYAKRLCQGRVYWEGPLAP---YADKPNKLEKEQTCKLMDTQQFLTELQGFIHHGRPMPRSQV
zebrafish irf4b LWMAPDGLYARRCCPCRVYWTGAHAP---PTDKPNKLEREQNCKLLDTHLFITELQSYTLHARPAPCSQV
grass carp irf7 LEVNQYGIYGFRQDKCKVFVSTSDPS-EIQNPEPRKLHQNSREQLFSFDKYIRDLMDFKENRRGSPDYTI
Atlantic cod irf7 LEVRESGLYACRQDRCHVFASTADPS-QAS-PDPQKLPQNTLVELLSFEKFVKELKEFKENRRGSPEYVV
Japanese flounder irf7 LEVCETGIYAWRQDRCHVFASTSDPS-VAL-PDPRKLPQNTMVQLLSFEKYVNELKKFKENNGGSPDYTI
Atlantic salmon irf7A LEVQNTGIYGYRQDKCHVFSSTSNPR-EAH-PEPRKMPQNEMVQLLNFQQYENELIAFKENRRGSPDYTI
Atlantic salmon irf7B LEVRNTGIYGYRQDKCHVFSSTSDPR-EAH-PEPRKMPQNEMVQLLSFDKYMTDLIAFKENRGGSPDYTI
zebrafish irf7 LEVNQYGIYGFRQDKCKVFVSTSDPC-EIQKPEPRKLQQNYKEQLLSFDKYIRDLLDFKENRGGSPDYTI
Atlantic cod irf8 VRANREGVFIKRLCQSRVFWSGHGEHGQHHGPVTCKLERDAVVKIFDTGRFLHALQLHQEGQIPAPDPTV
Japanese flounder irf8 VRSNQEGIFIKRLCQSRVFWSGLGDVGSPYSSVPCKLERDAVVKIFDTGRFLQAVQLYQEGQLPAPDPTV
rock bream irf8 VRANQEGIFIKRLCQSRVFWSGLGEVGSQYSPMPCKLERDAVVKIFDTERFLQALQLYQEGQFPAPDPTV
zebrafish irf8 LRANREGIFIKRLCQSRVFWIGQDAR---YN--PCKLERDAVVKIFDTARFLQALQLYQDGHYQAPEPTV
grass carp irf10 VWVAPDGVFIKRFCQGRVYWDGPLAE---HRQKPNKLERERTCKLLDMTIFMQELQSHQQATGPEPRYTV
Atlantic cod irf10-v1 LWVAPDGLFIKRFCQGRVYWSGPLAP---HTEKPNKLERDRTCKLLDMPVFVNELQNYMQRKGPQPNYEI
Japanese flounder irf10 LWVAPDGVFIKRFCQGRVYWSGPLAQ---HTDAPNKLEREKTFKLLDIPRFVSELQRSLWGKGPAPSYEI
zebrafish irf10 LWVAPDGLFIKRFCQGRVYWDGPLAE---HRHKPNKLERERTCKLLDMKIFSQELLNYRQGIGPEPQYIV


430 440 450 460 470 480 490
grass carp irf1 LM--------NSSTITPTM-CPL----------
Atlantic cod irf1 PTPIY-----------QQINFHP----------
Japanese flounder irf1 YHQVN------------SLL----------
rock bream irf1 CQQIPPCSESSRTGKDSSLTLWTF----------
Atlantic salmon irf1 YWTLN----NNTSSYPQQITCPL----------
zebrafish irf1a YD----------
zebrafish irf1b LW--------NSSSMPQSI-C----------
Atlantic cod irf4a ----------
Atlantic cod irf4b LLCFGDECVDTE--RPRRSLTVQVEPLFARQLFYYAQ--QTGGHYYRGYEHH--GVPEH---ISPFEDYQ
Japanese flounder irf4 LLSFGDECLDPQ--RQRRTLSVQVEPLFARQLLYYAQ--QTGGHYYRSYDLP--GVTDH---FNASEDFQ
rock bream irf4 VLCFGDEYPDPQ--RPRKMITAQVEPVFARKLVYYYQ--QNNGHYLRGYDHIQEQNTSP------AIDYP
Atlantic salmon irf4 LLSFGDECLDPQ--RQR-TLTVQVEPMFARQLLYYTQHQQTSGHYYRSYDIPLPGVTEHSMTPSVTEDYQ
zebrafish irf4a ILCFGDEFPDPQ--RQSKMITAQVEPMFARQLLYFAS--QTNGHYLRSY-ELQTPGSLP------VEDY-
zebrafish irf4b LLFFEDESTEGQ--RPRRTYTVQVEPLFARQLLILTH--PGSMNYIRSHELQH-LPPEHS--LSPTQDYH
grass carp irf7 YLCFGEKLPDGKP-LEKKLITVKVVPLICRELHERAQMEGASSLR-DNVSLQIS-HNSLFDLINS-LGLP
Atlantic cod irf7 NMCFGEKFPDGKP-LEKKLIVVKVVPLICRYFYEMAQVEGASSLDSTNVSLQIS-HDSLYDLISSAFGLP
Japanese flounder irf7 NMCFGEKFPDGKP-LEKKLITVKVVPLICRHFHEMAQMEGASSLHSANVSLQMS-HNSLYDLINSVFGLP
Atlantic salmon irf7A HMCFGEKFPDGKP-PEKKLIVVKVVPLICRYFHEVAQEEGASSLQND-ISLQISHHNSLMELINATWPDG
Atlantic salmon irf7B HMCFGEKFPDGKP-LEKKLIIVKVVPLICRHFHEVAQGEGASSLQNDNISLQISHHNSLMELISATWPDG
zebrafish irf7 YLCFGEKLHDGKP-LEKKLITVKVVPLICRELHERAQMEGASSLRNDNVSLQIS-HNSLYDLINS-LGLP
Atlantic cod irf8 TLCFGEELHDLSN-AKNKLILVQITAMNCQQLLEAVNMRAVQSYNHSPSVEMSDEMASDQMARIYQDLCS
Japanese flounder irf8 TLCFGEELHNLNN-AKSKLLIVQITVVNCQHLLEAVNMRRSQPYCNNPNLDMSDAATNEQMAHIYQDLCS
rock bream irf8 TLCFGEELHDVSN-AKGKLIIVQITVVNCQHLLDAVNMRRTQPFCNNPNLDMSDNVATDQMARIYQDLCS
zebrafish irf8 TLCFGEEFNDFST-VKSKLIIVEITAWNCQQLLNAVTARRTQ--CSSGNMEISDNLVSDQMACIYQDLCS
grass carp irf10 DLCFGEEFPDPSQPKNKKLITAQVIPLFAVECLRRHN--ASNNVEMKQSPPHRKTND----------
Atlantic cod irf10-v1 DLCFGEEYPDAKVSKTMKLITVHVVPLFAMELLQRFQ--LERVEAEPDVHTPKEAKDEM----------
Japanese flounder irf10 ELCFGEEYPDPHVVKTRKLIMAQVVPLFAVELLQKFN--PGASEEKRSNLSSNSVGEKL----------
zebrafish irf10 ELCFGEEFPDPTQPKNKKLIRAQVTPMFAVDALRKLK--ADNNVEMKPPHPLAQENQ----------
IAD1

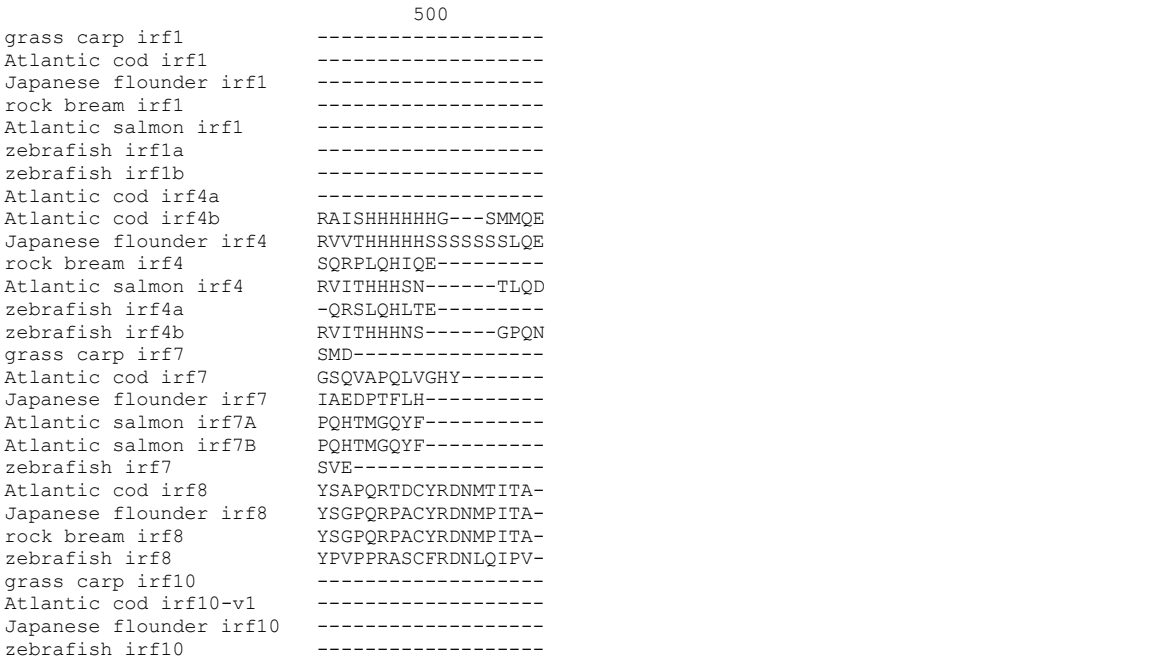

500
grass carp irf1 ----------
Atlantic cod irf1 ----------
Japanese flounder irf1 ----------
rock bream irf1 ----------
Atlantic salmon irf1 ----------
zebrafish irf1a ----------
zebrafish irf1b ----------
Atlantic cod irf4a ----------
Atlantic cod irf4b RAISHHHHHHG---SMMQE
Japanese flounder irf4 RVVTHHHHHSSSSSSSLQE
rock bream irf4 SQRPLQHIQE---------
Atlantic salmon irf4 RVITHHHSN------TLQD
zebrafish irf4a -QRSLQHLTE---------
zebrafish irf4b RVITHHHNS------GPQN
grass carp irf7 SMD----------
Atlantic cod irf7 GSQVAPQLVGHY-------
Japanese flounder irf7 IAEDPTFLH----------
Atlantic salmon irf7A PQHTMGQYF----------
Atlantic salmon irf7B PQHTMGQYF----------
zebrafish irf7 SVE----------
Atlantic cod irf8 YSAPQRTDCYRDNMTITA-
Japanese flounder irf8 YSGPQRPACYRDNMPITA-
rock bream irf8 YSGPQRPACYRDNMPITA-
zebrafish irf8 YPVPPRASCFRDNLQIPV-
grass carp irf10 ----------
Atlantic cod irf10-v1 ----------
Japanese flounder irf10 ----------
zebrafish irf10 ----------

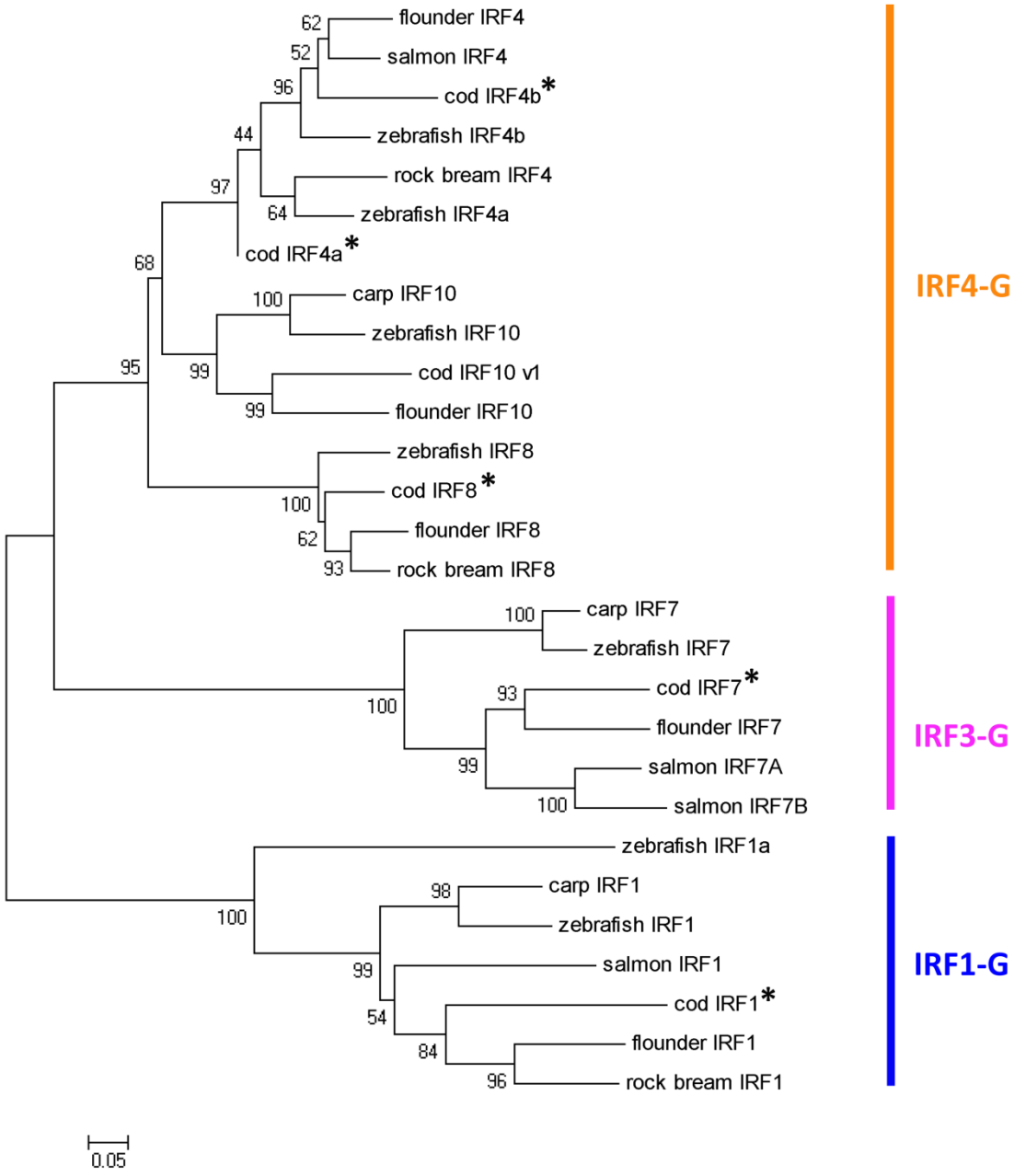
Table 6: Fish IRF amino acid sequences used in multiple sequence alignment and phylogenetic analysis

Protein name	Species name (common name)	GenBank accession no.
IRF1	*Ctenopharyngodon idella* (grass carp)	ADF57571.1
	Gadus morhua (Atlantic cod)	ADG85733.1
	Paralichthys olivaceus (Japanese flounder)	BAA83468.1
	Oplegnathus fasciatus (rock bream)[1]	ADJ21809.1
	Salmo salar (Atlantic salmon)	NP_001117117.1
IRF1a	*Danio rerio* (zebrafish)	NP_001035442.1
IRF1b[2]		AAH85555.1
IRF4	*Paralichthys olivaceus* (Japanese flounder)	AEY55358
	Oplegnathus fasciatus (rock bream)	AFU81289
	Salmo salar (Atlantic salmon)	NP_001133454.1
IRF4a	*Danio rerio* (zebrafish)	NP_001116182.1
IRF4b		CAI11951.1
IRF7	*Ctenopharyngodon idella* (grass carp)	ACS34986
	Paralichthys olivaceus (Japanese flounder)	ACY69214.1
	Danio rerio (zebrafish)	NP_956971.1
IRF7A	*Salmo salar* (Atlantic salmon)	NP_001130020.1
IRF7B		NP_001165321.1
IRF8	*Paralichthys olivaceus* (Japanese flounder)	AFE18694
	Oplegnathus fasciatus (rock bream)	AFU81290
	Danio rerio (zebrafish)	NP_001002622
IRF10	*Ctenopharyngodon idella* (grass carp)	ACT83676.1
	Paralichthys olivaceus (Japanese flounder)	BAI63219
	Danio rerio (zebrafish)	NP_998044

[1]*Oplegnathus fasciatus* is more commonly called barred knifejaw or striped beakfish, but is called rock bream in publications describing IRF genes in that species.

[2]Zebrafish IRF1b is also called IRF11

Figure 14: Phylogenetic analysis of Atlantic cod IRF family members. Putative cod amino acid sequences were aligned with IRF proteins from selected other teleost fish species using MEGA5 software (Tamura *et al.,* 2011). Based on the multiple sequence alignment, the evolutionary history was inferred using the neighbour-joining method. The bootstrap consensus tree was constructed from 5000 replicates, where numbers at the branch points represent percentage of replicates in which the associated taxa grouped together. Branch lengths are proportional to calculated evolutionary distances. Sequences determined from this study are indicated with an asterisk (*). IRF family subgroups are indicated using colours matching Figure 1.


flounder IRF4
salmon IRF4
cod IRF4b*
zebrafish IRF4b
rock bream IRF4
zebrafish IRF4a
cod IRF4a*
carp IRF10
zebrafish IRF10
cod IRF10 v1
flounder IRF10
zebrafish IRF8
cod IRF8*
flounder IRF8
rock bream IRF8
carp IRF7
zebrafish IRF7
cod IRF7*
flounder IRF7
salmon IRF7A
salmon IRF7B
zebrafish IRF1a
carp IRF1
zebrafish IRF1
salmon IRF1
cod IRF1*
flounder IRF1
rock bream IRF1
62
52
96
44
97
64
68
100
95
99
99
100
62
93
100
100
93
99
100
98
100
99
54
84
96
0.05
IRF4-G
IRF3-G
IRF1-G

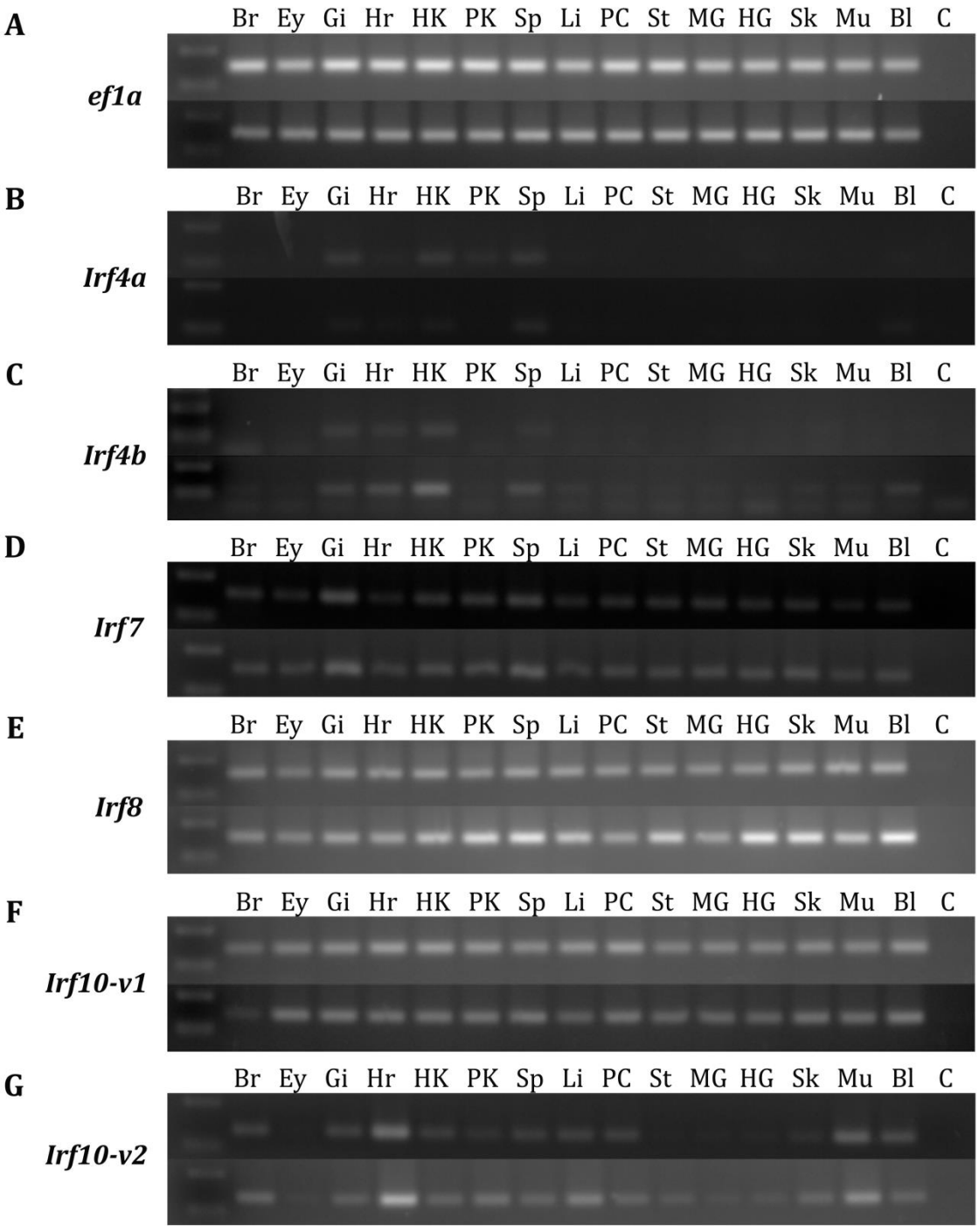
level in all tissues (using *ef1α* as a reference/housekeeping gene with relatively stable transcript expression in all tissues tested), *Irf4a*, *Irf4b* and *Irf10-v2* were absent or expressed at very low levels in some tissues (Figure 15). Interestingly, splice variants *Irf10-v1* and *Irf10-v2* appeared to have different patterns of expression: the shorter variant (v2) is apparently vastly reduced or absent in some digestive tissues (stomach, midgut, and hindgut) and in the eye, while the longer transcript is relatively evenly expressed in all 15 tissues. *Irf10-v2* was also unique among the transcripts studied in that the highest transcript expression appeared to be in the heart and skeletal muscle.

As a goal of this study was to better understand the roles of IRF-encoding transcripts in cod immune responses, expression in immune-relevant tissues (i.e. spleen, hematopoietic [head] kidney, blood) was of particular interest. All six transcripts were expressed in spleen and head kidney, and all except *Irf4a* were expressed in blood (*Irf4a* was faintly detected in only one replicate blood sample). All transcripts were also expressed in gill and heart tissues, although *Irf4a* expression in heart appeared to be much lower than that of the other transcripts (Figure 15B). The constitutive expression of all IRF transcripts in spleen supported the use of that organ for subsequent QPCR expression analyses.

3.3 Spleen transcript expression response to viral and bacterial antigens and increased temperature

Expression of cod IRF transcripts in response to injection with viral [poly(I:C)] and bacterial (ASAL) antigens at 10°C and 16°C was analyzed by QPCR. Interestingly, spleen transcript expression of *Irf4a*, the shorter *Irf4* paralogue, was observed to be

Figure 15: Composite agarose gel image of IRF family member transcript expression in 15 tissues of juvenile Atlantic cod. All gels are 1.7% agarose in TAE buffer, using 1 kb plus ladder (Invitrogen) as a size marker (100 bp and 200 bp bands are shown). PCR was carried out using samples from two fish; in each panel fish 1 is the top row and fish 2 is the bottom row. Br=brain, Ey=eye, Gi=gill, Hr=heart, HK=hematopoietic (head) kidney, PK=posterior (trunk) kidney, Sp=spleen, Li=liver, PC=pyloric caecum, St=stomach, MG=midgut, HG=hindgut, Sk=skin, Mu=skeletal muscle, Bl=blood, C=no-template control.


A
Br Ey Gi Hr HK PK Sp Li PC St MG HG Sk Mu Bl C
ef1a
B
Br Ey Gi Hr HK PK Sp Li PC St MG HG Sk Mu Bl C
Irf4a
C
Br Ey Gi Hr HK PK Sp Li PC St MG HG Sk Mu Bl C
Irf4b
D
Br Ey Gi Hr HK PK Sp Li PC St MG HG Sk Mu Bl C
Irf7
E
Br Ey Gi Hr HK PK Sp Li PC St MG HG Sk Mu Bl C
Irf8
F
Br Ey Gi Hr HK PK Sp Li PC St MG HG Sk Mu Bl C
Irf10-v1
G
Br Ey Gi Hr HK PK Sp Li PC St MG HG Sk Mu Bl C
Irf10-v2

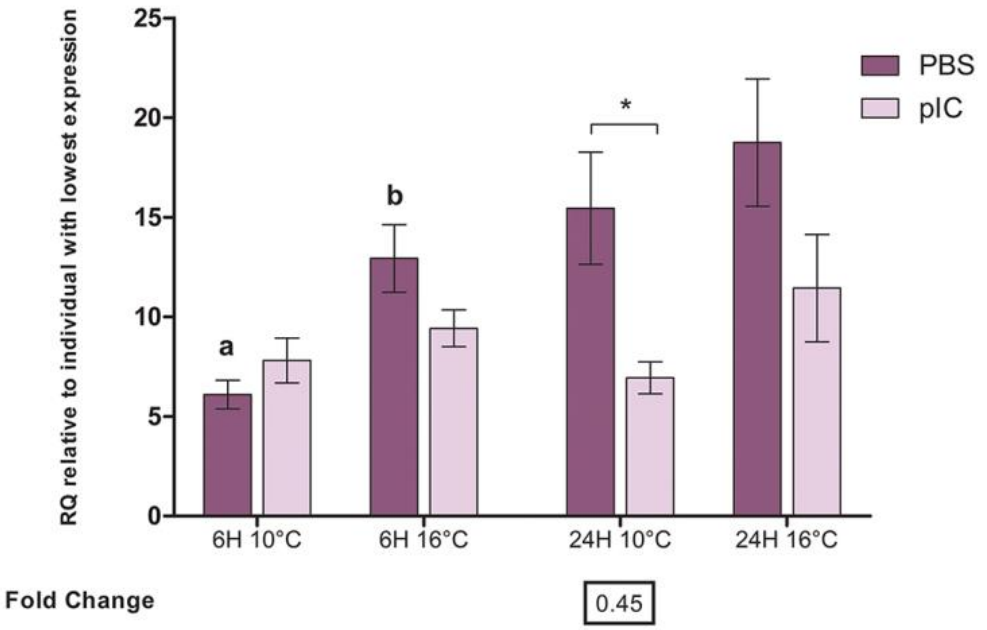
approximately 2-fold lower in poly(I:C) injected fish than PBS control fish at 10°C sampled 24 hours post injection (HPI) (Figure 16A). While there was no significant response to poly(I:C) for this paralogue at either temperature at the 6 HPI time-point, there was a significant increase in *Irf4a* transcript expression in the control (PBS) injected fish at 16°C compared with PBS fish at 10°C at that time point (Figure 16A). This temperature-dependant response of *Irf4a* (i.e. higher expressed at the elevated temperature at 6HPI) was also seen in ASAL-injected fish (Figure 16B), although ASAL injection itself did not have a significant effect (compared to time- and temperature-matched PBS controls) on *Irf4a* expression at the time points/temperatures studied.

Transcript expression of *Irf4b*, the longer IRF4 paralogue, was significantly upregulated in response to both poly(I:C) and ASAL injection at 6HPI compared with PBS controls (Figure 17). For poly(I:C) the change was seen only for fish held at 16°C (2.23-fold upregulated), while for ASAL it was observed at both 10°C and 16°C (1.98-fold and 3.41-fold upregulated, respectively). For both treatments the response was no longer observed at the 24HPI time-point. Changes in *Irf4b* transcript expression were also seen in response to increased temperature at the later time point, as expression was lower at 16°C than 10°C at 24HPI for all three [PBS, ASAL, poly(I:C)] treatment groups (Figure 17).

As noted above, the responses of Atlantic cod *Irf7* transcript expression to poly(I:C) and/or elevated temperatures have previously been investigated (Rise *et al.*, 2008; Hori *et al.*, 2012); therefore only the response to ASAL at two different temperatures was investigated in the current study for this transcript (Figure 18). Spleen

Figure 16: Spleen transcript expression responses of *Irf4a* to viral and bacterial antigens measured by QPCR. Data are presented as mean ± SEM, normalized to *ef1α* expression, with the lowest expressing sample set to RQ=1. Different letters represent significant differences between fish injected with PBS (lower case) or ASAL (upper case) at different temperatures within the same time-point. An asterisk (*) represents a significant difference between a poly(I:C) injected group and the time- and temperature-matched PBS injected group ($p < 0.05$). Fold change is calculated as [mean poly(I:C) RQ]/(mean PBS RQ). **A** = poly(I:C), **B** = ASAL. Note that poly(I:C) is abbreviated as pIC in the figure to conserve space.

A



Fold Change 0.45

B

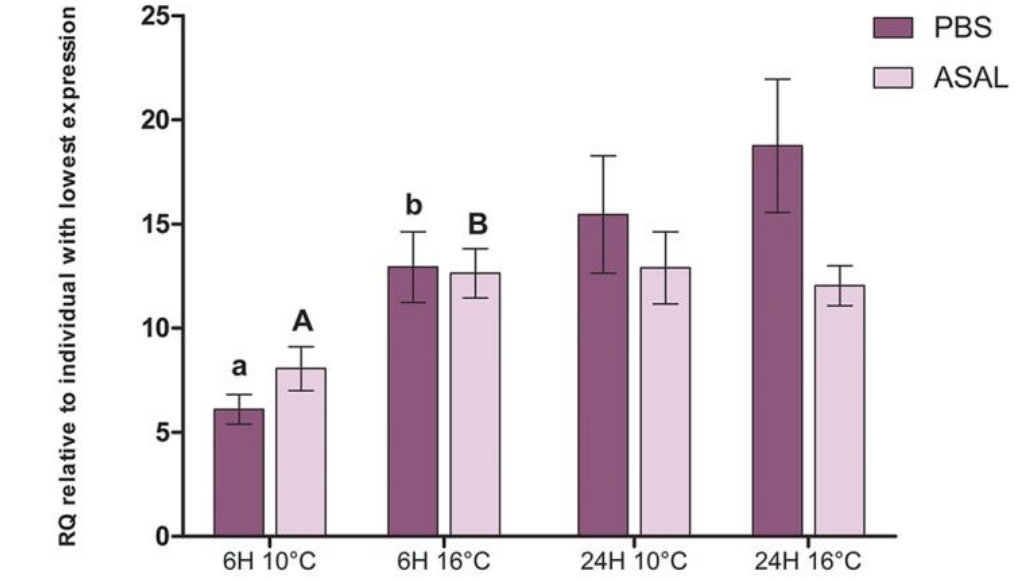
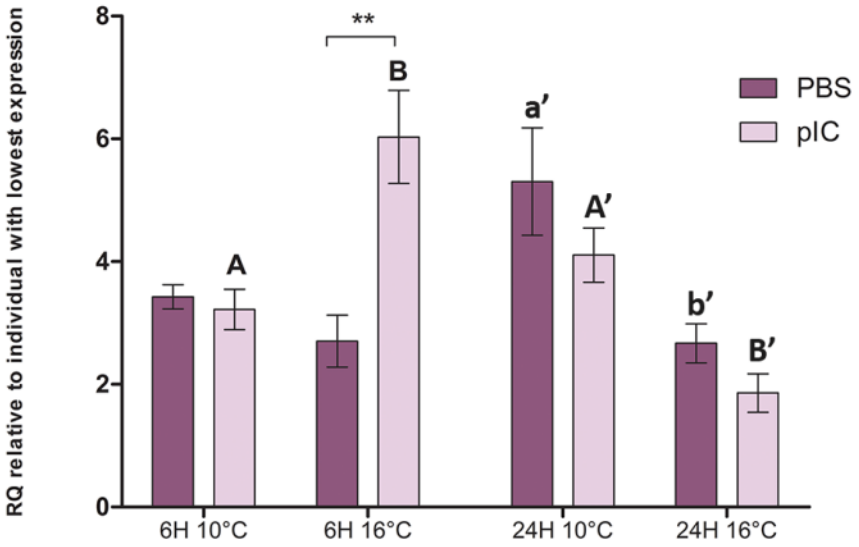


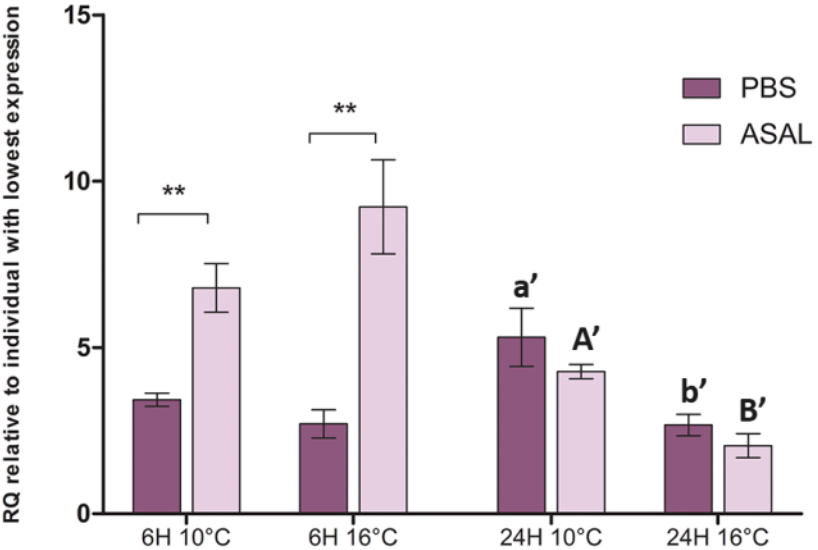
Figure 17: Spleen transcript expression responses of *Irf4b* to viral and bacterial antigens measured by QPCR. Data is presented as mean ± SEM, normalized to *ef1α* expression, with the lowest expressing sample set to RQ=1. Different letters represent significant differences between fish injected with PBS (lower case), or poly(I:C) or ASAL (upper case) at different temperatures within the same time point. Asterisks (*) represent significant differences between a poly(I:C) or ASAL injected group and the time- and temperature-matched PBS injected group (*p <0.05, ** p < 0.01). Fold change is calculated as [mean poly(I:C) or ASAL RQ] /(mean PBS RQ). **A** = poly(I:C), **B** = ASAL. Note that poly(I:C) is abbreviated as pIC in the figure to conserve space.

A



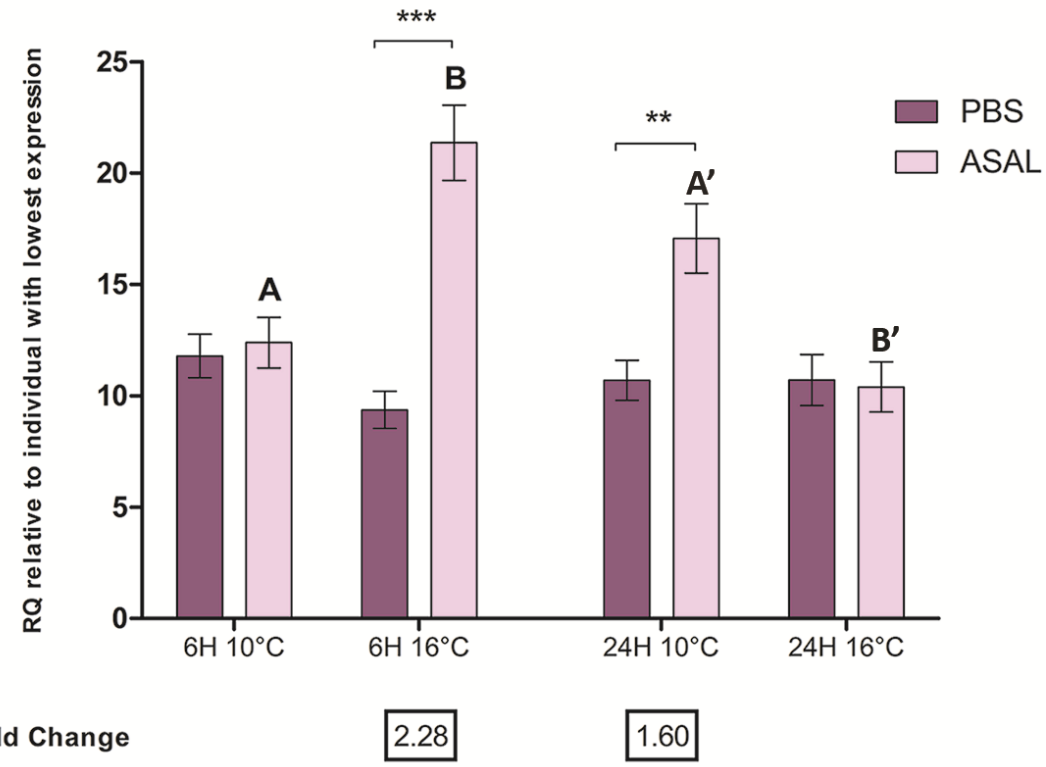
Fold Change		2.23		

B



Fold Change	1.98	3.41		

Figure 18: Spleen transcript expression responses of *Irf7* to bacterial antigens measured by QPCR. Data is presented as mean ± SEM, normalized to *ef1α* expression, with the lowest expressing sample set to RQ=1. Different letters represent significant differences between fish injected with ASAL at different temperatures within the same time point. Asterisks (*) represent significant differences between an ASAL injected group and the time- and temperature-matched PBS injected group (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). Fold change is calculated as (mean ASAL RQ)/(mean PBS RQ).


***
**
B
A'
A
B'
PBS
ASAL
RQ relative to individual with lowest expression
25
20
15
10
5
0
6H 10°C
6H 16°C
24H 10°C
24H 16°C
Fold Change
2.28
1.60

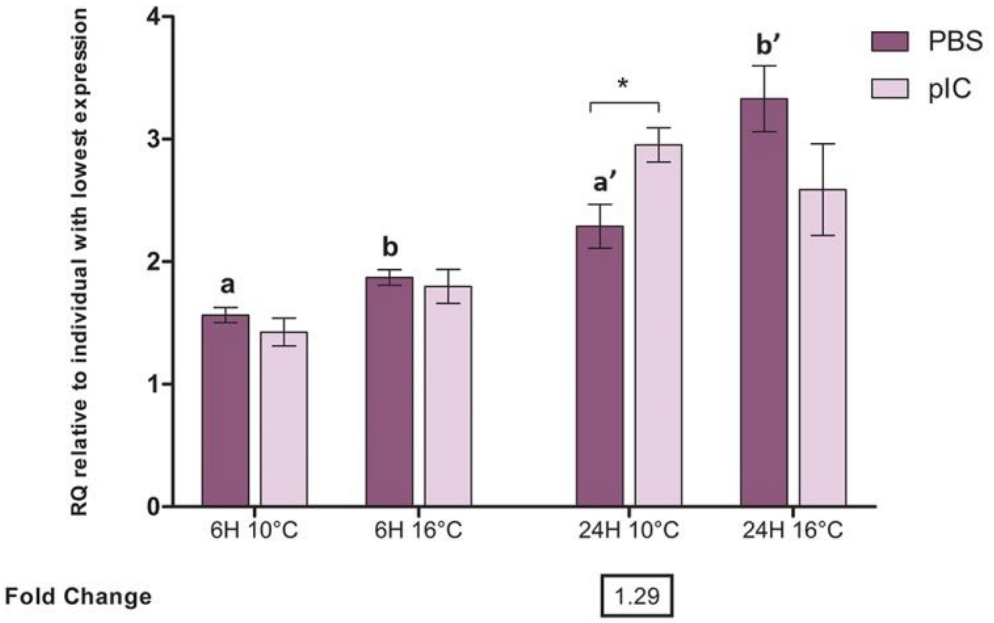
Irf7 transcript expression was seen to increase in response to ASAL injection (compared with time- and temperature-matched PBS controls) in fish held at both temperatures, although the increase was only at 6HPI for 16°C fish (2.28-fold upregulated), and only at24HPI for 10°C fish (1.60-fold upregulated) (Figure 18). Differences were also seen in ASAL injected fish at the same time point held at different temperatures, with transcript expression being higher at 16°C than 10°C at 6HPI and lower at 16°C than 10°C at 24HPI. In summary, the *Irf7* transcript expression response to bacterial antigens appeared to occur earlier at the elevated temperature.

Irf8 transcript expression was observed to increase in response to poly(I:C) only at 24HPI (1.29-fold upregulated compared with time and temperature matched PBS controls), in fish held at 10°C (Figure 19A). Response to ASAL injection, however, was similar to that of *Irf7*, as an increase in *Irf8* expression was observed at 6HPI for fish held at 16°C (2.17-fold) and at 24HPI for fish held at 10°C (1.45-fold) (Figure 19B). Interestingly, there was a small (1.25-fold) but statistically significant decrease in *Irf8* transcript expression in ASAL compared to PBS fish at 6HPI and 10°C. A response to temperature change was also seen in both ASAL and PBS injected fish, as *Irf8* expression was higher at 16°C than 10°C (at 6HPI for ASAL and at both time-points for PBS) (Figure 19B).

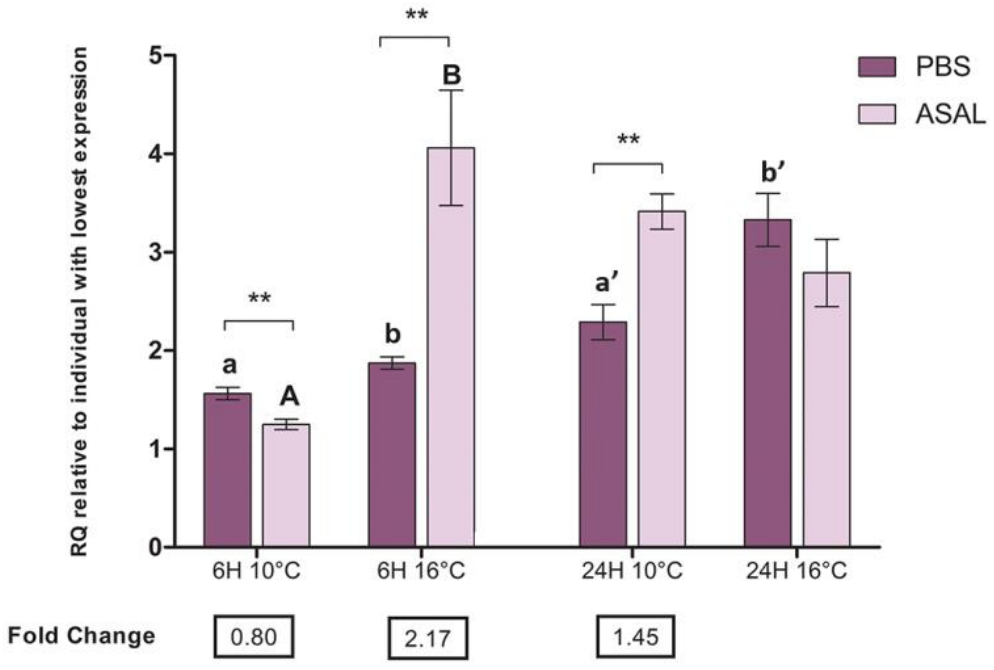
As with *Irf7*, the spleen transcript expression responses of cod *Irf10-v1* (the longer *Irf10* splice variant) to poly(I:C) and/or elevated temperature have previously been investigated (Rise *et al.*, 2008; Hori *et al.*, 2012). Therefore, only ASAL responsiveness of this transcript at the two temperatures was investigated in the current study.

Figure 19: Spleen transcript expression responses of *Irf8* to viral and bacterial antigens measured by QPCR. Data is presented as mean ± SEM, normalized to *ef1α* expression, with the lowest expressing sample set to RQ=1. Different letters represent significant differences between fish injected with PBS (lower case) or ASAL (upper case) at different temperatures within the same time point. Asterisks (*) represent significant differences between a poly(I:C) or ASAL injected group and the time- and temperature-matched PBS injected group (*p <0.05, ** p < 0.01). Fold change is calculated as [mean poly(I:C) or ASAL RQ] /(mean PBS RQ). **A** = poly(I:C), **B** = ASAL. Note that poly(I:C) is abbreviated as pIC in the figure to conserve space.

A


RQ relative to individual with lowest expression
4
3
2
1
0
a
b
a'
b'
*
PBS
pIC
6H 10°C
6H 16°C
24H 10°C
24H 16°C
Fold Change
1.29

B


RQ relative to individual with lowest expression
5
4
3
2
1
0
**
**
**
a
A
b
B
a'
b'
PBS
ASAL
6H 10°C
6H 16°C
24H 10°C
24H 16°C
Fold Change
0.80
2.17
1.45

Upregulation (2.40-fold) of *Irf10-v1* was observed only at 24HPI in fish held at 10°C compared to time and temperature matched PBS controls; no significant response was observed at 6HPI at either temperature. ASAL injected fish at 24HPI also had higher expression of *Irf10-v1* at 10°C than at 16°C (Figure 20), similar to the response observed for *Irf7* (Figure 18). Notably, *Irf10-v2* (the shorter *Irf10* splice variant) showed a significant increase in transcript expression response to ASAL at 6HPI at both temperatures (Figure 21B), unlike the longer *Irf10* splice variant which was non-responsive to ASAL at 6HPI. Significant upregulation of *Irf10-v2* was also seen in response to poly(I:C) injection compared with time- and temperature-matched PBS controls, at 6HPI for fish held at both temperatures (7.80-fold at 10°C and 10.76-fold at 16°C), and at 24HPI for fish held at 10°C (4.08-fold) (Figure 21A). Notably, the fold change values observed for *Irf10-v2* in response to poly(I:C) were the highest of any of the IRF family members included in this QPCR study. An effect of temperature on *Irf10-v2* transcript expression was observed in both ASAL and poly(I:C) injected fish, where expression at 6HPI was higher in 16°C fish and expression at 24HPI was higher in 10°C fish (Figure 21); this was similar to the effect of temperature on both *Irf4b* and *Irf7* transcript expression (Figures 17 and 18).

3.4 Developmental transcript expression analysis

Expression of cod IRF paralogues in embryos and larvae from 0 days post fertilization (dpf) to 17 dpf was studied using RT-PCR. Samples from three replicate incubators were observed under compound microscope each day to confirm that development was synchronous, and representative images were compiled (Figure 22). For

Figure 20: Spleen transcript expression response of *Irf10-v1* to bacterial antigens measured by QPCR. Data is presented as mean ± SEM, normalized to *ef1α* expression, with the lowest expressing sample set to RQ=1. Different letters represent significant differences between fish injected with ASAL at different temperatures within the same time point. Asterisks (*) represent significant differences between an ASAL injected group and the time- and temperature-matched PBS injected group (***p <0.001). Fold change is calculated as (mean ASAL RQ)/(mean PBS RQ).

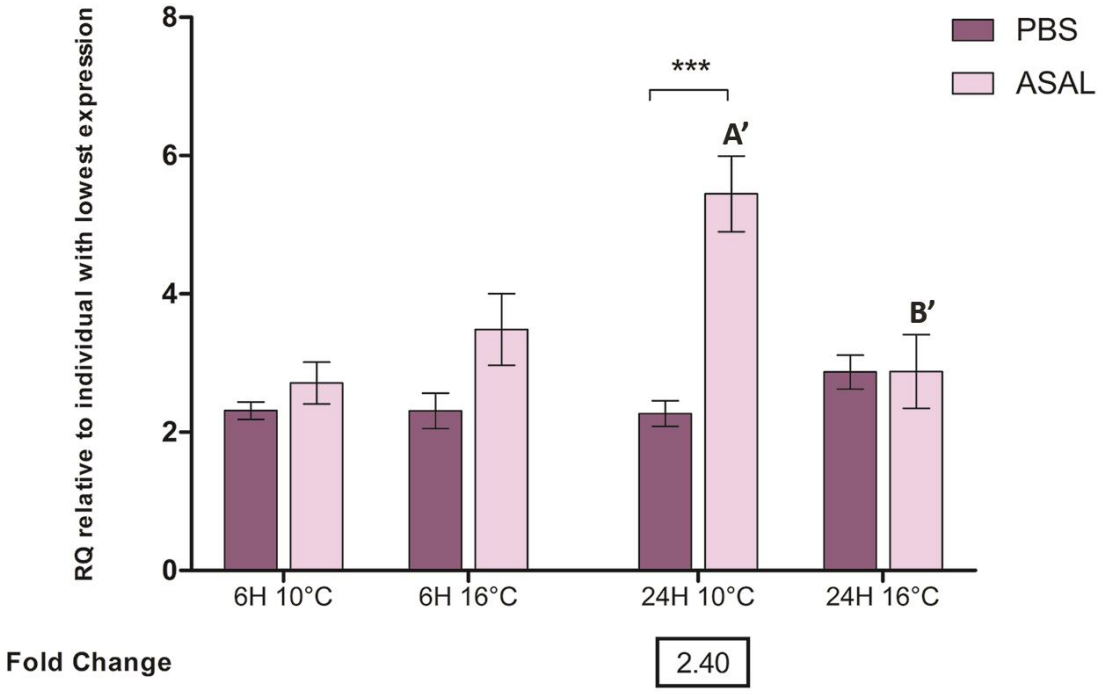
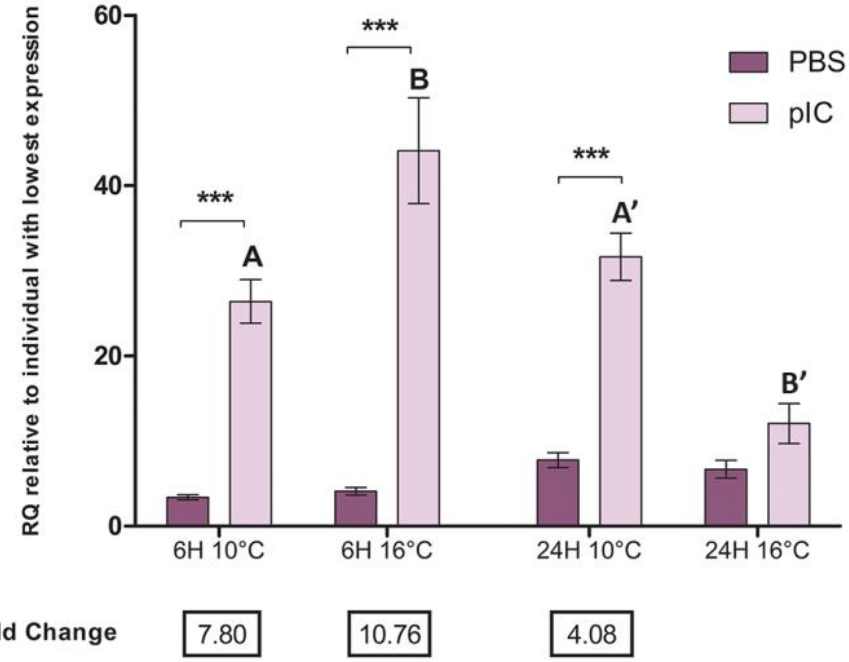

RQ relative to individual with lowest expression
8
6
4
2
0
PBS
ASAL
***
A'
B'
6H 10°C
6H 16°C
24H 10°C
24H 16°C
Fold Change
2.40

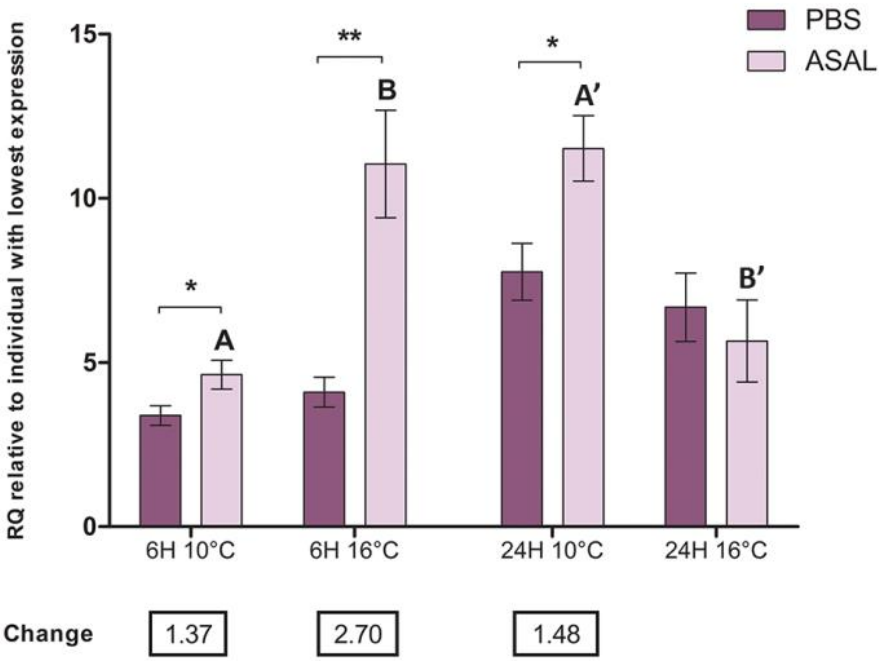
Figure 21: Spleen transcript expression responses of *Irf10-v2* to viral and bacterial antigens measured by QPCR. Data is presented as mean ± SEM, normalized to *ef1α* expression, with the lowest expressing sample set to RQ=1. Different letters represent significant differences between fish injected with poly(I:C) or ASAL at different temperatures within the same time point. Asterisks (*) represent significant differences between a poly(I:C) or ASAL injected group and the time- and temperature-matched PBS injected group (*p <0.05, ** p < 0.01, ***p < 0.001). Fold change is calculated as [mean poly(I:C) or ASAL RQ]/(mean PBS RQ). **A** = poly(I:C), **B** = ASAL. Note that poly(I:C) is abbreviated as pIC in the figure to conserve space.

A



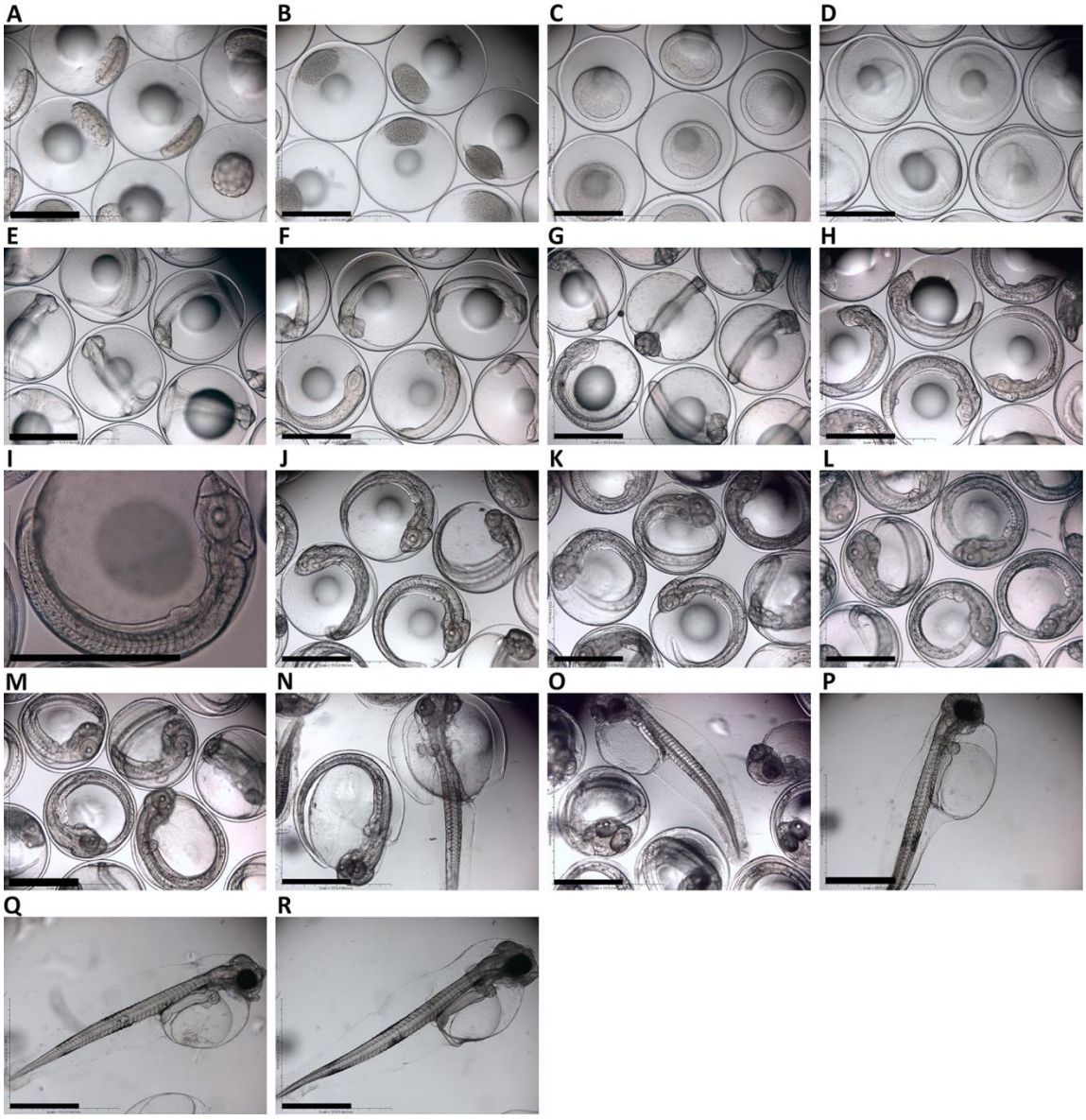
Fold Change	7.80	10.76	4.08

B



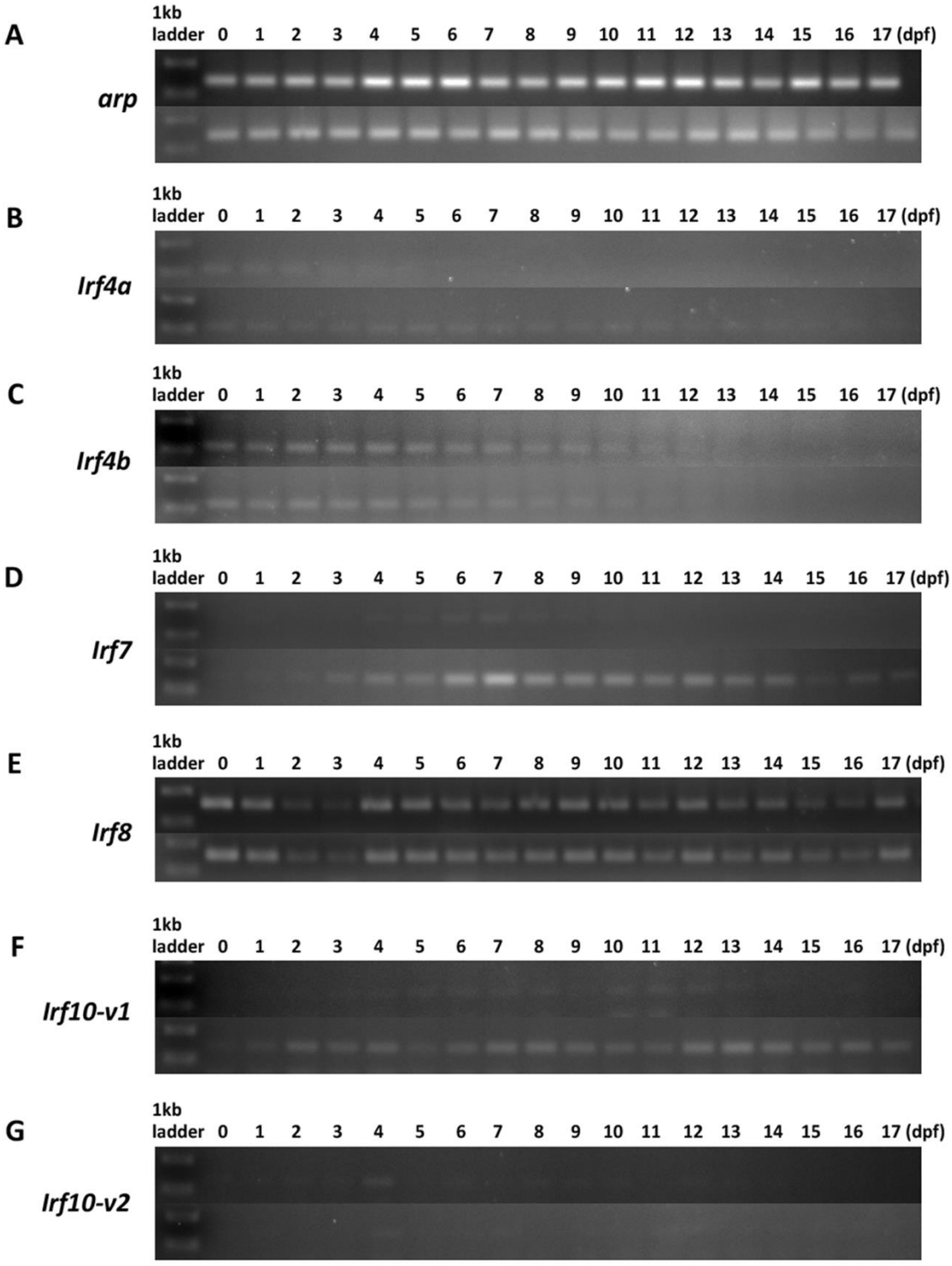
Fold Change	1.37	2.70	1.48

Figure 22: Representative images of Atlantic cod embryos and larvae sampled from 0 to 17 days post fertilization. Size bar = 1 mm. Embryos at 0 dpf (A) were observed to have some variation in stage, but most were at the 64 to 128 cell stage. Gastrulation was observed to be complete at 5 dpf (F). Hatching began at 13 dpf (N) and was complete at 15 dpf (P). Determination of developmental stages was based on descriptions by Hall *et al.*, (2004).


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R

RT-PCR analysis, samples from 2 of the 3 replicate incubators were used, and expression profiles between replicates were observed to be quite similar overall (Figure 23). In some cases however (i.e. *Irf7, Irf10-v1*), transcript expression in one replicate group appearedto be greater than the other (Figure 23D,F). Acidic ribosomal protein (arp) was chosen as a normalizer as the most stable of several potential normalizers tested, although some variation was still observed. As seen in the gel images, several cod IRF paralogues appear to have quite low transcript expression throughout embryonic development; this prevented analysis by QPCR in this study since acceptable standard curves were not produced in primer testing with these samples. Transcript expression profiles appear to be quite different among IRF paralogues. *Irf4a* and *Irf4b* (Figure 23A,B) transcripts appear to be most highly expressed in early embryonic stages and decrease over time, whereas *Irf10-v1* appears to have very little transcript expression during the first two days of development but remains relatively stable from 2 dpf to 17 dpf (Figure 23F). *Irf10-v2* appears to have little to no detectable transcript expression throughout most of the developmental stages included, with the exception of a visible band at 4 dpf (Figure 23G). *Irf7* and *Irf8* have interesting expression profiles based on this RT-PCR analysis, as *Irf7* expression (Figure 23D) appears to increase to a peak at 7 dpf and then decrease again (previously noted by Rise *et al*. (2012) based on QPCR analysis), and *Irf8* (Figure 23E) appears to have relatively high transcript expression at 0-1 dpf which drops drastically at 2-3 dpf and then increases again.

Figure 23: Composite agarose gel image of IRF family member transcript expression throughout Atlantic cod embryonic and early larval development. All gels are 1.7% agarose in TAE buffer, using 1 kb plus ladder (Invitrogen) as a size marker (100 bp and 200 bp bands are shown). PCR was carried out using samples from two replicate incubators/tanks (for each panel, "tank 1" = top and "tank 2" = bottom). Samples in each row are 0 dpf – 17 dpf from left to right.


A
1kb ladder 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 (dpf)
arp
B
1kb ladder 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 (dpf)
Irf4a
C
1kb ladder 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 (dpf)
Irf4b
D
1kb ladder 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 (dpf)
Irf7
E
1kb ladder 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 (dpf)
Irf8
F
1kb ladder 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 (dpf)
Irf10-v1
G
1kb ladder 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 (dpf)
Irf10-v2

4. Discussion

4.1 Overview

A better understanding of fish immune responses in general, and of the specific genes and molecular pathways involved in those responses, is of great value in furthering our knowledge of comparative vertebrate immunology and in improving aquaculture practices. The IRF gene family, which encodes transcription factors that are known to be important regulators of the vertebrate immune response to viral infection, have been studied in several fish species in recent years [e.g. *Irf*3 and *Irf7* in rainbow trout (Holland *et al.,* 2008), *Irf5* in grass carp (Xu *et al.,* 2010), *Irf1*, *Irf2*, *Irf3*, and *Irf7* in Atlantic salmon (Bergan *et al.,* 2010), *Irf1*, *Irf2*, and *Irf5* in paddlefish (Xiaoni *et al.,* 2011), and *Irf4* and *Irf8* in rock bream (Bathidge *et al.*, 2012); see Table 2 for summary], often with a focus on the transcript expression response to bacterial or viral stimulation. Since most Atlantic cod IRF family members had not previously been characterized or widely studied prior to the current study, the goal of this research was to fully characterize several cod IRFs at the cDNA level, to investigate how their transcript expression responds to immune stimulation, and to study expression in various tissues and developmental stages that may suggest potential roles of those genes and their encoded proteins.

4.2 mRNA characterization and phylogenetic analysis

In this study, complete cDNA sequences were obtained for Atlantic cod *Irf4a*, *Irf4b*, *Irf7*, *Irf8* and two *Irf10* splice variants, starting with partial cod IRF sequences from GenBank, and using RACE and other standard molecular techniques. The

identification of complete cDNA and predicted amino acid sequences of several cod IRFs allowed for molecular phylogenetic analysis to be conducted to study evolutionary relationships between these sequences and IRFs from other vertebrate species.

Two paralogous cod *Irf4* cDNA sequences were identified. The presence of additional IRF paralogues in a teleost species was not unexpected, as phylogenetic analysis of this gene family shows it has undergone expansion and diversification several times throughout vertebrate evolution (Nehyba *et al.,* 2002; 2009). Nehyba *et al.,* (2009) traced all IRF genes in humans to 4 of the 17 ancestral chordate linkage groups described by Putnam *et al.,* (2008), and noted that the 4 groups correspond to the 4 IRF subfamilies in vertebrates (see Figure 1). They concluded that the expansion from 4 to 10 IRF family members in most vertebrates likely resulted from the two rounds of whole genome duplication that are believed to have occurred in early vertebrate evolution. Interestingly, *Irf10*, present in chicken and teleost fish, appears to have been lost in humans and other mammals sometime after the second whole genome duplication event in the early vertebrate lineage (Nehyba *et al.,* 2009). Evidence suggests a third whole genome duplication occurred in the teleost fish lineage shortly after their divergence from lobe-finned fishes (Amores *et al.,* 1998), which could explain why some fish species show further expansion within the IRF family. For example, zebrafish has two *Irf1*-like genes (named *Irf1a* and *Irf1b* or *Irf1* and *Irf11* by different sources; Stein *et al.,* 2007; Huang *et al.,* 2010), and also has two *Irf4* paralogues, named *Irf4a* and *Irf4b* (Stein *et al.,* 2007). Stickleback (*Gasterosteus aculeatus*) is also predicted to have 2 *Irf1*-like and 2 *Irf4*-like genes (Huang *et al.,* 2010). Atlantic salmon has two *Irf7* paralogues (Bergan *et al.,* 2010),

although these likely arose after another putative whole genome duplication in the salmonid lineage (Allendorf and Thorgaard, 1984) which led to further expansion of many gene familes. As seen in Figure 14, phylogenetic analysis indicates the salmon IRF7 paralogues are more closely related to each other than to IRF7 protein sequences from other teleosts; however, zebrafish IRF4a and IRF4b are more closely related to rock bream IRF4 and cod IRF4b, respectively, than to each other. Therefore it is likely that the salmon IRF7 paralogues arose from duplication in the salmon lineage while the zebrafish IRF4 paralogues arose before the species diverged from the other teleosts included in this analysis.

Based on alignment with putative zebrafish orthologues, the shorter cod *Irf4* sequence identified in this study was named *Irf4a*, and the longer paralogue named *Irf4b*, being most similar to zebrafish *Irf4a* and *Irf4b*, respectively. Cod IRF4b does appear more closely related to zebrafish IRF4b than IRF4a in the phylogenetic tree depicted in Figure 14 (based on amino acid sequences), but appears to be most closely related to the Atlantic salmon and flounder IRF4 sequences. Cod IRF4a is shorter than the other amino acid sequences included in the analysis, which likely affected its placement on the phylogenetic tree on a separate branch from all of the other IRF4-like sequences. An alternate tree based on alignment of the same teleost IRF sequences trimmed to the length of cod IRF4a (144 AA) does show some differences from Figure 14 (particularly showing cod IRF4a and rock bream IRF4 sharing a branch and grouping separately from all other IRF4 sequences; see Appendix 9). The length of cod *Irf4a*, along with its lower expression compared to the other transcripts studied (below), suggests that a longer splice

variant of the *Irf4a* transcript exists but was not identified in the current study. Ensembl predicts a 954 bp cod *Irf4a* transcript (ENSGMOT00000005509), which is quite similar to the sequence obtained in the current research up to the end of exon 2.

Further studies to determine if a longer *Irf4a* splice variant exists in Atlantic cod would be of interest, as two different cod *Irf10* splice variants were identified in this study. It is therefore possible that alternate splicing may occur in other cod IRF family members as well. In humans, multiple splice variants of *Irf1* (Lee *et al.*, 2006), *Irf3* (Li *et al.*, 2011), *Irf5* (Graham *et al.*, 2006) and *Irf7* (Zhang and Pagano, 1998) have been identified, and several of these variants were found to have significant differences in function. For example, Lee *et al.* (2006) showed that alternative splicing of human *Irf1* negatively regulated wild type *Irf1* in cervical cancer tissue. They suggested that the more stable variant protein competes with the wild type IRF1 and decreases its functionality. Interestingly, although there are currently no studies about IRF splice variants in Atlantic cod, recent study of piscidins (a group of antimicrobial peptides) suggested that a splice variant of cod piscidin2 is produced by intron retention (Ruangsri *et al*., 2012), similar to *Irf10-v2* in the current study. The authors of that study suggested such a splice variant may regulate wild type expression through nonsense mediated decay. As IRFs and piscidins are both important to innate immune responses, future studies comparing expression and the roles of splice variants in the two groups in Atlantic cod would be interesting. Furthermore, as no evidence is present in the literature to indicate multiple splice variants of *Irf10* in any other species, the presence of differently expressed splice

variants in cod as indicated in the present study is of particular interest, as discussed below.

Phylogenetic analysis of predicted IRF amino acid sequences in cod along with those of other teleost species supported the division of IRFs into "IRF1-SG" and "IRF4-SG" supergroups, as described by Nehyba *et al.,* (2002), which can be distinguished by the presence of the IRF-association domain 1 (IAD1) in IRF4-SG (i.e., all IRFs except IRF1 and IRF2). The IAD, found in the middle to carboxyl region of the protein, is important for interaction with other IRF family members and other transcription factors (Meraro *et al.,* 1999). An IAD found in IRF1 and IRF2 (IAD2) was also identified by Meraro *et al.,* (1999); however, a consensus sequence for IAD2 was not found in the literature, and the domain is not listed in protein domain databases (e.g. NCBI, ExPASy).

All cod IRFs studied herein contain the amino terminal DNA binding domain (DBD) and associated conserved tryptophan residues found in all IRFs (Figure 13). While most mammalian IRFs contain five conserved Trp residues (Taniuchi *et al.,* 2001), there appears to be more variation in fish IRFs, with IRF1s having six and IRF7s having only four. As described above, the DBD binds specific enhancer-like elements in the promoters of type I IFNs or other target genes. The helix-loop-helix motif recognizes a sequence containing GAAA repeats and binds through three of the conserved tryptophan residues (Escalante *et al.,* 1998). The importance of this domain is highlighted by its high level of conservation among all IRFs in all species, even as evolution of the carboxyl terminal region has allowed this group of transcription factors to take on diverse roles in

biological processes such as development and oncogenesis (reviewed in Honda and Taniguchi, 2006; Ozato *et al.,* 2007; Savitsky *et al.,* 2010).

4.3 Expression analysis in juvenile cod tissues

To better understand the possible roles of IRFs in Atlantic cod, the constitutive expression of each transcript characterized above was investigated by RT-PCR in 15 different tissues of juvenile fish. As expected, the expression of all IRF transcripts was observed in spleen and hematopoietic (head or anterior) kidney, two important tissues in the teleost immune system. The spleen is a major site for the trapping and presentation of antigens for recognition by lymphocytes, and like the anterior kidney, is a site of hematopoiesis and the removal of aged or damaged blood cells (Zapata *et al.,* 1996). Both tissues are therefore of particular interest in immunological studies in teleosts. In the current study, all transcripts except *Irf4a* and *Irf4b* appeared to be expressed at some level in all of the included tissues (Figure 15). Studies of selected IRFs in rainbow trout (Holland *et al.,* 2010), yellow croaker (Yao *et al.,* 2010), turbot, Japanese flounder (Hu *et al.,* 2011a, b) and rock bream (Bathidge *et al.,* 2012) using QPCR have shown similar patterns of constitutive expression in most tissues with higher expression in spleen, head kidney and often gill and/or blood. The ubiquitous expression of cod *Irf7* and *Irf8* transcripts agrees with studies of those genes in other fish species [e.g. mandarin fish (Sun *et al.,* 2007), and Japanese flounder (Hu *et al.,* 2010; 2013)], where constitutive expression was seen in various tissues.

Irf4a appeared to have the lowest expression of all the transcripts included in the juvenile tissue panel RT-PCR study, and was only observed to be expressed in gill, head

kidney and spleen, with low expression in posterior kidney and blood in one replicate each (Figure 15B). IRF4 is known to be important to blood cell differentiation in human and mouse, particularly for dendritic cell development (Tamura *et al.*, 2005); therefore, it is not surprising that *Irf4a* appears to have higher transcript expression in hematopoietic tissues (i.e. spleen and kidney) than in most other tissues. Cod *Irf4b* appeared to be more widely expressed, although expression was low in several tissues (e.g. eye, posterior kidney, and stomach). Some discrepancy was observed however between the two biological replicates, particularly the replicate blood samples, for this transcript (Figure 15C). Constitutive transcript expression of cod *Irf4*-like genes agreed in general with previous studies in rainbow trout (Holland *et al.,* 2010) and rock bream (Bathige *et al.,* 2012) in which *Irf4* expression was relatively high in spleen and head kidney.

Importantly, different transcript expression profiles were observed for the two *Irf10* transcript variants in each of the expression studies carried out. The longer splice variant (named *Irf10-v1*) was observed to be constitutively expressed in all 15 tissues at a similar level overall. The shorter splice variant (*Irf10-v2*), however, appeared to have very low expression in eye and in most digestive tissues (stomach, midgut, and hindgut) and highest expression in the heart and skeletal muscle. It is therefore possible that the two splice variants have different functions, and as suggested above, that *Irf10-v2* may regulate *Irf10-v1* in some way. *Irf10* transcript expression has been investigated in very few other species. In chicken (*Gallus gallus*), this gene was observed to be most highly expressed in white blood cells, with relatively high transcript expression in spleen and thymus but little expression in other investigated tissues based on Northern blot analysis

(Nehyba *et al.*, 2002). In contrast, while both cod *Irf10* splice variants were expressed in hematopoietic tissues (spleen, head kidney) and blood, that expression was not observably higher than in other tissues. In Japanese flounder, *Irf10* mRNA was more widely expressed: in gill, heart, head and posterior (trunk) kidney, intestine and stomach (Suzuki *et al.*, 2011), which is comparable to the ubiquitous expression of *Irf10-v1* observed in the current study. Further studies using techniques such as *in situ* hybridization and immunohistochemistry should be carried out in the future to confirm differential constitutive expression of these cod *Irf10* splice variants, and suggest where (i.e. which tissues) and when (i.e. during different stages of development) each variant could function.

4.4 Spleen transcript expression response to immune stimulation

Previous to this study, transcript expression of cod *Irf7* and *Irf10* (*Irf10-v1*) had been observed to increase in spleen following intraperitoneal (IP) injection of the viral mimic poly(I:C) (Rise *et al.*, 2008; Hori *et al.*, 2012). Both transcripts, along with *Irf1*, had significantly higher transcript expression response to poly(I:C) at 16°C than 10°C at an earlier (6HPI) time point but a higher transcript expression response at 10°C than 16°C at a later (24HPI) time point (Hori *et al.*, 2012). However, neither the responsiveness of Atlantic cod *Irf4* or *Irf8* (or *Irf10-v2* which had not yet been identified) to poly(I:C), nor the transcript expression response to bacterial antigens of any transcript included in the current study had been previously investigated.

The response of IRF transcript expression to immune stimulation has been investigated in several other teleost species as described below, although to our

knowledge the effect of temperature on teleost IRF transcript expression response has only been investigated in our laboratory (Hori *et al.*, 2012, 2013) and in a zebrafish study which looked at the expression of *Irf3* along with several other antiviral genes (Dios *et al.*, 2010). An understanding of how changing temperatures may affect both the susceptibility of fish to infectious diseases and the function of immune responsive genes is of particular importance for Atlantic cod aquaculture, since cod that are confined to sea cages may be unable to move to an area of preferred temperature, and often experience seasonal temperature fluctuations (i.e. summer water temperatures of up to 20°C with short-term temperature fluctuations of up to 10°C; Gollock *et al.*, 2006). A primary goal of Hori *et al.* (2012, 2013) was therefore to determine if a gradual temperature increase (from 10°C to 16°C, 1°C every 5 days), comparable to that experienced by cod in the spring/summer Newfoundland climate, would modulate the anti-viral and anti-bacterial immune responses of cod and thereby potentially influence their susceptibility to infectious diseases. The current study uses the same temperature regime and samples as Hori *et al.*, (2012, 2013), but investigates the impact of elevated temperature and/or immune stimulation on the transcript expression of newly characterized IRF paralogues.

As in the constitutive tissue distribution study, differences were observed in the spleen transcript expression profiles of cod *Irf4a* and *Irf4b*, in response to both poly(I:C) and ASAL injection (Figures 16 and 17). For example, while *Irf4b* transcript expression increased in response to poly(I:C) (at 6HPI and 16°C; compared to time- and temperature-matched PBS control), *Irf4a* expression had no response at to poly(I:C) at 6HPI at either temperature or at 24HPI at 16°C, and was lower in poly(I:C) than PBS at

24HPI at 10°C. Interestingly, a similar transcript expression profile to cod *Irf4a* was observed for *Irf4* in rock bream injected with poly(I:C) (Bathige *et al.,* 2012). In that study, which included time points from 0HPI to 48HPI, the only significant response to poly(I:C) stimulation in spleen was a decrease at 12HPI. Based on the phylogenetic analysis (Figure 14), rock bream *Irf4* did appear to be more closely related to zebrafish *Irf4a* than to zebrafish or cod *Irf4b*, supporting its similar expression profile to cod *Irf4a* in response to poly(I:C). However, while ASAL stimulation was not included for the rock bream study, the effects of two other bacterial pathogens, *Edwardsiella tarda* and *Streptococcus iniae* were investigated, and both caused an initial decrease in *Irf4* expression at 3HPI, followed by an increase at 12HPI and then another decrease at the final (48HPI) time point, with similar expression profiles in spleen and head kidney (Bathige *et al.*, 2012). In cod, *Irf4b* was responsive to stimulation with ASAL while *Irf4a* was not, indicating that cod *Irf4b* may also share some similarity in function with the rock bream orthologue. Since cod *Irf4b* showed increased spleen transcript expression in response to ASAL at 6HPI (at both temperatures, compared to time- and temperature-matched PBS controls), but no response to ASAL at 24HPI at either temperature, it would be of interest to repeat this experiment using additional sampling time points from 3HPI to 48HPI to determine whether a similar pattern to that seen in rock bream *Irf4* following bacterial stimulation may occur.

Immune responsiveness of *Irf4* has also been studied in rainbow trout, where no response to poly(I:C) stimulation was observed in cultured splenocytes (Holland *et al.*, 2010). ASAL was again not used in that study, although stimulation with

lipopolysaccharide (LPS) produced a decrease in *Irf4* transcript expression. No data on *Irf4* transcript expression response to immune stimuli could be found for zebrafish or any other species with multiple *Irf4* paralogues, and therefore it is unknown whether the differing profiles observed in this study are unique to Atlantic cod. The very different transcript expression profiles of cod *Irf4a* and *Irf4b* (i.e. up-regulation of *Irf4b*, but not *Irf4a*, in response to both viral and bacterial antigens) provides evidence of regulatory divergence of these paralogues (i.e. gene duplication and divergence), even though they are quite similar over the length of the shorter *Irf4a* (74% identical overall and 81% identical over the DBD at the amino acid level, see Appendix 8). It also suggests the two genes may have different roles in immune responses to pathogens and/or pathogen-associated molecular patterns (PAMPs) such as poly(I:C).

Both the rainbow trout and rock bream studies discussed above investigated *Irf8* expression along with *Irf4*, as these two genes belong to the same sub-family (IRF4-G) and are more closely related to each other than to other IRFs, as indicated by phylogenetic analysis. In each species, up-regulation of *Irf8* transcript expression after poly(I:C) stimulation was observed, although in the current study the response was at 24HPI (at 10°C; compared to time- and temperature-matched PBS control) while in rock bream (Bathige *et al.*, 2012) the increase occurred at 3HPI, 12HPI and 24HPI time points (the trout study only included one sampling point at 4 hours post-stimulation). It should also be noted that in both the current study and the rock bream study, the increased *Irf8* transcript expression was quite subtle, indicated as fold changes of 2 or less compared to time matched PBS controls (Figure 19A; Bathige *et al.*, 2012). A 5-fold increase in *Irf8*

expression was observed in response to poly(I:C) in trout, although this study included cultured splenocytes rather than whole spleen tissue (Holland *et al.*, 2010). Responsiveness of *Irf8* to poly(I:C) has also been observed in the spleen of turbot (Chen *et al.*, 2012) and Japanese flounder (Hu *et al.*, 2013). In turbot, *Irf8* transcript expression was increased at 12HPI but not at 24HPI or 48HPI, while in flounder *Irf8* transcript expression peaked at 3HPI. Thus, the timing of the immune response may be different in each species, although differences in poly(I:C) dosage, fish age and/or size, and other factors must be considered.

The cod *Irf8* response to ASAL (at 10°C) appears to follow a similar pattern to the rock bream *Irf8* (and *Irf4*) response to bacterial pathogens: in both cases there is an initial decrease in transcript expression and then an increase compared to PBS controls. However, the transcript expression profile at 16°C for cod *Irf8* was quite different, showing an increase at 6HPI and no significant difference at 24HPI in response to ASAL compared to PBS controls (Figure 19B); unfortunately no other studies of *Irf8* transcript expression include multiple temperatures for comparison. As noted above, our study did not include a 48HPI time point, and therefore it is unknown whether a later decrease in expression may occur in Atlantic cod *Irf8* at either temperature. Bathige *et al.*, (2012) suggested the initial decrease observed in their study may have been caused by the immune suppressive capability of live pathogens; however, this explanation would not apply to killed pathogens (i.e. ASAL) as used in the current study. Interestingly, while Hori *et al.*, (2013) found the effect of temperature increase on overall immune-relevant transcript expression to be much greater in poly(I:C) vs. ASAL stimulated cod, the

greatest response of *Irf8* to ASAL stimulation (a 2.17 fold increase; Figure 19B) was observed at 6HPI at the elevated temperature in the current study, while no significant response was observed at the elevated temperature in poly(I:C) injected fish.

Atlantic cod *Irf7* transcript expression has been shown to increase in response to poly(I:C) exposure at 6HPI and 24HPI time points (Rise *et al*, 2008), with a greater response at 16°C at the earlier time-point and a greater response at 10°C at the later time-point (Hori *et al.*, 2012). *Irf7* has also been observed to be poly(I:C) responsive in head kidney and gill in Japanese flounder (Hu *et al.*, 2010), in rainbow trout cell lines (Holland *et al.*, 2008), and in liver and head kidney of Atlantic salmon (Kileng *et al.*, 2009), although spleen expression was not studied in these species. In the mandarin fish, spleen transcript expression of *Irf7* was studied and found to increase with poly(I:C) stimulation, peaking at 12HPI, with similar responses in gill and liver (Sun *et al.*, 2007). Response to ASAL was not investigated in any of these species, although a different study in the orange-spotted grouper showed that *Irf7* expression in spleen increased in response to injection with the bacterium *Vibrio vulnificus* (Cui *et al.*, 2011). In Atlantic cod, increased *Irf7* transcript expression in the brain (based on microarray data) has been observed in response to injection with nervous necrosis virus, and QPCR analysis showed a response to poly(I:C) in cod cell lines (Krasnov *et al*., 2012). In the current study, an increase in *Irf7* transcript expression in response to ASAL injection (at 16°C for the 6HPI time-point and at 10°C for the 24HPI time-point) was observed, indicating that this gene (along with all other genes in this study except *Irf4b*) likely plays a role in the immune response to both viral and bacterial infection in this species. The temperature-dependant

expression profile of *Irf7* observed in response to ASAL injection is similar to that observed in response to poly(I:C) by Hori et al (2012) for *Irf7* and several other immune-relevant cod transcripts (i.e. earlier response at elevated temperature). The results of the current study build on those of Hori *et al.,* (2012, 2013) by showing that a moderate temperature increase also modulates the cod spleen transcript expression response of multiple IRF genes (*Irf7*, *Irf8* and both *Irf10* splice variants) to bacterial antigens.

The response of *Irf10-v1* to poly(I:C) was also investigated by Hori *et al.,* (2012), where (as with *Irf7*) the increase in transcript expression was greater at 16°C for the 6HPI time point and at 10°C for the 24HPI time point. This transcript was shown to be responsive to ASAL injection as well in the current study, although expression only increased (compared to the time- and temperature-matched PBS control) at 10°C, and only at the 24HPI time-point (Figure 20). Interestingly, the second *Irf10* splice variant (*Irf10-v2*) showed different expression profiles from *Irf10-v1* in response to both poly(I:C) and ASAL. While the greatest response to poly(I:C) was observed at 24HPI and 10°C for *Irf10-v1* (9-fold increase; Hori *et al.,* 2012), the responses of *Irf10-v2* at 6HPI were both greater (~8-fold increase at 10°C and ~11-fold at 16°C) than the response at 24HPI (4-fold at 10°C, with no response at 16°C; Figure 21A). Increases in *Irf10-v2* expression were observed in ASAL injected fish at both 6HPI and 24HPI at 10°C, although the increase was greater in the later time point (Figure 21B), consistent with the common profile (i.e. later responses at the lower temperature) observed by Hori *et al.,* (2012, 2013). The study of *Irf10* expression response to immune stimulation in this experiment indicated that while the two splice variants of this gene in Atlantic cod are

responsive to both poly(I:C) and ASAL stimulation, there are observable differences in the timing and intensity of those responses. Along with the tissue distribution data above, this suggests that the two splice variants may have distinct roles in the immune response, which will be an area of particular interest for further study. Very little study of *Irf10* in other species has been carried out to date, and therefore it is unknown whether the presence of such splice variants is unique to Atlantic cod.

4.5 Developmental transcript expression analysis

Since IRF family members are known in several species to function in the development of innate and adaptive immunity (reviewed in Ozato *et al.*, 2007), and because cod *Irf1* and *Irf7* have previously been shown to be maternal transcripts with dynamic expression profiles during embryonic development (Rise *et al.,* 2012), the expression of all IRF transcripts included in the current study throughout early development was also investigated. Although QPCR studies were not completed using these samples, RT-PCR did indicate several expression profiles that will be of interest for further study; notably, *Irf7* expression was similar to that seen by Rise *et al.,* (2012) using QPCR, with an apparent peak in early segmentation [6 dpf in the previous study; 7 dpf in the current study (Figure 23D)]. This indicates a possible important role for IRF7 [and IRF1, as hypothesized by Rise *et al.,* (2012)] in this stage of development, which could be investigated further in the future (e.g. using morpholino injection for gene knockdown). Very little information is found in the literature about the role of the IRF7 transcription factor in development, although one study indicates it is required for the development of medullary thymic epithelial cells in mice (Otero *et al.,* 2013).

Investigation of the role of IRF7 in early embryonic development in cod and other teleosts (e.g. Atlantic salmon or zebrafish) as well as in other vertebrate species will therefore be of particular interest in ongoing research.

Atlantic cod *Irf4a* and *Irf4b* transcript expression levels appear to decrease throughout embryonic development (Figure 23B,C), suggesting both may be maternal transcripts (present in the unfertilized egg), and possibly have an important role in the early embryo. Future QPCR studies could include unfertilized egg in addition to embryonic/larval stages to further investigate this possibility. *Irf8* appears to have its highest transcript expression at 0 dpf as well, although this transcript has a unique expression profile; it is expressed throughout the developmental stages included in the current study, from 0 dpf to 17 dpf, but appears to drop suddenly at 2-3 dpf before increasing again at 4 dpf. As noted above (Table 1), both IRF4 and IRF8 are known to be important to the differentiation of different cell types in mammals. For example, mice deficient in IRF4, which in mammals is only expressed in lymphoid and myeloid cells, show impaired activation and differentiation of B and T cells (Mittrucker *et al*., 1997); and the transcription factor is required for B cells to undergo isotype switching and differentiation into plasma cells (Sciammas *et al*., 2006). IRF8 has been shown to be required for the differentiation of myeloid progenitor cells into macrophages as opposed to granulocytes, with IRF8 knockout mice developing immunodeficiency (Tamura and Ozato, 2002). In another study, IRF8-deficient mice were shown to have increased numbers of microglia with altered morphology compared to wild type mice, indicating the transcription factor has an important role in the development of those cells in the

brain (Minten *et al.,* 2012). The role of these genes in immune system development in fish is less well studied, although IRF8 has been shown to regulate the differentiation of myeloid cells during zebrafish development, as knockdown of its expression produced embryos with depleted macrophage but expanded neutrophil populations (Li *et al.,* 2011). Investigation into the role of each of these genes during Atlantic cod development using knockdown studies will be of interest for further research.

As seen with the immune stimulation QPCR studies above, the two cod *Irf10* splice variants again appear to have different transcript expression profiles in the developmental series RT-PCR study. *Irf10-v2* showed only very faint expression throughout development, with a peak at 4 dpf, possibly indicating a role in late gastrulation, while *Irf10-v1* expression appeared to increase with time (Figure 23F,G). However, because some discrepancy is visible between replicates (particularly in *Irf10-v1*), further studies (i.e. using QPCR) will be necessary to confirm all developmental expression profiles.

4.6 Conclusions

The main objectives of this research were to characterize multiple Atlantic cod IRF family members at the cDNA and putative amino acid levels; to investigate the constitutive expression of those transcripts; and to expand on the findings of earlier studies in our laboratory (Hori *et al.,* 2012, 2013) about the effect of temperature on the immune response to viral and bacterial antigens. Six Atlantic cod IRF transcripts were characterized, including a novel *Irf10* splice variant, and the *Irf10* genomic region was sequenced. RT-PCR analysis showed that all of these transcripts were expressed in

spleen, head kidney and gill, and most were ubiquitously expressed in the tissues studied. The second RT-PCR study indicated that different IRF transcripts have unique developmental expression profiles and that some IRFs (e.g. *Irf7*, *Irf10-v2*) may have an important function at specific stages of development.

QPCR analysis of spleen expression confirmed that all transcripts were responsive to poly(I:C) and all except *Irf4a* were responsive to ASAL stimulation; and the effect of increased temperature previously observed (leading to an earlier transcriptional response to immune stimulation; Hori *et al.*, 2012, 2013) was seen in several cases. As noted by Hori *et al.*, (2012), these findings indicate that while increased summer temperatures in themselves may not be lethal for Atlantic cod, the effect of such temperatures on immune responses will be of particular importance to future Atlantic cod aquaculture.

4.7 Future Research

This study adds to our knowledge of molecular immunology in fish and of the IRF gene family, and provides many avenues for further investigation. For example, further sequencing at the genomic DNA level is of interest for each of these genes, both to confirm the placement of introns and to further characterize the 5' upstream regions as the 5'UTR obtained using RACE techniques were as short as 36 bp. Analysis of the upstream regions would aid in our understanding of how IRF expression is regulated, and how IRF family members interact with each other, with other transcription factors, and with IFN. In particular, it would be interesting to sequence and analyze the proximal promoters of cod *Irf4a* and *Irf4b* to determine if there are differences in regulatory sequences (e.g. putative transcription factor binding sites) that may explain the

differences in regulation of these paralogues observed above. It would also be valuable to use QPCR to study the transcript expression of both *Irf4* paralogues and *Irf8* in blood cells and hematopoietic kidney, since these genes are known to be important to hematopoiesis in other species. It will be important to investigate transcript expression in different classes of cod leukocytes [e.g. using Fluorescence Activated Cell Sorting (FACS)] to determine if similar functions are carried out during hematopoiesis in cod. As QPCR studies of cod IRF family members during embryonic and larval development were not successful in this study due to low levels of transcript expression, future research will utilize modified QPCR methods (e.g. use of amplified RNA) which may allow this experiment to be completed. Furthermore, techniques such as *in situ* hybridization, immunohistochemistry and gene knockdown by morpholino injection will be used in ongoing research following from this study to better understand the roles of IRF genes, particularly in development.

Literature Cited

Allendorf FW and Thorgaard GH. 1984. Tetraploidy and the evolution of salmonid fishes. Pp. 1-53 *In*: *Evolutionary genetics of fishes* (ed. B.J. Turner). Plenum Press, New York, USA.

Altmann SM, Mellon MT, Distel DL, Kim CH. 2003. Molecular and functional analysis of an interferon gene from the zebrafish, *Danio rerio*. J Virol 77:1992-2002.

Amores A, Force A, Yan Y, Joly L, Amemiya C, Fritz A, Ho RK, Langeland J, Prince V, Wang Y, Westerfield M, Ekker M, Postlethwait JH. 1998. Zebrafish *hox* clusters and vertebrate genome evolution. Science 282:1711-14.

Au WC, Moore PA, Lowther W, Juang YT, Pitha PM. 1995. Identification of a member of the interferon regulatory factor family that binds to the interferon-stimulated response element and activates expression of interferon-induced genes. Proc Natl Acad Sci USA 92:11657-61.

Barnes B, Lubyova B, Pitha PM. 2002. On the role of IRF in host defense. J Interferon Cytokine Res 22:59-71.

Bathige SDNK, Whang I, Umasuthan N, Lim BS, Park MA, Kim E, Park HC, Lee J. 2012. Interferon regulatory factors 4 and 8 in rock bream, *Oplegnathus fasciatus*: Structural and expressional evidence for their antimicrobial role in teleosts. Fish Shellfish Immunol 33:857-871.

Ben J, Wang Jabs E, and Chong SS. 2005. Genomic, cDNA and embryonic expression analysis of zebrafish IRF6, the gene mutated in the human oral clefting disorders Van der Woude and popliteal pterygium syndromes. Gene Expression Patterns 5:629-638.

Bergan V, Kileng Ø, Sun B, Robertsen B. 2010. Regulation and function of interferon regulatory factors of Atlantic salmon. Mol Immunol 47:2005-14.

Booman M, Borza T, Feng CY, Hori TS, Higgins B, Culf A, Léger D, Chute IC, Belkaid A, Rise M, Gamperl AK, Hubert S, Kimball J, Ouellette RJ, Johnson SC, Bowman S, Rise ML. 2011. Development and experimental validation of a 20K Atlantic cod (*Gadus morhua*) oligonucleotide microarray based on a collection of over 150,000 ESTs. Mar Biotechnol 13:733-50.

Booman M and Rise ML. 2012. Genomic tools for understanding the molecular basis of production-relevant traits in finfish. In: *Aquaculture Biotechnology* (eds. G.L. Fletcher and M.L. Rise), Wiley-Blackwell, Ames, IA. pp. 3-20.

Bowman S, Hubert S, Higgins B, Stone C, Kimball J., Borza T, Bussey J, Simpson G, Kozera C, Curtis BA, Hall JR, Hori TS, Feng CY, Rise M, Booman M, Gamperl AK, Trippel E, Symonds J, Johnson SC, and Rise ML. 2011. An integrated approach to gene discovery and marker development in Atlantic cod (*Gadus morhua*). Mar Biotechnol 13:242-255.

Brown JA, Minkoff G, Puvanendran V. 2003. Larviculture of Atlantic cod (*Gadus morhua*): progress, protocols and problems. Aquaculture 227:357-372.

Cham CM, Ko K, Niewold TB. 2012. Interferon regulatory factor 5 in the pathogenesis of systemic lupus erythematosus. Clinical Dev Immunol 2012:780436.

Chen X, Hu G, Dong X, Liu Q, Zhang S. 2012. Molecular cloning and expression analysis of interferon regulatory factor 8 (IRF8) in turbot, Scophthalmus maximus. Vet Immunol Immunopathol 149:143-50.

Colgan DF and Manley JL. 1997. Mechanism and regulation of mRNA polyadenylation. Genes Dev 11:2755-66.

Cui H, Yan Y, Wei J, Huang X, Huang Y, Ouyang Z, Qin Q. 2011. Identification and functional characterization of an interferon regulatory factor 7-like (IRF7-like) gene from orange-spotted grouper, *Epinephelus coioides*. Dev Comp Immunol 35:672-84.

Darnell JE, Kerr IM, Stark GR. 1994. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. Science 264:1415-21.

Dios S, Romero A, Chamorro R, Figueras A, Novoa B. 2010. Effect of the temperature during antiviral immune response ontogeny in teleosts. Fish Shellfish Immunol 29:1019-27.

Di Paolo NC, Doronin K, Baldwin LK, Papayannopoulou T, Shayakhmetov DM. 2013. The transcription factor IRF3 triggers "defensive suicide" necrosis in response to viral and bacterial pathogens. Cell Reports 3:1840-46.

Escalante CR, Yie J, Thanos D, Aggarwal AK. 1998. Structure of IRF-1 with bound DNA reveals determinants of interferon regulation. Nature 391:103-6.

Feng CY, Johnson SC, Hori TS, Rise M, Hall JR, Gamperl AK. 2009. Identification and analysis of differentially expressed genes in immune tissues of Atlantic cod stimulated with formalin-killed, atypical *Aeromonas salmonicida*. Physiol Genomics 37:149-63.

Gabriele L and Ozato K. 2007. The role of the interferon regulatory (IRF) family in dendritic cell development and function. Cytokine Growth Factor Rev 18:503-10.

Gollock MJ, Currie S, Petersen LH, Gamperl AK. 2006. Cardiovascular and haematological responses of Atlantic cod (*Gadus morhua*) to acute temperature increase. J Exp Biol 209:2961-70.

Graham RR, Kyogoku C, Sigurdsson S, Vlasova IA, Davies LRL, Baechler EC, Plenge RM, Koeuth T, Ortmann WA, Hom G, *et al.*, 2007. Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. Proc Natl Acad Sci USA 104:6758-63.

Hall TE, Smith P, and Johnson IA. 2004. Stages of embryonic development in the Atlantic cod *Gadus morhua*. J Morphol 259:255-0.

Holland JW, Bird S, Williamson B, Woudstra C, Mustafa A, Wang T, Zou J, Blaney SC, Collet B, Secombes CJ. 2008. Molecular characterization of IRF3 and IRF7 in rainbow trout, *Oncorhynchus mykiss*: Functional analysis and transcriptional modulation. Mol Immunol 46:269-85.

Honda K, Yanai H, Negishi H, Asagiri M, Sato M, Mizutani T, Shimada N, OhbaY, Takaoka A, Yoshida N, Taniguchi T. 2005. IRF-7 is the master regulator of type-I interferon-dependent immune responses. Nature 434:772-77.

Honda K and Taniguchi T. 2006. IRFs: Master regulators of signalling by Toll-like receptors and cytosolic pattern-recognition receptors. Nat Rev Immunol 6:644-58.

Hori TS, Gamperl AK, Booman M, Nash GW, Rise ML. 2012. A moderate increase in ambient temperature modulates the Atlantic cod (*Gadus morhua*) spleen transcriptome response to intraperitoneal viral mimic injection. BMC Genomics 13:431.

Hori TS, Gamperl AK, Nash G, Booman M, Barat A, Rise ML. 2013. The impact of a moderate chronic temperature increase on spleen immune-relevant gene transcription depends on whether Atlantic cod (*Gadus morhua*) are stimulated with bacterial versus viral antigens. Genome 56:567-76.

Hu G, Yin X, Xia J, Dong X, Zhang J, Liu Q. 2010. Molecular cloning and characterization of interferon regulatory factor 7 (IRF-7) in Japanese flounder, *Paralichthys olivaceus*. Fish Shellfish Immunol 29:963-71.

Hu G, Xia J, Lou H, Chen X, Li J, Liu Q. 2011a. An IRF-3 homolog that is up-regulated by DNA virus and poly I:C in turbot, *Scophthalmus maximus.* Fish Shellfish Immunol 31:1224-31.

Hu G, Yin X, Lou H, Xia J, Dong X, Zhang J, Liu Q. 2011b. Interferon regulatory factor

3 (IRF-3) in Japanese flounder, *Paralichthys olivaceus*: Sequencing, limited tissue distribution, inducible expression and induction of fish type I interferon promoter. Dev Comp Immunol 35:164-73.

Hu G, Xia J, Lou Q, Lin J, Yin X, Dong X. 2011c. Cloning and expression analysis of interferon regulatory factor 7 (IRF-7) in turbot, *Scophthalamus maximus.* Dev Comp Immunol 35:416-20.

Hu G, Chen X, Gong Q, Liu Q, Zhang S, Dong X. 2013. Structural and expression studies of interferon regulatory factor 8 in Japanese flounder, *Paralichthys olivaceus.* Fish Shellfish Immunol 35:1016-24.

Huang B, Qi Z, Xu Z, Nie P. 2010. Global characterization of interferon regulatory factor (IRF) genes in vertebrates: Glimpse of the diversification in evolution. BMC Immunol 11:22.

Kileng O, Bergan V, Workenhe ST, Robertsen B. 2009. Structural and functional studies of an IRF-7-like gene from Atlantic salmon. Dev Comp Immunol 33:18-27.

Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, Howard E, Ferreira de Lima RL, Daack-Hirsch S, *et al.*, 2002. Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. Nat Genet 32:285-89.

Krasnov A, Kileng O, Skugor S, Jorgensen SM, Afanasyev S, Timmerhaus G, Sommer AI, Jensen I. 2013. Genomic analysis of the host response to nervous necrosis virus in Atlantic cod (*Gadus morhua*) brain. Mol Immunol 54:443-52.

Lanes CFC, Fernandes JMO, Kiron V, Babiak I. 2012. Profiling key apoptotic, stress, and immune-related transcripts during embryonic and postembryonic development of Atlantic cod (*Gadus morhua* L.). Theriogenology 78:1583-96.

Lang AS, Rise ML, Culley AI, Steward GF. 2009. RNA viruses in the sea. FEMS Microbiol Rev 33:295-323.

Lee E, Jo M, Park J, Zhang W, Lee J. 2006. Alternative splicing variants of IRF-1 lacking exons 7, 8, and 9 in cervical cancer. Biochem Biophys Res Comm 347:882-88.

Li L, Jin H, Xu J, Shi Y, Wen Z. 2011. Irf8 regulates macrophage versus neutrophil fate during zebrafish primitive myelopoiesis. Blood 117:1359-69.

Li Y, Hu X, Song Y, Lu Z, Ning T, Cai H, Ke Y. 2011. Identification of novel alternative splicing variants of interferon regulatory factor 3. Biochem Biophys Acta 1809:166-75.

Lien C, Fang CM, Huso D, Livak F, Lu R, Pitha PM. 2010. Critical role of IRF-5 in regulation of B-cell differentiation. Proc Natl Acad Sci USA 107:4664-68.

Lutfalla G, Hugues RC, Stange-Thomann N, Jaillon O, Mogensen K, Monneron D. 2003. Comparative genomic analysis reveals independent expansion of a lineage-specific gene family in vertebrates: The class II cytokine receptors and their ligands in mammals and fish. BMC Genomics 4:29.

MacDonald CC, and Redondo JL. 2002. Reexamining the polyadenylation signal: were we wrong about AAUAAA? Mol Cell Endocrinol 190:1-8.

Marteinsdottir G, Ruzzante D, Nielsen EE. 2005. History of the North Atlantic cod stocks. ICES CM AA:19.

Matsuyama T, Grossman A, Mittrucker HW, Siderovski DP, Kiefer F, Kawakami T, Richardson CD, Taniguchi T, Yoshinaga SK, Mak TW. 1995. Molecular cloning of LSIRF, a lymphoid-specific member of the interferon regulatory factor family that binds the interferon-stimulated response element (ISRE). Nucleic Acids Res 23:2127-36.

Meraro D, Hashmueli S, Koren B, Azriel A, Oumard A, Kirchhoff S, Hauser H, Nagulapalli S, Atchison ML, Levi B. 1999. Protein-protein and DNA-protein interactions affect the activity of lymphoid-specific IFN regulatory factors. J Immunol 163:6468-78.

Minten C, Terry R, Deffrasnes C, King NJC, Campbell IL. 2012. IFN regulatory factor 8 is a key constitutive determinant of the morphological and molecular properties of microglia in the CNS. PLoS One 7:e49851.

Mittrucker HW, Matsuyama T, Grossman A, Kundig TM, Potter J, Shahinian A, Wakeham A, Patterson B, Ohashi PS, Mak TW. 1997. Requirement for the transcription factor LSIRF/IRF4 for mature B and T lymphocyte function. Science 275:540-43.

Negishi H, Ohba Y, Yanai H, Takaoka A, Honma K, Yui K, Matsuyama T, Taniguchi T, Honda K. 2005. Negative regulation of toll-like-receptor signaling by IRF-4. Proc Natl Acad Sci USA 102:15989-94.

Nehyba J, Hrdlickova R, Burnside J, Bose HR. 2002. A novel interferon regulatory factor (IRF), IRF-10, has a unique role in immune defense and is induced by the v-Rel oncoprotein. Mol Cell Biol 22:3942-57.

Nehyba J, Hrdlickova R, Bose HR. 2009. Dynamic evolution of immune system regulators: The history of the interferon regulatory factor family. Mol Biol Evol 26:2539-50.

Nelson N, Kanno Y, Hong C, Contursi C, Fujita T, Fowlkes BJ, O'Connell E, Hu-Li J, Paul WE, Jankovic D, Sher AF, Coligan JE, Thornton A, Appella E, Yang Y, Ozato K. 1996. Expression of IFN regulatory factor family proteins in lymphocytes. Induction of Stat-1 and IFN consensus sequence binding protein expression by T cell activation. J Immunol 156:3711-20.

Otero DC, Baker DP, David M. 2013. IRF7-dependent IFN-B production in response to RANKL promotes medullary thymic epithelial cell development. J Immunol 190:3289-98.

Ozato K, Tailor P, Kubota T. 2007. The interferon regulatory factor family in host defense: mechanism of action. J Biol Chem 282:20065-69.

Perez-Casanova JC, Rise ML, Dixon B, Alfonso LOB, Hall JR, Johnson SC,, Gamperl AK. 2008. The immune and stress response of Atlantic cod to long-term increases in water temperature. Fish Shellfish Immunol 24:600-09.

Putnam NH, Butts T, Ferrier DEK, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu J, *et al.*, 2008. The amphioxus genome and the evolution of the chordate karyotype. Nature 453:1064-71.

Restivo G, Nguyen B, Dziunycz P, Ristorcelli E, Ryan RJH, Ozuysal OY, Di Piazza M, Radtke F, *et al.*, 2011. IRF6 is a mediator of Notch pro-differentiation and tumor suppressive function in keratinocytes. EMBO Journal 30:4571-85.

Rise ML, Hall J, Rise M, Hori T, Gamperl AK, Kimball J, Hubert S, Bowman S, and Johnson SC. 2008. Functional genomic analysis of the response of Atlantic cod (*Gadus morhua*) spleen to the viral mimic polyriboinosinic polyribocytidylic acid (pIC). Dev Comp Immunol 32:916-31.

Rise ML, Hall JR, Alcock BP, Hori TS. 2012. Dynamic expression profiles of virus-responsive and putative antimicrobial peptide-encoding transcripts during Atlantic cod (*Gadus morhua*) embryonic and early larval development. Gene 509:232-246.

Robertsen B. 2006. The interferon system of teleost fish. Fish Shellfish Immunol 20:172-91.

Rosenlund G, and Hallorsson O. 2007. Cod juvenile production: Research and commercial developments. Aquaculture 268:188-194.

Ruangsri J, Salger SA, Caipang CMA, Kiron V, Fernandes JMO. 2012. Differential expression and biological activity of two piscidin paralogues and a novel splice variant in Atlantic cod (*Gadus morhua* L.). Fish Shellfish Immunol 32:396-406.

Sabel JL, d'Alençon C, O'Brien EK, Otterloo EV, Lutz K, Cuykendall TN, Schutte BC, Houston DW, Cornell RA. 2009. Maternal interferon regulatory factor 6 is required for the differentiation of primary superficial epithelia in Danio and Xenopus embryos. Dev Biol 325:249-62.

Savitsky D, Yanai H, Taniguchi T. 2010. Regulation of immunity and oncogenesis by the IRF transcription factor family. Cancer Immunol Immunother 59:489-510.

Sciammas R, Shaffer AL, Schatz JH, Zhao H, Staudt LM, Singh H. 2006. Graded expression of interferon regulatory factor-4 coordinates isotype switching with plasma cell differentiation. Immunity 25:225-36.

Shi J, Zhang Y, Liu T, Sun F, Gui J. 2012. Subcellular localization and functional characterization of a fish IRF9 from crucian carp *Carassius auratus*. Fish Shellfish Immunol 33:258-66.

Solem ST, and Stenvik J. 2006. Antibody repertoire development in teleosts – a review with emphasis on salmonids and *Gadus morhua L*. Dev Comp Immunol 30:57-76.

Star B, Nederbragt AJ, Jentoft S, Grimholt U, Malmstrom M, Gregers TF, Rounge TB, Paulsen J, *et al.*, The genome sequence of Atlantic cod reveals a unique immune system. Nature 477:207-10.

Stein C, Caccamo M, Laird G, Leptin M. 2007. Conservation and divergence of gene families encoding components of innate immune response systems in zebrafish. Genome Biol 8:R251.

Sun BJ, Chang MX, Song Y, Yao WJ, Nie P. 2007. Gene structure and transcription of IRF-1 and IRF-7 in the mandarin fish *Siniperca chuatsi*. Vet Immunol Immunopathol 116:26-36.

Sun F, Zhang Y, Liu T, Gan L, Yu F, Liu Y, Gui J. 2010. Characterization of fish IRF3 as an IFN-inducible protein reveals evolving regulation of IFN response in vertebrates. J Immunol 185:7573-82.

Suzuki Y, Yasuike M, Kondo H, Aoki T, Hirono I. 2011. Molecular cloning and expression analysis of interferon regulatory 10 (IRF10) in Japanese flounder, *Paralichthys olivaceus*. Fish Shellfish Immunol 30:67-76.

Takaoka A, Yanai H, Kondo S, Duncan G, Negishi H, Mizutani T, Kano S, Honda K, Ohba Y, Mak TW, Taniguchi T. 2005. Integral role of IRF-5 in the gene induction programme activated by Toll-like receptors. Nature 434:243-49.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731-39.

Tamura T, and Ozato K. 2002. ICSBP/IRF-8: its regulatory roles in the development of myeloid cells. J Interferon Cytokine Res 22:145-52.

Tamura T, Tailor P, Yamaoka K, Kong HJ, Tsujimura H, O'Shea JJ, Singh H, Ozato K. 2005. IFN regulatory factor-4 and -8 govern dendritic cell subset development and their functional diversity. J Immunol 174:2573-81.

Tanaka N, Ishihara M, Lamphier MS, Nozawa H, Matsuyama T, Mak TW, Aizawa S, Tokino T, Oren M, Taniguchi T. 1996. Cooperation of the tumor suppressors IRF-1 and p53 in response to DNA damage. Nature 382:816-18.

Taniguchi T, Ogasawara K, Takaoka A, Tanaka N. 2001. IRF family of transcription factors as regulators of host defense. Annu Rev Immunol 19:623-55.

Verrier ER, Langevin C, Benmansour A, Boudinot P. 2011. Early antiviral response and virus-induced genes in fish. Dev Comp Immunol 35:1204-14.

Workenhe ST, Rise ML, Kibenge MJT, Kibenge FSB. 2010. The fight between the teleost fish immune response and aquatic viruses. Mol Immunol 47:2525-36.

Xia J, Hu G, Dong X, Liu Q, Zhang S. 2012. Molecular characterization and expression analysis of interferon regulatory factor 5 (IRF-5) in turbot, *Scophthalamus maximus*. Fish Shellfish Immunol 32:211-18.

Xiang Z, Dong C, Qi L, Chen W, Huang L, Li Z, Xia Q, Liu D, Huang M, Weng S, *et al.*, 2010. Characteristics of the interferon regulatory factor pairs zfIRF5/7 and their stimulation expression by ISKNV infection in zebrafish (*Danio rerio*). Dev Comp Immunol 34:1263-73.

Xiaoni G, Zhuo C, Xuzhen W, Dengqiang W, Xinwen C. 2011. Molecular cloning and characterization of interferon regulatory factor 1 (IRF-1), IRF-2 and IRF-5 in the chondrostean paddlefish *Polyodon spathula* and their phylogenetic importance in the Osteichthyes. Dev Comp Immunol 36:74-84.

Xu Q, Chang MX, Xiao FS, Huang B, Nie P. 2010. The gene and virus-induced expression of IRF-5 in grass carp *Ctenopharyngodon idella*. Vet Immunol Immunopathol 134:269-78.

Yanai H, Negishi H, Taniguchi T. 2012. The IRF family of transcription factors: inception, impact and implications in oncogenesis. OncoImmunology 1:1376-86.

Yao CL, Kong P, Huang XN, Wang ZY. 2010. Molecular cloning and expression of IRF1 in large yellow croaker, *Pseudosciaena crocea*. Fish Shellfish Immunol 28:654-60.

Zapata AG, Chiba A, Varas A. 1996. Cells and tissues of the immune system of fish. pp. 1-40 *In*: *The Fish Immune System: Organism, Pathogen, and Environment*. (eds. G. Iwama and T. Nakanishi). Academic Press, Inc. San Diego, CA, USA.

Zhang L and Pagano JS. 1997. IRF-7, a new interferon regulatory factor associated with Epstein-Barr virus latency. Mol Cell Biol 17:5748-57.

Web References

http://codgene.ca/; the Atlantic Cod Genomics and Broodstock Development Project (CGP).

https://www.ebi.ac.uk/Tools/msa/clustalo/; Web interface for Clustal Omega.

http://frodo.wi.mit.edu/; Web interface for Primer3.

http://www.ncbi.nlm.nih.gov/; National Center for Biotechnology Information web interface.

http://web.expasy.org/translate/; Web interface for ExPASy Translate tool.

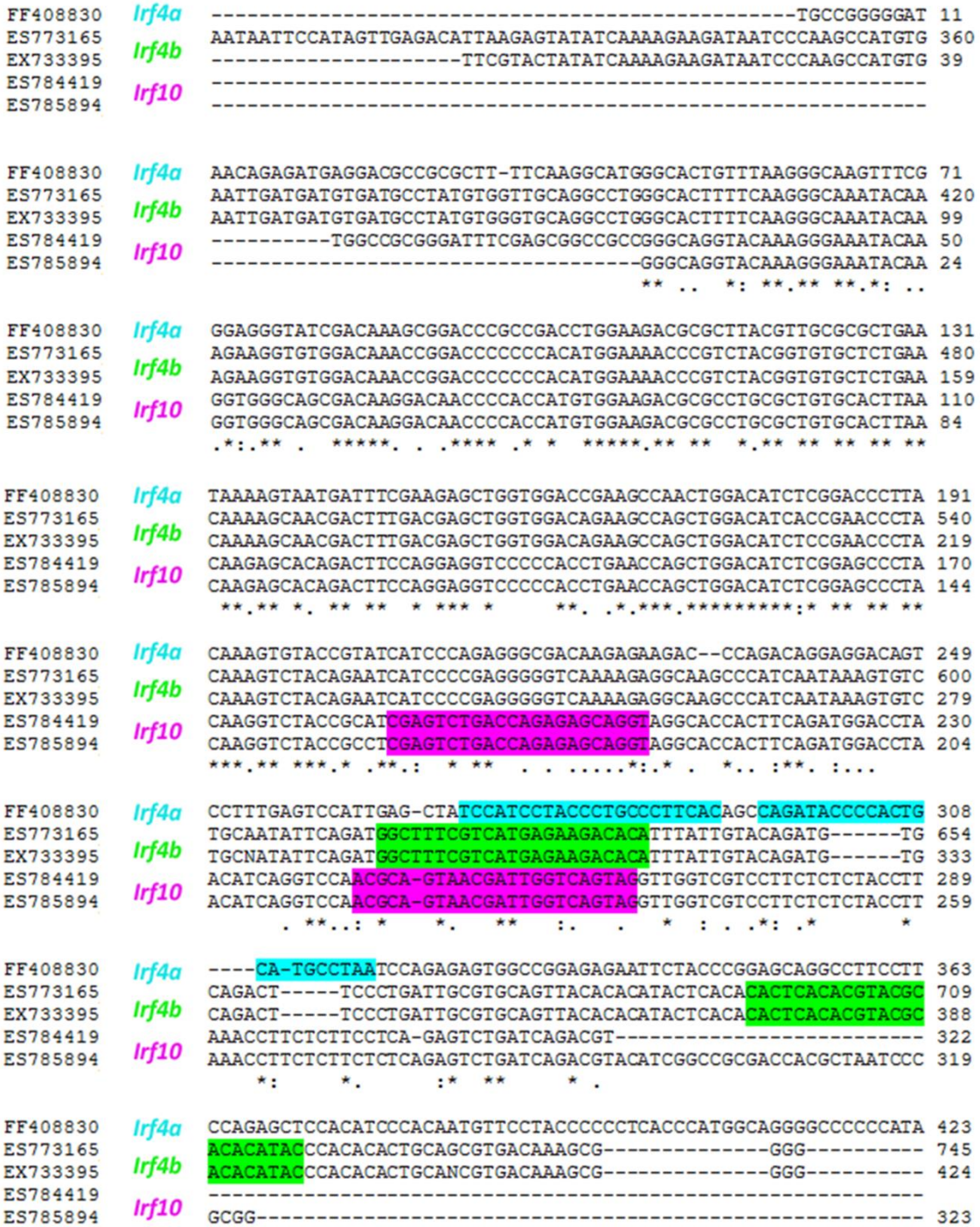
Appendix 1: *Irf4*-like Atlantic cod ESTs used to design paralogue specific RACE primers for characterization of cod *Irf4a, Irf4b* and *Irf10-v2.* A) Table summarizing cod *Irf4*-like ESTs found in dbEST. B) Partial alignment of Irf4-like ESTs. Conserved nucleotides are marked by an asterisk (*). The locations of RACE primers are indicated in blue for *Irf4a*, green for *Irf4b*, and purple for *Irf10-v2*. Alignment was constructed using Clustal Omega software (see Web References).

Appendix 1 A*: Gadus morhua* ESTs representing *Irf4*

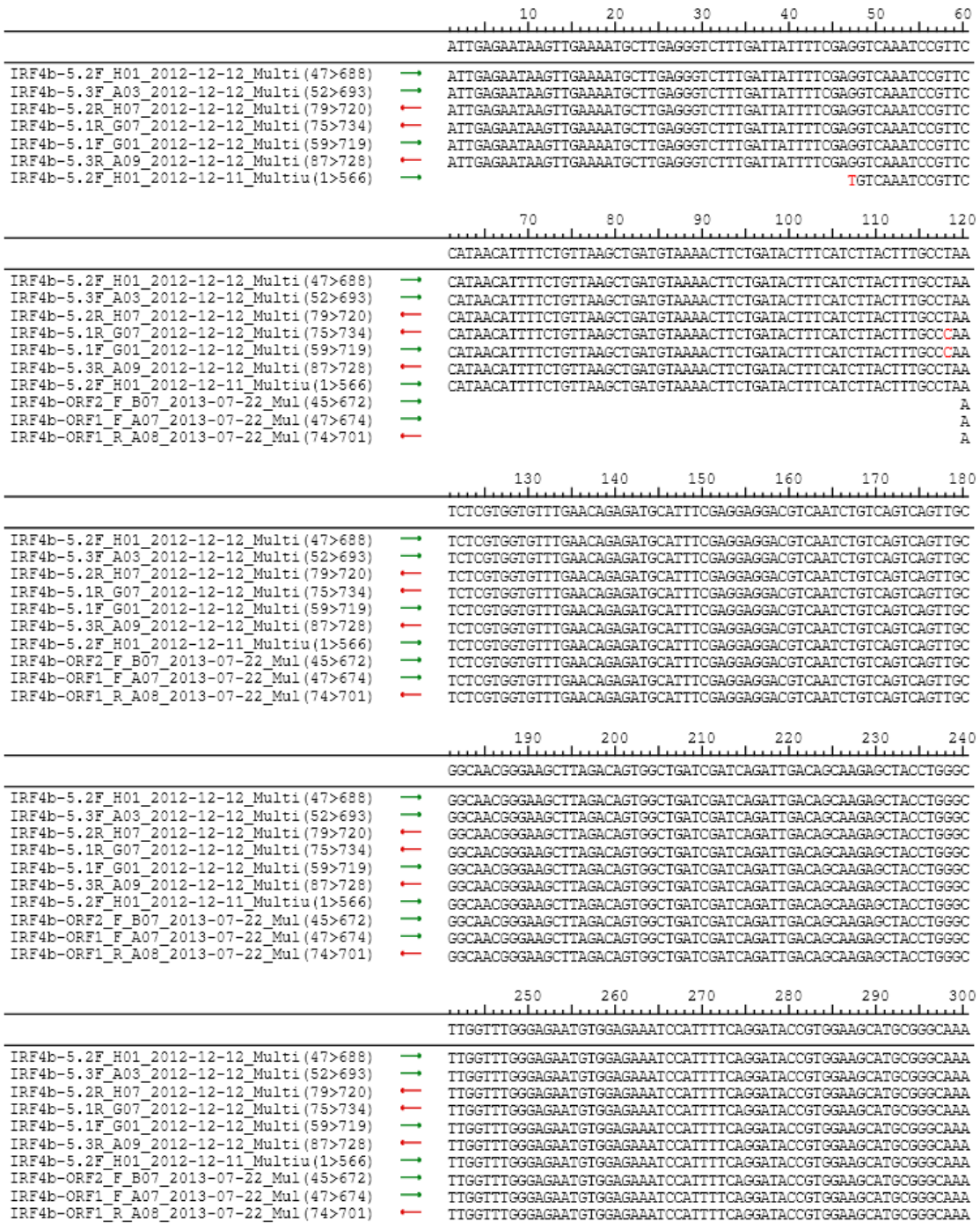
Genbank Accession Number	Library name	Tissue	Treatment	Best BLASTx hit
FF408830	gmapte	testis	none	IRF4 [*Paralichthys olivaceus*] (E= 2e-78). AEY55358
EX733395	ZNKAA	kidney	none	IRF4-like [*Oreochromis niloticus*] predicted (E= 8e-33). XP_003437930
ES773165	gmnbhkas	head kidney	ASAL	IRF4-like [*Oreochromis niloticus*] predicted (E= 3e-31). XP_003437930
*ES784419	gmnlsfic	spleen	Poly(I:C)	IRF4-like [*Oreochromis niloticus*] predicted (E=8e-22). XP_005448898
*ES785894	gmnlsfic	spleen	Poly(I:C)	IRF4-like [*Oreochromis niloticus*] predicted (E= 6e-22). XP_005448898

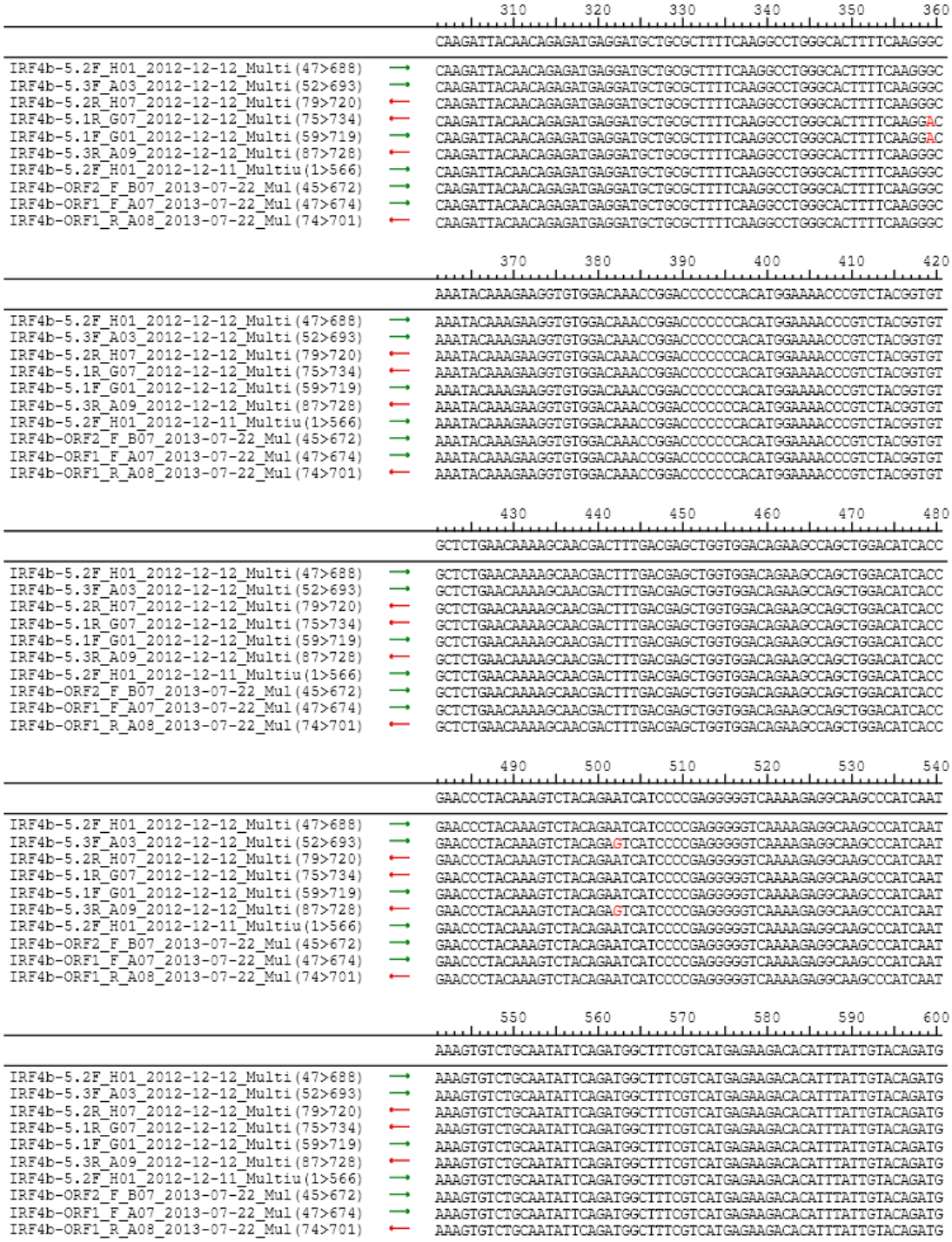
*BLASTx search returned many hits for *Irf10*, but *Irf4*-like sequence had the lowest E-value.

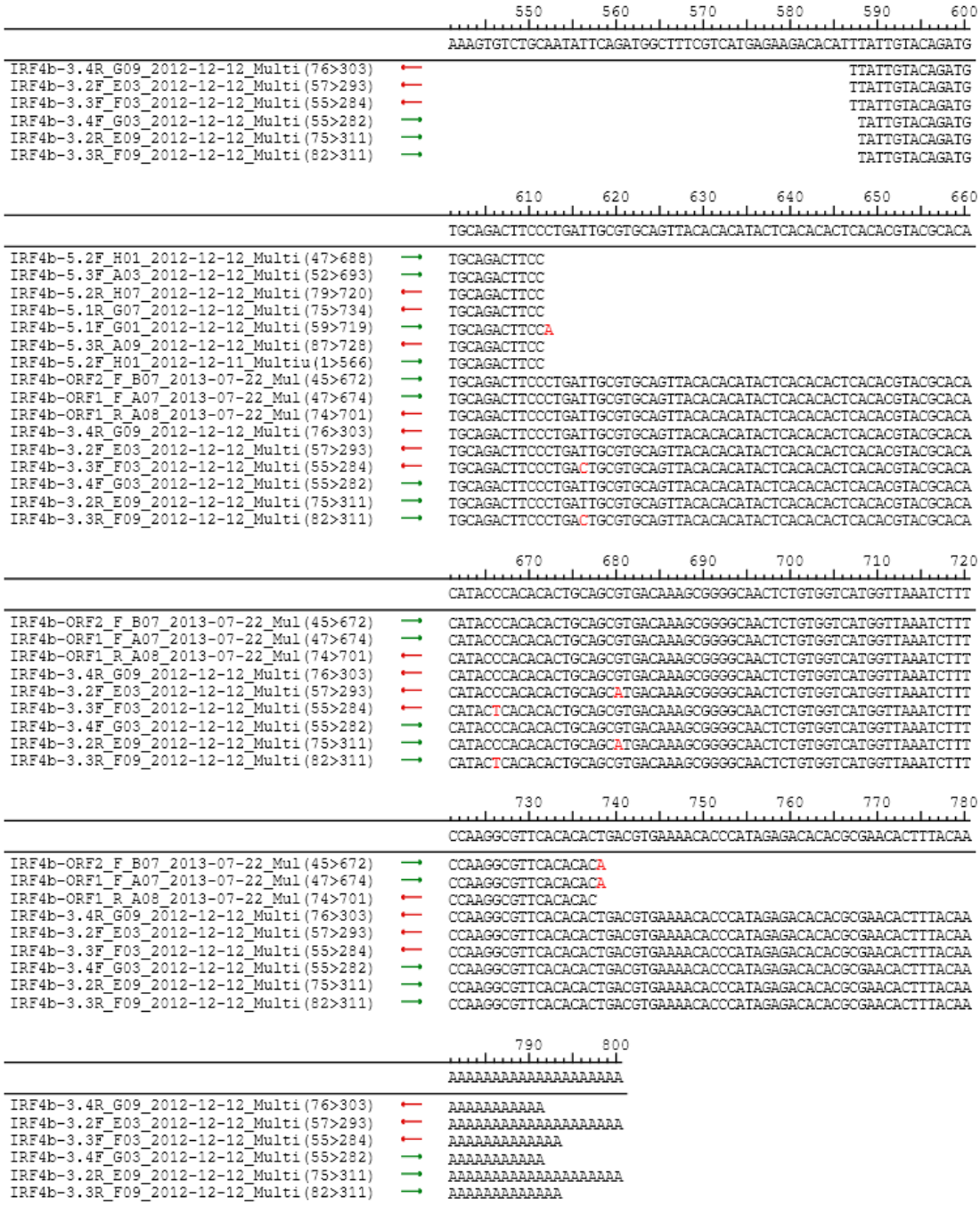
Appendix 1 B:


FF408830 Irf4a
ES773165 Irf4b
EX733395 Irf4b
ES784419 Irf10
ES785894 Irf10

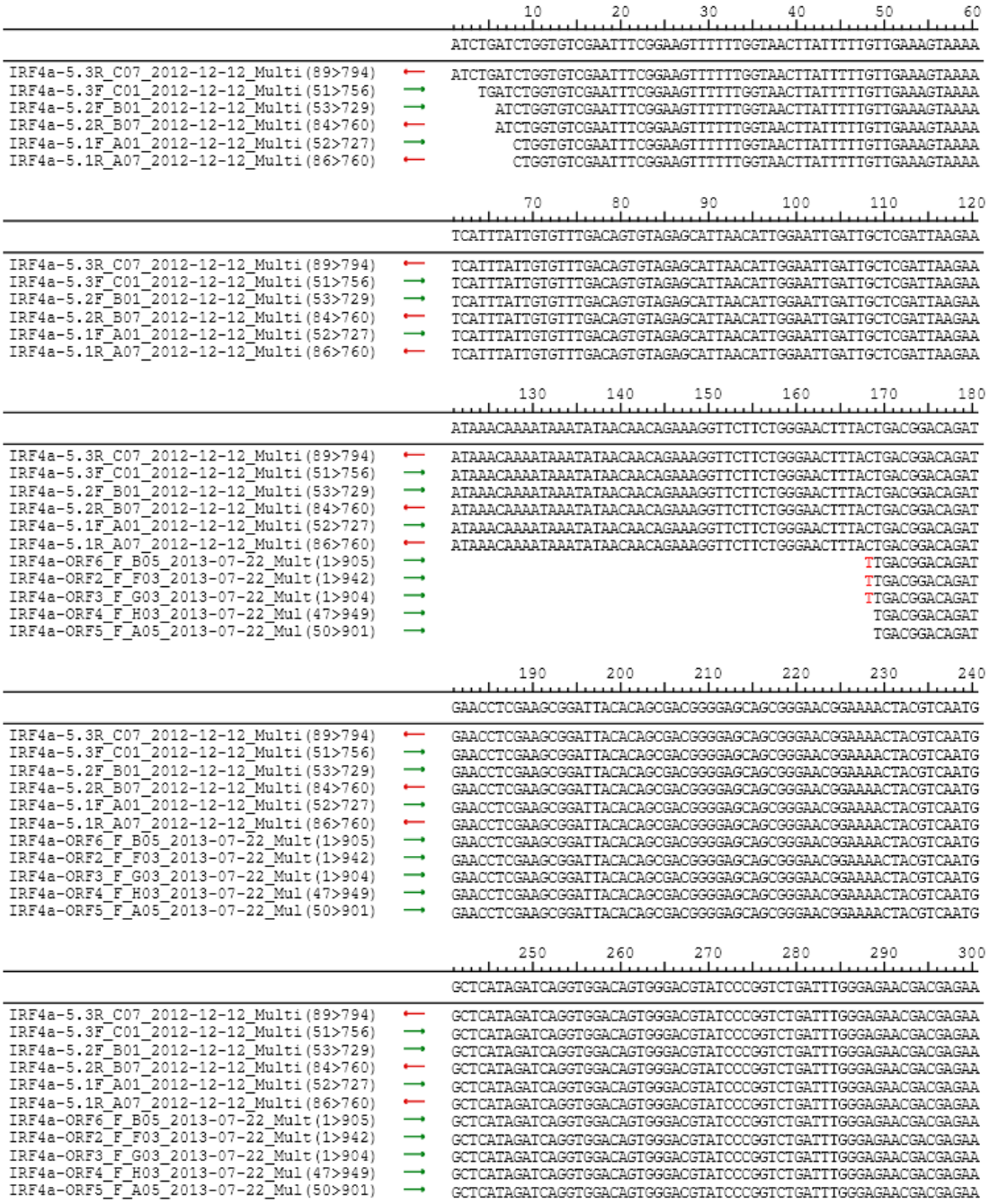
Appendix 2: Assembly of Atlantic cod *Irf4a* RACE and ORF PCR sequencing reads. Sequencing methods are described in section 2.1.2. Sequence data was assembled using Lasergene SeqMan Pro software (DNASTAR). Consensus sequence is indicated between horizontal lines. Note that naming appears incorrect as *Irf4a* and *Irf4b* names were switched after phylogenetic analysis based on similarity to zebrafish *Irf4* paralogues.

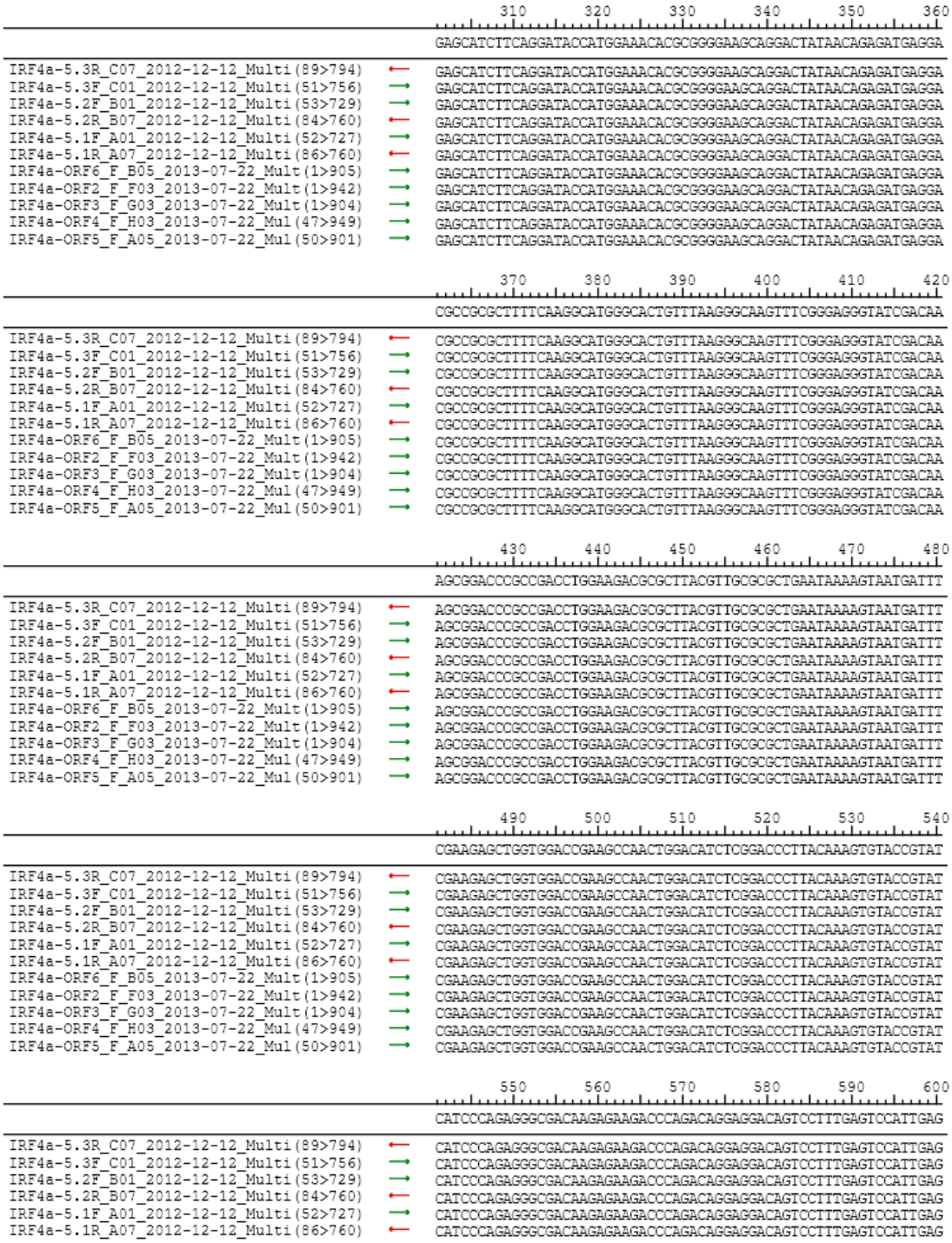

10 20 30 40 50 60
ATTGAGAATAAGTTGAAAATGCTTGAGGGTCTTTGATTATTTTCGAGGTCAAATCCGTTC
IRF4b-5.2F_H01_2012-12-12_Multi(47>688) ATTGAGAATAAGTTGAAAATGCTTGAGGGTCTTTGATTATTTTCGAGGTCAAATCCGTTC
IRF4b-5.3F_A03_2012-12-12_Multi(52>693) ATTGAGAATAAGTTGAAAATGCTTGAGGGTCTTTGATTATTTTCGAGGTCAAATCCGTTC
IRF4b-5.2R_H07_2012-12-12_Multi(79>720) ATTGAGAATAAGTTGAAAATGCTTGAGGGTCTTTGATTATTTTCGAGGTCAAATCCGTTC
IRF4b-5.1R_G07_2012-12-12_Multi(75>734) ATTGAGAATAAGTTGAAAATGCTTGAGGGTCTTTGATTATTTTCGAGGTCAAATCCGTTC
IRF4b-5.1F_G01_2012-12-12_Multi(59>719) ATTGAGAATAAGTTGAAAATGCTTGAGGGTCTTTGATTATTTTCGAGGTCAAATCCGTTC
IRF4b-5.3R_A09_2012-12-12_Multi(87>728) ATTGAGAATAAGTTGAAAATGCTTGAGGGTCTTTGATTATTTTCGAGGTCAAATCCGTTC
IRF4b-5.2F_H01_2012-12-11_Multiu(1>566) TGTCAAATCCGTTC
70 80 90 100 110 120
CATAACATTTTCTGTTAAGCTGATGTAAAACTTCTGATACTTTCATCTTACTTTGCCTAA
IRF4b-5.2F_H01_2012-12-12_Multi(47>688) CATAACATTTTCTGTTAAGCTGATGTAAAACTTCTGATACTTTCATCTTACTTTGCCTAA
IRF4b-5.3F_A03_2012-12-12_Multi(52>693) CATAACATTTTCTGTTAAGCTGATGTAAAACTTCTGATACTTTCATCTTACTTTGCCTAA
IRF4b-5.2R_H07_2012-12-12_Multi(79>720) CATAACATTTTCTGTTAAGCTGATGTAAAACTTCTGATACTTTCATCTTACTTTGCCTAA
IRF4b-5.1R_G07_2012-12-12_Multi(75>734) CATAACATTTTCTGTTAAGCTGATGTAAAACTTCTGATACTTTCATCTTACTTTGCCCAA
IRF4b-5.1F_G01_2012-12-12_Multi(59>719) CATAACATTTTCTGTTAAGCTGATGTAAAACTTCTGATACTTTCATCTTACTTTGCCCAA
IRF4b-5.3R_A09_2012-12-12_Multi(87>728) CATAACATTTTCTGTTAAGCTGATGTAAAACTTCTGATACTTTCATCTTACTTTGCCTAA
IRF4b-5.2F_H01_2012-12-11_Multiu(1>566) CATAACATTTTCTGTTAAGCTGATGTAAAACTTCTGATACTTTCATCTTACTTTGCCTAA
IRF4b-ORF2_F_B07_2013-07-22_Mul(45>672) A
IRF4b-ORF1_F_A07_2013-07-22_Mul(47>674) A
IRF4b-ORF1_R_A08_2013-07-22_Mul(74>701) A
130 140 150 160 170 180
TCTCGTGGTGTTTGAACAGAGATGCATTTCGAGGAGGACGTCAATCTGTCAGTCAGTTGC
IRF4b-5.2F_H01_2012-12-12_Multi(47>688) TCTCGTGGTGTTTGAACAGAGATGCATTTCGAGGAGGACGTCAATCTGTCAGTCAGTTGC
IRF4b-5.3F_A03_2012-12-12_Multi(52>693) TCTCGTGGTGTTTGAACAGAGATGCATTTCGAGGAGGACGTCAATCTGTCAGTCAGTTGC
IRF4b-5.2R_H07_2012-12-12_Multi(79>720) TCTCGTGGTGTTTGAACAGAGATGCATTTCGAGGAGGACGTCAATCTGTCAGTCAGTTGC
IRF4b-5.1R_G07_2012-12-12_Multi(75>734) TCTCGTGGTGTTTGAACAGAGATGCATTTCGAGGAGGACGTCAATCTGTCAGTCAGTTGC
IRF4b-5.1F_G01_2012-12-12_Multi(59>719) TCTCGTGGTGTTTGAACAGAGATGCATTTCGAGGAGGACGTCAATCTGTCAGTCAGTTGC
IRF4b-5.3R_A09_2012-12-12_Multi(87>728) TCTCGTGGTGTTTGAACAGAGATGCATTTCGAGGAGGACGTCAATCTGTCAGTCAGTTGC
IRF4b-5.2F_H01_2012-12-11_Multiu(1>566) TCTCGTGGTGTTTGAACAGAGATGCATTTCGAGGAGGACGTCAATCTGTCAGTCAGTTGC
IRF4b-ORF2_F_B07_2013-07-22_Mul(45>672) TCTCGTGGTGTTTGAACAGAGATGCATTTCGAGGAGGACGTCAATCTGTCAGTCAGTTGC
IRF4b-ORF1_F_A07_2013-07-22_Mul(47>674) TCTCGTGGTGTTTGAACAGAGATGCATTTCGAGGAGGACGTCAATCTGTCAGTCAGTTGC
IRF4b-ORF1_R_A08_2013-07-22_Mul(74>701) TCTCGTGGTGTTTGAACAGAGATGCATTTCGAGGAGGACGTCAATCTGTCAGTCAGTTGC
190 200 210 220 230 240
GGCAACGGGAAGCTTAGACAGTGGCTGATCGATCAGATTGACAGCAAGAGCTACCTGGGC
IRF4b-5.2F_H01_2012-12-12_Multi(47>688) GGCAACGGGAAGCTTAGACAGTGGCTGATCGATCAGATTGACAGCAAGAGCTACCTGGGC
IRF4b-5.3F_A03_2012-12-12_Multi(52>693) GGCAACGGGAAGCTTAGACAGTGGCTGATCGATCAGATTGACAGCAAGAGCTACCTGGGC
IRF4b-5.2R_H07_2012-12-12_Multi(79>720) GGCAACGGGAAGCTTAGACAGTGGCTGATCGATCAGATTGACAGCAAGAGCTACCTGGGC
IRF4b-5.1R_G07_2012-12-12_Multi(75>734) GGCAACGGGAAGCTTAGACAGTGGCTGATCGATCAGATTGACAGCAAGAGCTACCTGGGC
IRF4b-5.1F_G01_2012-12-12_Multi(59>719) GGCAACGGGAAGCTTAGACAGTGGCTGATCGATCAGATTGACAGCAAGAGCTACCTGGGC
IRF4b-5.3R_A09_2012-12-12_Multi(87>728) GGCAACGGGAAGCTTAGACAGTGGCTGATCGATCAGATTGACAGCAAGAGCTACCTGGGC
IRF4b-5.2F_H01_2012-12-11_Multiu(1>566) GGCAACGGGAAGCTTAGACAGTGGCTGATCGATCAGATTGACAGCAAGAGCTACCTGGGC
IRF4b-ORF2_F_B07_2013-07-22_Mul(45>672) GGCAACGGGAAGCTTAGACAGTGGCTGATCGATCAGATTGACAGCAAGAGCTACCTGGGC
IRF4b-ORF1_F_A07_2013-07-22_Mul(47>674) GGCAACGGGAAGCTTAGACAGTGGCTGATCGATCAGATTGACAGCAAGAGCTACCTGGGC
IRF4b-ORF1_R_A08_2013-07-22_Mul(74>701) GGCAACGGGAAGCTTAGACAGTGGCTGATCGATCAGATTGACAGCAAGAGCTACCTGGGC
250 260 270 280 290 300
TTGGTTTGGGAGAATGTGGAGAAATCCATTTTCAGGATACCGTGGAAGCATGCGGGCAAA
IRF4b-5.2F_H01_2012-12-12_Multi(47>688) TTGGTTTGGGAGAATGTGGAGAAATCCATTTTCAGGATACCGTGGAAGCATGCGGGCAAA
IRF4b-5.3F_A03_2012-12-12_Multi(52>693) TTGGTTTGGGAGAATGTGGAGAAATCCATTTTCAGGATACCGTGGAAGCATGCGGGCAAA
IRF4b-5.2R_H07_2012-12-12_Multi(79>720) TTGGTTTGGGAGAATGTGGAGAAATCCATTTTCAGGATACCGTGGAAGCATGCGGGCAAA
IRF4b-5.1R_G07_2012-12-12_Multi(75>734) TTGGTTTGGGAGAATGTGGAGAAATCCATTTTCAGGATACCGTGGAAGCATGCGGGCAAA
IRF4b-5.1F_G01_2012-12-12_Multi(59>719) TTGGTTTGGGAGAATGTGGAGAAATCCATTTTCAGGATACCGTGGAAGCATGCGGGCAAA
IRF4b-5.3R_A09_2012-12-12_Multi(87>728) TTGGTTTGGGAGAATGTGGAGAAATCCATTTTCAGGATACCGTGGAAGCATGCGGGCAAA
IRF4b-5.2F_H01_2012-12-11_Multiu(1>566) TTGGTTTGGGAGAATGTGGAGAAATCCATTTTCAGGATACCGTGGAAGCATGCGGGCAAA
IRF4b-ORF2_F_B07_2013-07-22_Mul(45>672) TTGGTTTGGGAGAATGTGGAGAAATCCATTTTCAGGATACCGTGGAAGCATGCGGGCAAA
IRF4b-ORF1_F_A07_2013-07-22_Mul(47>674) TTGGTTTGGGAGAATGTGGAGAAATCCATTTTCAGGATACCGTGGAAGCATGCGGGCAAA
IRF4b-ORF1_R_A08_2013-07-22_Mul(74>701) TTGGTTTGGGAGAATGTGGAGAAATCCATTTTCAGGATACCGTGGAAGCATGCGGGCAAA

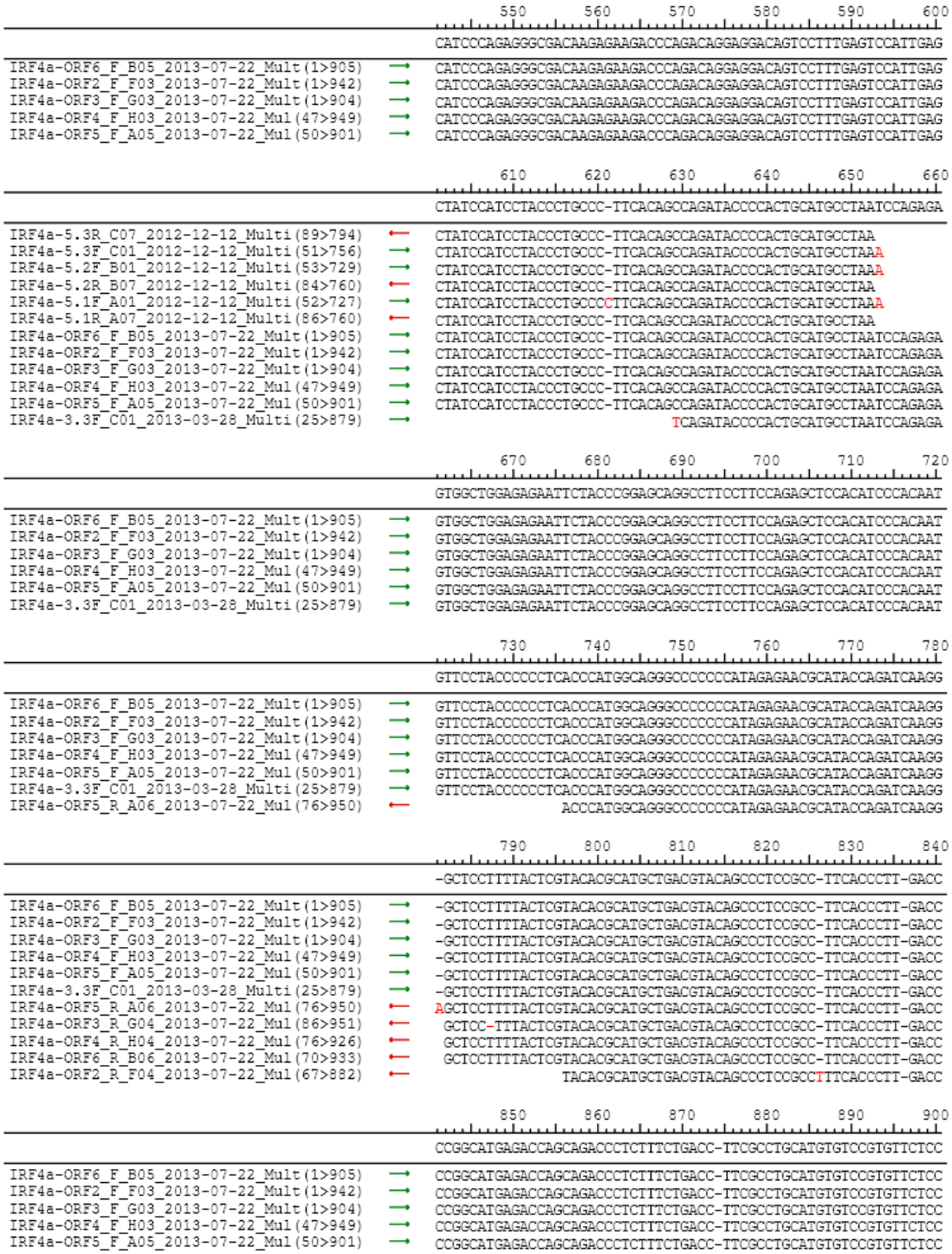

310 320 330 340 350 360
CAAGATTACAACAGAGATGAGGATGCTGCGCTTTTCAAGGCCTGGGCACTTTTCAAGGGC
IRF4b-5.2F_H01_2012-12-12_Multi(47>688)
IRF4b-5.3F_A03_2012-12-12_Multi(52>693)
IRF4b-5.2R_H07_2012-12-12_Multi(79>720)
IRF4b-5.1R_G07_2012-12-12_Multi(75>734)
IRF4b-5.1F_G01_2012-12-12_Multi(59>719)
IRF4b-5.3R_A09_2012-12-12_Multi(87>728)
IRF4b-5.2F_H01_2012-12-11_Multiu(1>566)
IRF4b-ORF2_F_B07_2013-07-22_Mul(45>672)
IRF4b-ORF1_F_A07_2013-07-22_Mul(47>674)
IRF4b-ORF1_R_A08_2013-07-22_Mul(74>701)
370 380 390 400 410 420
AAATACAAAGAAGGTGTGGACAAACCGGACCCCCCCCACATGGAAAACCCGTCTACGGTGT
430 440 450 460 470 480
GCTCTGAACAAAAGCAACGACTTTGACGAGCTGGTGGACAGAAGCCAGCTGGACATCACC
490 500 510 520 530 540
GAACCCTACAAAGTCTACAGAATCATCCCCGAGGGGGTCAAAAGAGGCAAGCCCATCAAT
550 560 570 580 590 600
AAAGTGTCTGCAATATTCAGATGGCTTTCGTCATGAGAAGACACATTTATTGTACAGATG

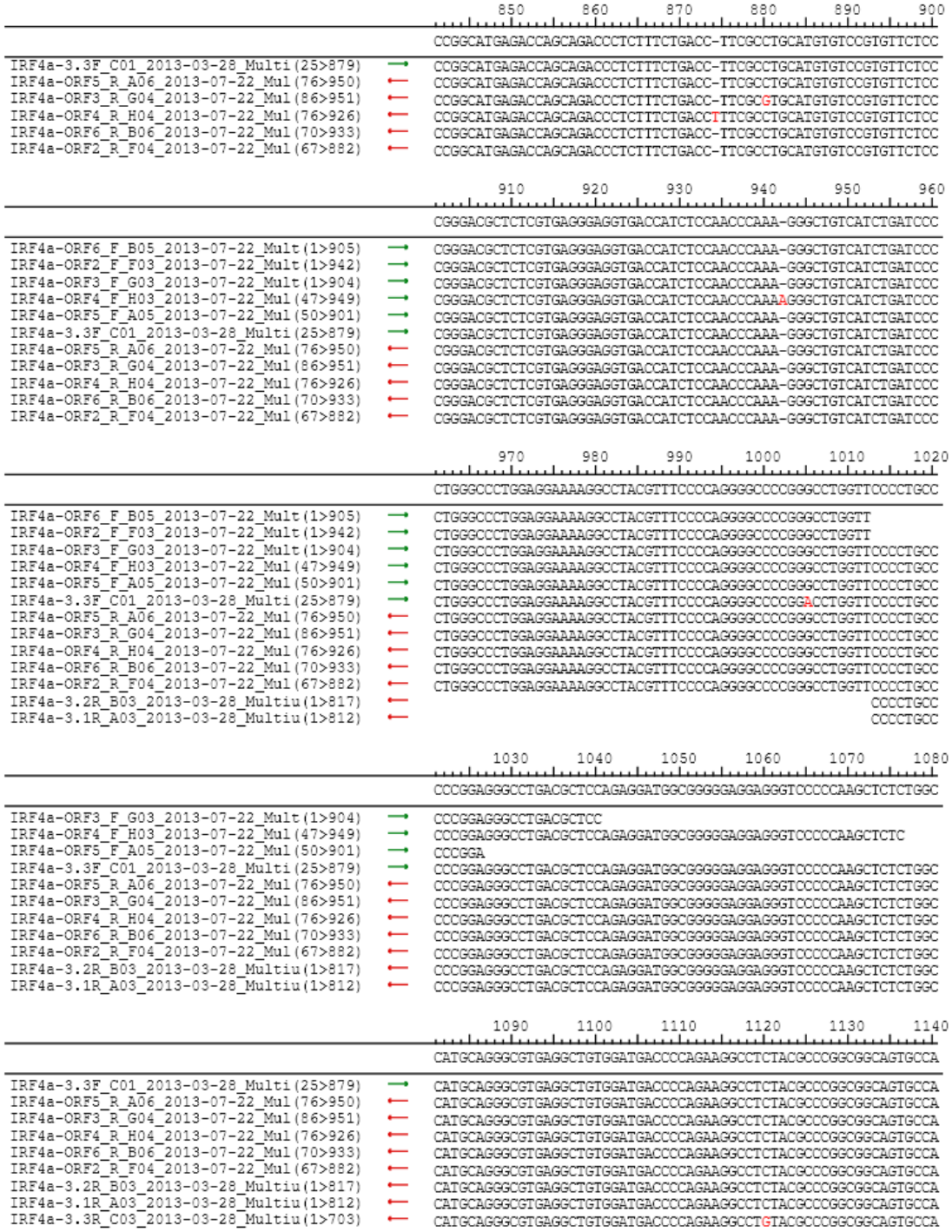

550 560 570 580 590 600
AAAGTGTCTGCAATATTCAGATGGCTTTCGTCATGAGAAGACACATTTATTGTACAGATG
IRF4b-3.4R_G09_2012-12-12_Multi(76>303) TTATTGTACAGATG
IRF4b-3.2F_E03_2012-12-12_Multi(57>293) TTATTGTACAGATG
IRF4b-3.3F_F03_2012-12-12_Multi(55>284) TTATTGTACAGATG
IRF4b-3.4F_G03_2012-12-12_Multi(55>282) TATTGTACAGATG
IRF4b-3.2R_E09_2012-12-12_Multi(75>311) TATTGTACAGATG
IRF4b-3.3R_F09_2012-12-12_Multi(82>311) TATTGTACAGATG
610 620 630 640 650 660
TGCAGACTTCCCTGATTGCGTGCAGTTACACACATACTCACACACTCACACGTACGCACA
IRF4b-5.2F_H01_2012-12-12_Multi(47>688) TGCAGACTTCC
IRF4b-5.3F_A03_2012-12-12_Multi(52>693) TGCAGACTTCC
IRF4b-5.2R_H07_2012-12-12_Multi(79>720) TGCAGACTTCC
IRF4b-5.1R_G07_2012-12-12_Multi(75>734) TGCAGACTTCC
IRF4b-5.1F_G01_2012-12-12_Multi(59>719) TGCAGACTTCCA
IRF4b-5.3R_A09_2012-12-12_Multi(87>728) TGCAGACTTCC
IRF4b-5.2F_H01_2012-12-11_Multiu(1>566) TGCAGACTTCC
IRF4b-ORF2_F_B07_2013-07-22_Mul(45>672) TGCAGACTTCCCTGATTGCGTGCAGTTACACACATACTCACACACTCACACGTACGCACA
IRF4b-ORF1_F_A07_2013-07-22_Mul(47>674) TGCAGACTTCCCTGATTGCGTGCAGTTACACACATACTCACACACTCACACGTACGCACA
IRF4b-ORF1_R_A08_2013-07-22_Mul(74>701) TGCAGACTTCCCTGATTGCGTGCAGTTACACACATACTCACACACTCACACGTACGCACA
IRF4b-3.4R_G09_2012-12-12_Multi(76>303) TGCAGACTTCCCTGATTGCGTGCAGTTACACACATACTCACACACTCACACGTACGCACA
IRF4b-3.2F_E03_2012-12-12_Multi(57>293) TGCAGACTTCCCTGATTGCGTGCAGTTACACACATACTCACACACTCACACGTACGCACA
IRF4b-3.3F_F03_2012-12-12_Multi(55>284) TGCAGACTTCCCTGACTGCGTGCAGTTACACACATACTCACACACTCACACGTACGCACA
IRF4b-3.4F_G03_2012-12-12_Multi(55>282) TGCAGACTTCCCTGATTGCGTGCAGTTACACACATACTCACACACTCACACGTACGCACA
IRF4b-3.2R_E09_2012-12-12_Multi(75>311) TGCAGACTTCCCTGATTGCGTGCAGTTACACACATACTCACACACTCACACGTACGCACA
IRF4b-3.3R_F09_2012-12-12_Multi(82>311) TGCAGACTTCCCTGACTGCGTGCAGTTACACACATACTCACACACTCACACGTACGCACA
670 680 690 700 710 720
CATACCCACACACTGCAGCGTGACAAAGCGGGGCAACTCTGTGGTCATGGTTAAATCTTT
IRF4b-ORF2_F_B07_2013-07-22_Mul(45>672) CATACCCACACACTGCAGCGTGACAAAGCGGGGCAACTCTGTGGTCATGGTTAAATCTTT
IRF4b-ORF1_F_A07_2013-07-22_Mul(47>674) CATACCCACACACTGCAGCGTGACAAAGCGGGGCAACTCTGTGGTCATGGTTAAATCTTT
IRF4b-ORF1_R_A08_2013-07-22_Mul(74>701) CATACCCACACACTGCAGCGTGACAAAGCGGGGCAACTCTGTGGTCATGGTTAAATCTTT
IRF4b-3.4R_G09_2012-12-12_Multi(76>303) CATACCCACACACTGCAGCGTGACAAAGCGGGGCAACTCTGTGGTCATGGTTAAATCTTT
IRF4b-3.2F_E03_2012-12-12_Multi(57>293) CATACCCACACACTGCAGCATGACAAAGCGGGGCAACTCTGTGGTCATGGTTAAATCTTT
IRF4b-3.3F_F03_2012-12-12_Multi(55>284) CATACTCACACACTGCAGCGTGACAAAGCGGGGCAACTCTGTGGTCATGGTTAAATCTTT
IRF4b-3.4F_G03_2012-12-12_Multi(55>282) CATACCCACACACTGCAGCGTGACAAAGCGGGGCAACTCTGTGGTCATGGTTAAATCTTT
IRF4b-3.2R_E09_2012-12-12_Multi(75>311) CATACCCACACACTGCAGCATGACAAAGCGGGGCAACTCTGTGGTCATGGTTAAATCTTT
IRF4b-3.3R_F09_2012-12-12_Multi(82>311) CATACTCACACACTGCAGCGTGACAAAGCGGGGCAACTCTGTGGTCATGGTTAAATCTTT
730 740 750 760 770 780
CCAAGGCGTTCACACACTGACGTGAAAACACCCATAGAGACACACGCGAACACTTTACAA
IRF4b-ORF2_F_B07_2013-07-22_Mul(45>672) CCAAGGCGTTCACACACA
IRF4b-ORF1_F_A07_2013-07-22_Mul(47>674) CCAAGGCGTTCACACACA
IRF4b-ORF1_R_A08_2013-07-22_Mul(74>701) CCAAGGCGTTCACACAC
IRF4b-3.4R_G09_2012-12-12_Multi(76>303) CCAAGGCGTTCACACACTGACGTGAAAACACCCATAGAGACACACGCGAACACTTTACAA
IRF4b-3.2F_E03_2012-12-12_Multi(57>293) CCAAGGCGTTCACACACTGACGTGAAAACACCCATAGAGACACACGCGAACACTTTACAA
IRF4b-3.3F_F03_2012-12-12_Multi(55>284) CCAAGGCGTTCACACACTGACGTGAAAACACCCATAGAGACACACGCGAACACTTTACAA
IRF4b-3.4F_G03_2012-12-12_Multi(55>282) CCAAGGCGTTCACACACTGACGTGAAAACACCCATAGAGACACACGCGAACACTTTACAA
IRF4b-3.2R_E09_2012-12-12_Multi(75>311) CCAAGGCGTTCACACACTGACGTGAAAACACCCATAGAGACACACGCGAACACTTTACAA
IRF4b-3.3R_F09_2012-12-12_Multi(82>311) CCAAGGCGTTCACACACTGACGTGAAAACACCCATAGAGACACACGCGAACACTTTACAA
790 800
AAAAAAAAAAAAAAAAAAAA
IRF4b-3.4R_G09_2012-12-12_Multi(76>303) AAAAAAAAAA
IRF4b-3.2F_E03_2012-12-12_Multi(57>293) AAAAAAAAAAAAAAAAAAAA
IRF4b-3.3F_F03_2012-12-12_Multi(55>284) AAAAAAAAAAAA
IRF4b-3.4F_G03_2012-12-12_Multi(55>282) AAAAAAAAAA
IRF4b-3.2R_E09_2012-12-12_Multi(75>311) AAAAAAAAAAAAAAAAAAAA
IRF4b-3.3R_F09_2012-12-12_Multi(82>311) AAAAAAAAAAAA

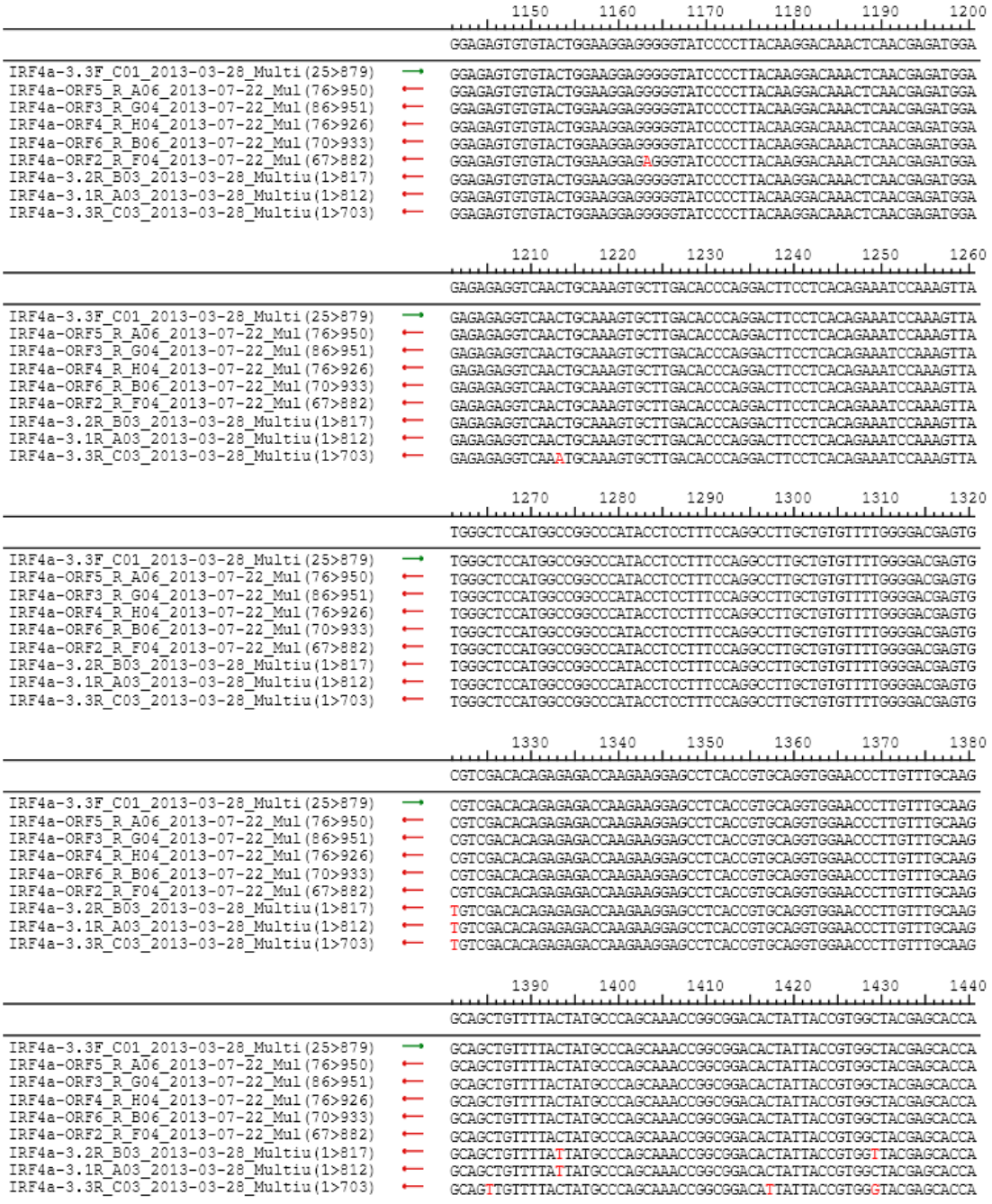
Appendix 3: Assembly of Atlantic cod *Irf4b* RACE and ORF PCR sequencing reads. Sequencing methods are described in section 2.1.2. Sequence data was assembled using Lasergene SeqMan Pro software (DNASTAR). Consensus sequence is indicated between horizontal lines. Note that naming appears incorrect as *Irf4a* and *Irf4b* names were switched after phylogenetic analysis based on similarity to zebrafish *Irf4* paralogues.

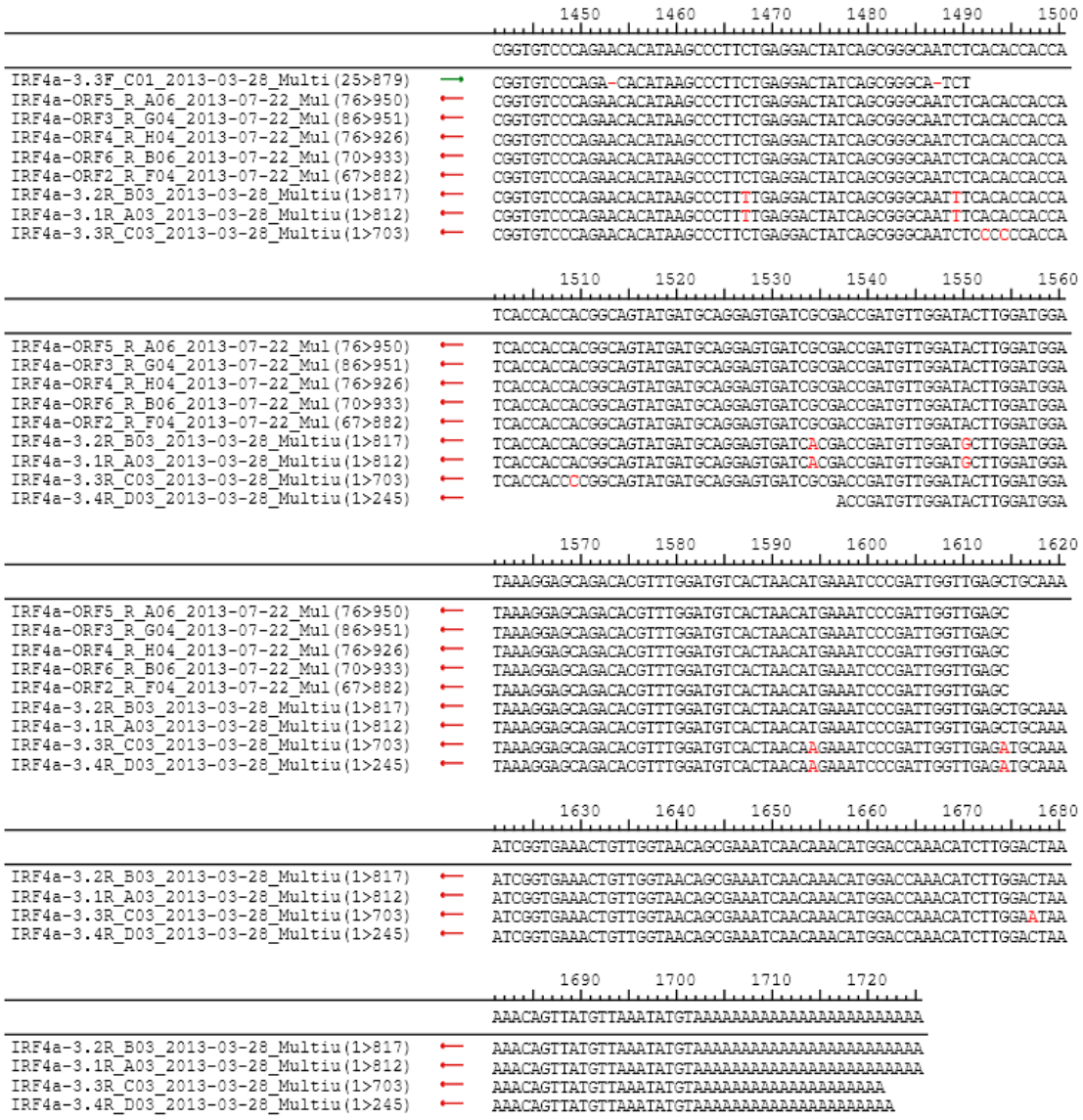

10 20 30 40 50 60
ATCTGATCTGGTGTCGAATTTCGGAAGTTTTTTGGTAACTTATTTTTGTTGAAAGTAAAA
IRF4a-5.3R_C07_2012-12-12_Multi(89>794)
IRF4a-5.3F_C01_2012-12-12_Multi(51>756)
IRF4a-5.2F_B01_2012-12-12_Multi(53>729)
IRF4a-5.2R_B07_2012-12-12_Multi(84>760)
IRF4a-5.1F_A01_2012-12-12_Multi(52>727)
IRF4a-5.1R_A07_2012-12-12_Multi(86>760)
70 80 90 100 110 120
TCATTTATTGTGTTTGACAGTGTAGAGCATTAACATTGGAATTGATTGCTCGATTAAGAA
IRF4a-5.3R_C07_2012-12-12_Multi(89>794)
IRF4a-5.3F_C01_2012-12-12_Multi(51>756)
IRF4a-5.2F_B01_2012-12-12_Multi(53>729)
IRF4a-5.2R_B07_2012-12-12_Multi(84>760)
IRF4a-5.1F_A01_2012-12-12_Multi(52>727)
IRF4a-5.1R_A07_2012-12-12_Multi(86>760)
130 140 150 160 170 180
ATAAACAAAATAAATATAACAACAGAAAGGTTCTTCTGGGAACTTTACTGACGGACAGAT
IRF4a-5.3R_C07_2012-12-12_Multi(89>794)
IRF4a-5.3F_C01_2012-12-12_Multi(51>756)
IRF4a-5.2F_B01_2012-12-12_Multi(53>729)
IRF4a-5.2R_B07_2012-12-12_Multi(84>760)
IRF4a-5.1F_A01_2012-12-12_Multi(52>727)
IRF4a-5.1R_A07_2012-12-12_Multi(86>760)
IRF4a-ORF6_F_B05_2013-07-22_Mult(1>905)
IRF4a-ORF2_F_F03_2013-07-22_Mult(1>942)
IRF4a-ORF3_F_G03_2013-07-22_Mult(1>904)
IRF4a-ORF4_F_H03_2013-07-22_Mul(47>949)
IRF4a-ORF5_F_A05_2013-07-22_Mul(50>901)
190 200 210 220 230 240
GAACCTCGAAGCGGATTACACAGCGACGGGGAGCAGCGGGAACGGAAAACTACGTCAATG
IRF4a-5.3R_C07_2012-12-12_Multi(89>794)
IRF4a-5.3F_C01_2012-12-12_Multi(51>756)
IRF4a-5.2F_B01_2012-12-12_Multi(53>729)
IRF4a-5.2R_B07_2012-12-12_Multi(84>760)
IRF4a-5.1F_A01_2012-12-12_Multi(52>727)
IRF4a-5.1R_A07_2012-12-12_Multi(86>760)
IRF4a-ORF6_F_B05_2013-07-22_Mult(1>905)
IRF4a-ORF2_F_F03_2013-07-22_Mult(1>942)
IRF4a-ORF3_F_G03_2013-07-22_Mult(1>904)
IRF4a-ORF4_F_H03_2013-07-22_Mul(47>949)
IRF4a-ORF5_F_A05_2013-07-22_Mul(50>901)
250 260 270 280 290 300
GCTCATAGATCAGGTGGACAGTGGGACGTATCCCGGTCTGATTTGGGAGAACGACGAGAA
IRF4a-5.3R_C07_2012-12-12_Multi(89>794)
IRF4a-5.3F_C01_2012-12-12_Multi(51>756)
IRF4a-5.2F_B01_2012-12-12_Multi(53>729)
IRF4a-5.2R_B07_2012-12-12_Multi(84>760)
IRF4a-5.1F_A01_2012-12-12_Multi(52>727)
IRF4a-5.1R_A07_2012-12-12_Multi(86>760)
IRF4a-ORF6_F_B05_2013-07-22_Mult(1>905)
IRF4a-ORF2_F_F03_2013-07-22_Mult(1>942)
IRF4a-ORF3_F_G03_2013-07-22_Mult(1>904)
IRF4a-ORF4_F_H03_2013-07-22_Mul(47>949)
IRF4a-ORF5_F_A05_2013-07-22_Mul(50>901)


310 320 330 340 350 360
GAGCATCTTCAGGATACCATGGAAACACGCGGGGAAGCAGGACTATAACAGAGATGAGGA
IRF4a-5.3R_C07_2012-12-12_Multi(89>794)
IRF4a-5.3F_C01_2012-12-12_Multi(51>756)
IRF4a-5.2F_B01_2012-12-12_Multi(53>729)
IRF4a-5.2R_B07_2012-12-12_Multi(84>760)
IRF4a-5.1F_A01_2012-12-12_Multi(52>727)
IRF4a-5.1R_A07_2012-12-12_Multi(86>760)
IRF4a-ORF6_F_B05_2013-07-22_Mult(1>905)
IRF4a-ORF2_F_F03_2013-07-22_Mult(1>942)
IRF4a-ORF3_F_G03_2013-07-22_Mult(1>904)
IRF4a-ORF4_F_H03_2013-07-22_Mul(47>949)
IRF4a-ORF5_F_A05_2013-07-22_Mul(50>901)
GAGCATCTTCAGGATACCATGGAAACACGCGGGGAAGCAGGACTATAACAGAGATGAGGA
370 380 390 400 410 420
CGCCGCGCTTTTCAAGGCATGGGCACTGTTTAAGGGCAAGTTTCGGGAGGGTATCGACAA
CGCCGCGCTTTTCAAGGCATGGGCACTGTTTAAGGGCAAGTTTCGGGAGGGTATCGACAA
430 440 450 460 470 480
AGCGGACCCGCCGACCTGGAAGACGCGCTTACGTTGCGCGCTGAATAAAAGTAATGATTT
AGCGGACCCGCCGACCTGGAAGACGCGCTTACGTTGCGCGCTGAATAAAAGTAATGATTT
490 500 510 520 530 540
CGAAGAGCTGGTGGACCGAAGCCAACTGGACATCTCGGACCCTTACAAAGTGTACCGTAT
CGAAGAGCTGGTGGACCGAAGCCAACTGGACATCTCGGACCCTTACAAAGTGTACCGTAT
550 560 570 580 590 600
CATCCCAGAGGGCGACAAGAGAAGACCCAGACAGGAGGACAGTCCTTTGAGTCCATTGAG
CATCCCAGAGGGCGACAAGAGAAGACCCAGACAGGAGGACAGTCCTTTGAGTCCATTGAG

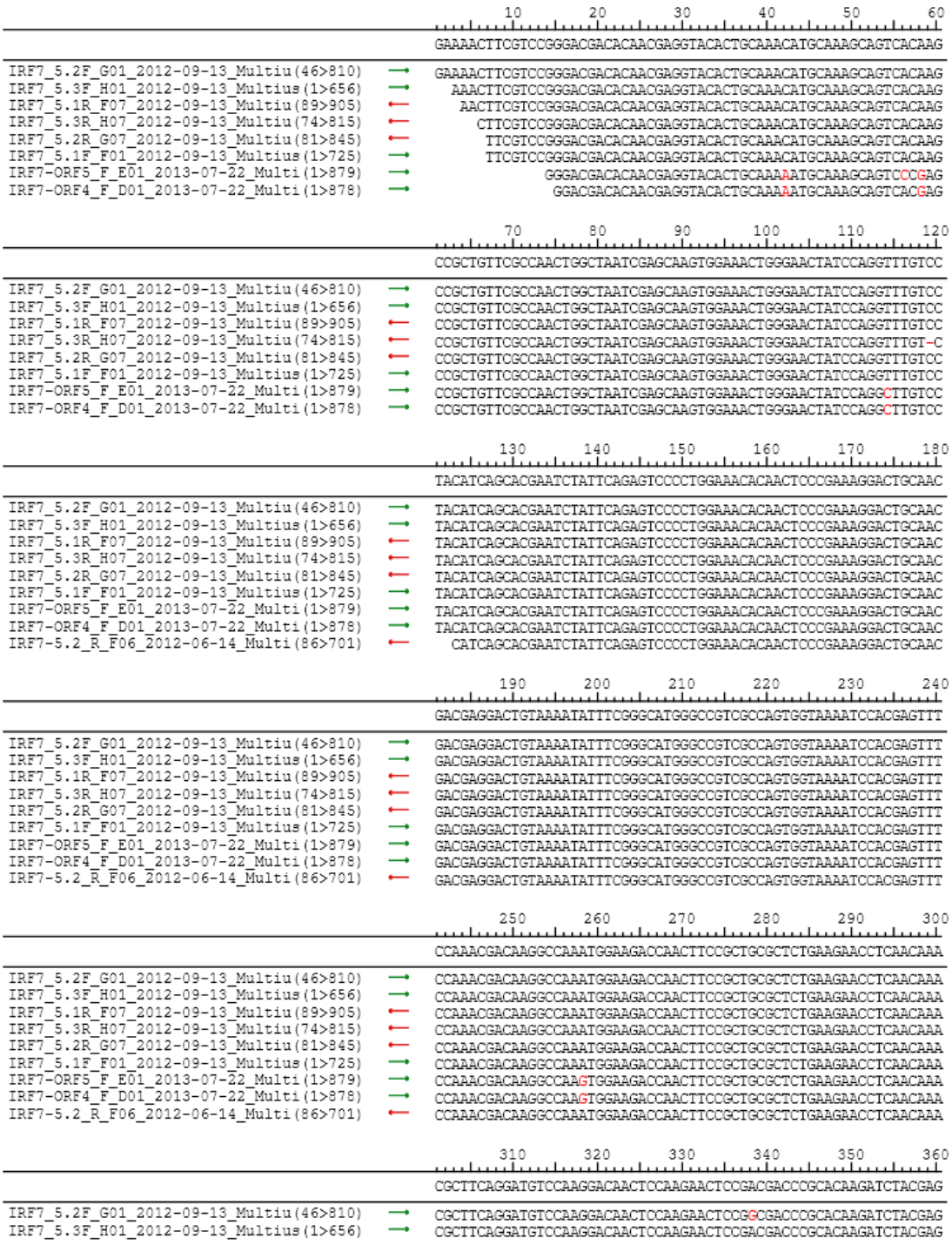

550 560 570 580 590 600
CATCCCAGAGGGCGACAAGAGAAGACCCAGACAGGAGGACAGTCCTTTGAGTCCATTGAG
IRF4a-ORF6_F_B05_2013-07-22_Mult(1>905) CATCCCAGAGGGCGACAAGAGAAGACCCAGACAGGAGGACAGTCCTTTGAGTCCATTGAG
IRF4a-ORF2_F_F03_2013-07-22_Mult(1>942) CATCCCAGAGGGCGACAAGAGAAGACCCAGACAGGAGGACAGTCCTTTGAGTCCATTGAG
IRF4a-ORF3_F_G03_2013-07-22_Mult(1>904) CATCCCAGAGGGCGACAAGAGAAGACCCAGACAGGAGGACAGTCCTTTGAGTCCATTGAG
IRF4a-ORF4_F_H03_2013-07-22_Mul(47>949) CATCCCAGAGGGCGACAAGAGAAGACCCAGACAGGAGGACAGTCCTTTGAGTCCATTGAG
IRF4a-ORF5_F_A05_2013-07-22_Mul(50>901) CATCCCAGAGGGCGACAAGAGAAGACCCAGACAGGAGGACAGTCCTTTGAGTCCATTGAG
610 620 630 640 650 660
CTATCCATCCTACCCTGCCC-TTCACAGCCAGATACCCCACTGCATGCCTAATCCAGAGA
IRF4a-5.3R_C07_2012-12-12_Multi(89>794) CTATCCATCCTACCCTGCCC-TTCACAGCCAGATACCCCACTGCATGCCTAA
IRF4a-5.3F_C01_2012-12-12_Multi(51>756) CTATCCATCCTACCCTGCCC-TTCACAGCCAGATACCCCACTGCATGCCTAAA
IRF4a-5.2F_B01_2012-12-12_Multi(53>729) CTATCCATCCTACCCTGCCC-TTCACAGCCAGATACCCCACTGCATGCCTAAA
IRF4a-5.2R_B07_2012-12-12_Multi(84>760) CTATCCATCCTACCCTGCCC-TTCACAGCCAGATACCCCACTGCATGCCTAA
IRF4a-5.1F_A01_2012-12-12_Multi(52>727) CTATCCATCCTACCCTGCCCCTTCACAGCCAGATACCCCACTGCATGCCTAAA
IRF4a-5.1R_A07_2012-12-12_Multi(86>760) CTATCCATCCTACCCTGCCC-TTCACAGCCAGATACCCCACTGCATGCCTAA
IRF4a-ORF6_F_B05_2013-07-22_Mult(1>905) CTATCCATCCTACCCTGCCC-TTCACAGCCAGATACCCCACTGCATGCCTAATCCAGAGA
IRF4a-ORF2_F_F03_2013-07-22_Mult(1>942) CTATCCATCCTACCCTGCCC-TTCACAGCCAGATACCCCACTGCATGCCTAATCCAGAGA
IRF4a-ORF3_F_G03_2013-07-22_Mult(1>904) CTATCCATCCTACCCTGCCC-TTCACAGCCAGATACCCCACTGCATGCCTAATCCAGAGA
IRF4a-ORF4_F_H03_2013-07-22_Mul(47>949) CTATCCATCCTACCCTGCCC-TTCACAGCCAGATACCCCACTGCATGCCTAATCCAGAGA
IRF4a-ORF5_F_A05_2013-07-22_Mul(50>901) CTATCCATCCTACCCTGCCC-TTCACAGCCAGATACCCCACTGCATGCCTAATCCAGAGA
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) TCAGATACCCCACTGCATGCCTAATCCAGAGA
670 680 690 700 710 720
GTGGCTGGAGAGAATTCTACCCGGAGCAGGCCTTCCTTCCAGAGCTCCACATCCCACAAT
IRF4a-ORF6_F_B05_2013-07-22_Mult(1>905) GTGGCTGGAGAGAATTCTACCCGGAGCAGGCCTTCCTTCCAGAGCTCCACATCCCACAAT
IRF4a-ORF2_F_F03_2013-07-22_Mult(1>942) GTGGCTGGAGAGAATTCTACCCGGAGCAGGCCTTCCTTCCAGAGCTCCACATCCCACAAT
IRF4a-ORF3_F_G03_2013-07-22_Mult(1>904) GTGGCTGGAGAGAATTCTACCCGGAGCAGGCCTTCCTTCCAGAGCTCCACATCCCACAAT
IRF4a-ORF4_F_H03_2013-07-22_Mul(47>949) GTGGCTGGAGAGAATTCTACCCGGAGCAGGCCTTCCTTCCAGAGCTCCACATCCCACAAT
IRF4a-ORF5_F_A05_2013-07-22_Mul(50>901) GTGGCTGGAGAGAATTCTACCCGGAGCAGGCCTTCCTTCCAGAGCTCCACATCCCACAAT
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) GTGGCTGGAGAGAATTCTACCCGGAGCAGGCCTTCCTTCCAGAGCTCCACATCCCACAAT
730 740 750 760 770 780
GTTCCTACCCCCCTCACCCATGGCAGGGCCCCCCCCATAGAGAACGCATACCAGATCAAGG
IRF4a-ORF6_F_B05_2013-07-22_Mult(1>905) GTTCCTACCCCCCTCACCCATGGCAGGGCCCCCCCCATAGAGAACGCATACCAGATCAAGG
IRF4a-ORF2_F_F03_2013-07-22_Mult(1>942) GTTCCTACCCCCCTCACCCATGGCAGGGCCCCCCCCATAGAGAACGCATACCAGATCAAGG
IRF4a-ORF3_F_G03_2013-07-22_Mult(1>904) GTTCCTACCCCCCTCACCCATGGCAGGGCCCCCCCCATAGAGAACGCATACCAGATCAAGG
IRF4a-ORF4_F_H03_2013-07-22_Mul(47>949) GTTCCTACCCCCCTCACCCATGGCAGGGCCCCCCCCATAGAGAACGCATACCAGATCAAGG
IRF4a-ORF5_F_A05_2013-07-22_Mul(50>901) GTTCCTACCCCCCTCACCCATGGCAGGGCCCCCCCCATAGAGAACGCATACCAGATCAAGG
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) GTTCCTACCCCCCTCACCCATGGCAGGGCCCCCCCCATAGAGAACGCATACCAGATCAAGG
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) ACCCATGGCAGGGCCCCCCCCATAGAGAACGCATACCAGATCAAGG
790 800 810 820 830 840
-GCTCCTTTTACTCGTACACGCATGCTGACGTACAGCCCTCCGCC-TTCACCCTT-GACC
IRF4a-ORF6_F_B05_2013-07-22_Mult(1>905) -GCTCCTTTTACTCGTACACGCATGCTGACGTACAGCCCTCCGCC-TTCACCCTT-GACC
IRF4a-ORF2_F_F03_2013-07-22_Mult(1>942) -GCTCCTTTTACTCGTACACGCATGCTGACGTACAGCCCTCCGCC-TTCACCCTT-GACC
IRF4a-ORF3_F_G03_2013-07-22_Mult(1>904) -GCTCCTTTTACTCGTACACGCATGCTGACGTACAGCCCTCCGCC-TTCACCCTT-GACC
IRF4a-ORF4_F_H03_2013-07-22_Mul(47>949) -GCTCCTTTTACTCGTACACGCATGCTGACGTACAGCCCTCCGCC-TTCACCCTT-GACC
IRF4a-ORF5_F_A05_2013-07-22_Mul(50>901) -GCTCCTTTTACTCGTACACGCATGCTGACGTACAGCCCTCCGCC-TTCACCCTT-GACC
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) -GCTCCTTTTACTCGTACACGCATGCTGACGTACAGCCCTCCGCC-TTCACCCTT-GACC
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) AGCTCCTTTTACTCGTACACGCATGCTGACGTACAGCCCTCCGCC-TTCACCCTT-GACC
IRF4a-ORF3_R_G04_2013-07-22_Mul(86>951) GCTCC-TTTACTCGTACACGCATGCTGACGTACAGCCCTCCGCC-TTCACCCTT-GACC
IRF4a-ORF4_R_H04_2013-07-22_Mul(76>926) GCTCCTTTTACTCGTACACGCATGCTGACGTACAGCCCTCCGCC-TTCACCCTT-GACC
IRF4a-ORF6_R_B06_2013-07-22_Mul(70>933) GCTCCTTTTACTCGTACACGCATGCTGACGTACAGCCCTCCGCC-TTCACCCTT-GACC
IRF4a-ORF2_R_F04_2013-07-22_Mul(67>882) TACACGCATGCTGACGTACAGCCCTCCGCCTTTCACCCTT-GACC
850 860 870 880 890 900
CCGGCATGAGACCAGCAGACCCTCTTTCTGACC-TTCGCCTGCATGTGTCCGTGTTCTCC
IRF4a-ORF6_F_B05_2013-07-22_Mult(1>905) CCGGCATGAGACCAGCAGACCCTCTTTCTGACC-TTCGCCTGCATGTGTCCGTGTTCTCC
IRF4a-ORF2_F_F03_2013-07-22_Mult(1>942) CCGGCATGAGACCAGCAGACCCTCTTTCTGACC-TTCGCCTGCATGTGTCCGTGTTCTCC
IRF4a-ORF3_F_G03_2013-07-22_Mult(1>904) CCGGCATGAGACCAGCAGACCCTCTTTCTGACC-TTCGCCTGCATGTGTCCGTGTTCTCC
IRF4a-ORF4_F_H03_2013-07-22_Mul(47>949) CCGGCATGAGACCAGCAGACCCTCTTTCTGACC-TTCGCCTGCATGTGTCCGTGTTCTCC
IRF4a-ORF5_F_A05_2013-07-22_Mul(50>901) CCGGCATGAGACCAGCAGACCCTCTTTCTGACC-TTCGCCTGCATGTGTCCGTGTTCTCC

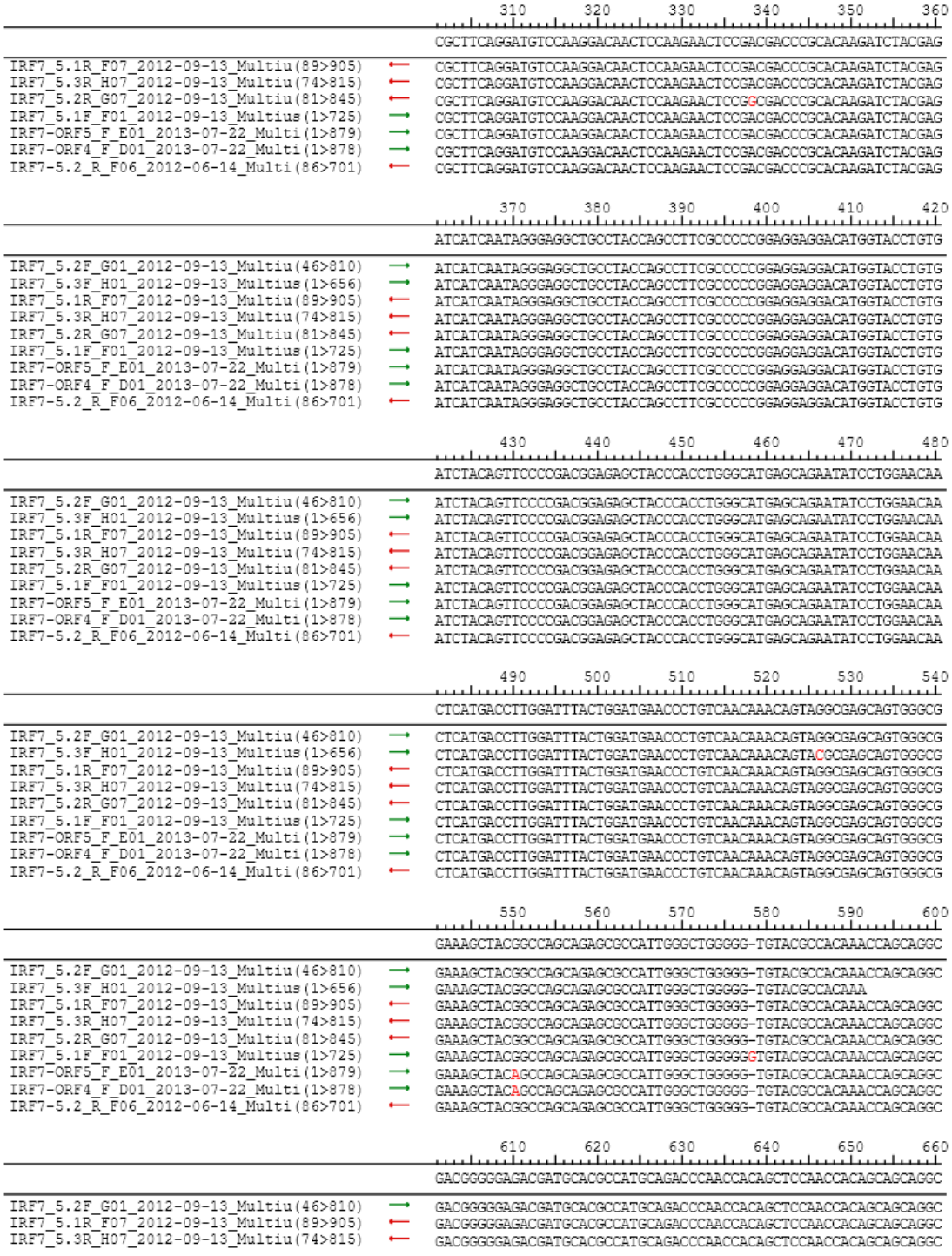

850 860 870 880 890 900
CCGGCATGAGACCAGCAGACCCTCTTTCTGACC-TTCGCCTGCATGTGTCCGTGTTCTCC
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) CCGGCATGAGACCAGCAGACCCTCTTTCTGACC-TTCGCCTGCATGTGTCCGTGTTCTCC
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) CCGGCATGAGACCAGCAGACCCTCTTTCTGACC-TTCGCCTGCATGTGTCCGTGTTCTCC
IRF4a-ORF3_R_G04_2013-07-22_Mul(86>951) CCGGCATGAGACCAGCAGACCCTCTTTCTGACC-TTCGCGTGCATGTGTCCGTGTTCTCC
IRF4a-ORF4_R_H04_2013-07-22_Mul(76>926) CCGGCATGAGACCAGCAGACCCTCTTTCTGACCTTTCGCCTGCATGTGTCCGTGTTCTCC
IRF4a-ORF6_R_B06_2013-07-22_Mul(70>933) CCGGCATGAGACCAGCAGACCCTCTTTCTGACC-TTCGCCTGCATGTGTCCGTGTTCTCC
IRF4a-ORF2_R_F04_2013-07-22_Mul(67>882) CCGGCATGAGACCAGCAGACCCTCTTTCTGACC-TTCGCCTGCATGTGTCCGTGTTCTCC
910 920 930 940 950 960
CGGGACGCTCTCGTGAGGGAGGTGACCATCTCCAACCCAAA-GGGCTGTCATCTGATCCC
IRF4a-ORF6_F_B05_2013-07-22_Mult(1>905) CGGGACGCTCTCGTGAGGGAGGTGACCATCTCCAACCCAAA-GGGCTGTCATCTGATCCC
IRF4a-ORF2_F_F03_2013-07-22_Mult(1>942) CGGGACGCTCTCGTGAGGGAGGTGACCATCTCCAACCCAAA-GGGCTGTCATCTGATCCC
IRF4a-ORF3_F_G03_2013-07-22_Mult(1>904) CGGGACGCTCTCGTGAGGGAGGTGACCATCTCCAACCCAAA-GGGCTGTCATCTGATCCC
IRF4a-ORF4_F_H03_2013-07-22_Mul(47>949) CGGGACGCTCTCGTGAGGGAGGTGACCATCTCCAACCCAAAAGGGCTGTCATCTGATCCC
IRF4a-ORF5_F_A05_2013-07-22_Mul(50>901) CGGGACGCTCTCGTGAGGGAGGTGACCATCTCCAACCCAAA-GGGCTGTCATCTGATCCC
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) CGGGACGCTCTCGTGAGGGAGGTGACCATCTCCAACCCAAA-GGGCTGTCATCTGATCCC
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) CGGGACGCTCTCGTGAGGGAGGTGACCATCTCCAACCCAAA-GGGCTGTCATCTGATCCC
IRF4a-ORF3_R_G04_2013-07-22_Mul(86>951) CGGGACGCTCTCGTGAGGGAGGTGACCATCTCCAACCCAAA-GGGCTGTCATCTGATCCC
IRF4a-ORF4_R_H04_2013-07-22_Mul(76>926) CGGGACGCTCTCGTGAGGGAGGTGACCATCTCCAACCCAAA-GGGCTGTCATCTGATCCC
IRF4a-ORF6_R_B06_2013-07-22_Mul(70>933) CGGGACGCTCTCGTGAGGGAGGTGACCATCTCCAACCCAAA-GGGCTGTCATCTGATCCC
IRF4a-ORF2_R_F04_2013-07-22_Mul(67>882) CGGGACGCTCTCGTGAGGGAGGTGACCATCTCCAACCCAAA-GGGCTGTCATCTGATCCC
970 980 990 1000 1010 1020
CTGGGCCCTGGAGGAAAAGGCCTACGTTTCCCCAGGGGCCCCGGGCCTGGTTCCCCTGCC
IRF4a-ORF6_F_B05_2013-07-22_Mult(1>905) CTGGGCCCTGGAGGAAAAGGCCTACGTTTCCCCAGGGGCCCCGGGCCTGGTT
IRF4a-ORF2_F_F03_2013-07-22_Mult(1>942) CTGGGCCCTGGAGGAAAAGGCCTACGTTTCCCCAGGGGCCCCGGGCCTGGTT
IRF4a-ORF3_F_G03_2013-07-22_Mult(1>904) CTGGGCCCTGGAGGAAAAGGCCTACGTTTCCCCAGGGGCCCCGGGCCTGGTTCCCCTGCC
IRF4a-ORF4_F_H03_2013-07-22_Mul(47>949) CTGGGCCCTGGAGGAAAAGGCCTACGTTTCCCCAGGGGCCCCGGGCCTGGTTCCCCTGCC
IRF4a-ORF5_F_A05_2013-07-22_Mul(50>901) CTGGGCCCTGGAGGAAAAGGCCTACGTTTCCCCAGGGGCCCCGGGCCTGGTTCCCCTGCC
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) CTGGGCCCTGGAGGAAAAGGCCTACGTTTCCCCAGGGGCCCCGGACCTGGTTCCCCTGCC
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) CTGGGCCCTGGAGGAAAAGGCCTACGTTTCCCCAGGGGCCCCGGGCCTGGTTCCCCTGCC
IRF4a-ORF3_R_G04_2013-07-22_Mul(86>951) CTGGGCCCTGGAGGAAAAGGCCTACGTTTCCCCAGGGGCCCCGGGCCTGGTTCCCCTGCC
IRF4a-ORF4_R_H04_2013-07-22_Mul(76>926) CTGGGCCCTGGAGGAAAAGGCCTACGTTTCCCCAGGGGCCCCGGGCCTGGTTCCCCTGCC
IRF4a-ORF6_R_B06_2013-07-22_Mul(70>933) CTGGGCCCTGGAGGAAAAGGCCTACGTTTCCCCAGGGGCCCCGGGCCTGGTTCCCCTGCC
IRF4a-ORF2_R_F04_2013-07-22_Mul(67>882) CTGGGCCCTGGAGGAAAAGGCCTACGTTTCCCCAGGGGCCCCGGGCCTGGTTCCCCTGCC
IRF4a-3.2R_B03_2013-03-28_Multiu(1>817) CCCCTGCC
IRF4a-3.1R_A03_2013-03-28_Multiu(1>812) CCCCTGCC
1030 1040 1050 1060 1070 1080
CCCGGAGGGCCTGACGCTCCAGAGGATGGCGGGGGAGGAGGGTCCCCCAAGCTCTCTGGC
IRF4a-ORF3_F_G03_2013-07-22_Mult(1>904) CCCGGAGGGCCTGACGCTCC
IRF4a-ORF4_F_H03_2013-07-22_Mul(47>949) CCCGGAGGGCCTGACGCTCCAGAGGATGGCGGGGGAGGAGGGTCCCCCAAGCTCTC
IRF4a-ORF5_F_A05_2013-07-22_Mul(50>901) CCCGGA
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) CCCGGAGGGCCTGACGCTCCAGAGGATGGCGGGGGAGGAGGGTCCCCCAAGCTCTCTGGC
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) CCCGGAGGGCCTGACGCTCCAGAGGATGGCGGGGGAGGAGGGTCCCCCAAGCTCTCTGGC
IRF4a-ORF3_R_G04_2013-07-22_Mul(86>951) CCCGGAGGGCCTGACGCTCCAGAGGATGGCGGGGGAGGAGGGTCCCCCAAGCTCTCTGGC
IRF4a-ORF4_R_H04_2013-07-22_Mul(76>926) CCCGGAGGGCCTGACGCTCCAGAGGATGGCGGGGGAGGAGGGTCCCCCAAGCTCTCTGGC
IRF4a-ORF6_R_B06_2013-07-22_Mul(70>933) CCCGGAGGGCCTGACGCTCCAGAGGATGGCGGGGGAGGAGGGTCCCCCAAGCTCTCTGGC
IRF4a-ORF2_R_F04_2013-07-22_Mul(67>882) CCCGGAGGGCCTGACGCTCCAGAGGATGGCGGGGGAGGAGGGTCCCCCAAGCTCTCTGGC
IRF4a-3.2R_B03_2013-03-28_Multiu(1>817) CCCGGAGGGCCTGACGCTCCAGAGGATGGCGGGGGAGGAGGGTCCCCCAAGCTCTCTGGC
IRF4a-3.1R_A03_2013-03-28_Multiu(1>812) CCCGGAGGGCCTGACGCTCCAGAGGATGGCGGGGGAGGAGGGTCCCCCAAGCTCTCTGGC
1090 1100 1110 1120 1130 1140
CATGCAGGGCGTGAGGCTGTGGATGACCCCAGAAGGCCTCTACGCCCGGCGGCAGTGCCA
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) CATGCAGGGCGTGAGGCTGTGGATGACCCCAGAAGGCCTCTACGCCCGGCGGCAGTGCCA
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) CATGCAGGGCGTGAGGCTGTGGATGACCCCAGAAGGCCTCTACGCCCGGCGGCAGTGCCA
IRF4a-ORF3_R_G04_2013-07-22_Mul(86>951) CATGCAGGGCGTGAGGCTGTGGATGACCCCAGAAGGCCTCTACGCCCGGCGGCAGTGCCA
IRF4a-ORF4_R_H04_2013-07-22_Mul(76>926) CATGCAGGGCGTGAGGCTGTGGATGACCCCAGAAGGCCTCTACGCCCGGCGGCAGTGCCA
IRF4a-ORF6_R_B06_2013-07-22_Mul(70>933) CATGCAGGGCGTGAGGCTGTGGATGACCCCAGAAGGCCTCTACGCCCGGCGGCAGTGCCA
IRF4a-ORF2_R_F04_2013-07-22_Mul(67>882) CATGCAGGGCGTGAGGCTGTGGATGACCCCAGAAGGCCTCTACGCCCGGCGGCAGTGCCA
IRF4a-3.2R_B03_2013-03-28_Multiu(1>817) CATGCAGGGCGTGAGGCTGTGGATGACCCCAGAAGGCCTCTACGCCCGGCGGCAGTGCCA
IRF4a-3.1R_A03_2013-03-28_Multiu(1>812) CATGCAGGGCGTGAGGCTGTGGATGACCCCAGAAGGCCTCTACGCCCGGCGGCAGTGCCA
IRF4a-3.3R_C03_2013-03-28_Multiu(1>703) CATGCAGGGCGTGAGGCTGTGGATGACCCCAGAAGGCCTGTACGCCCGGCGGCAGTGCCA

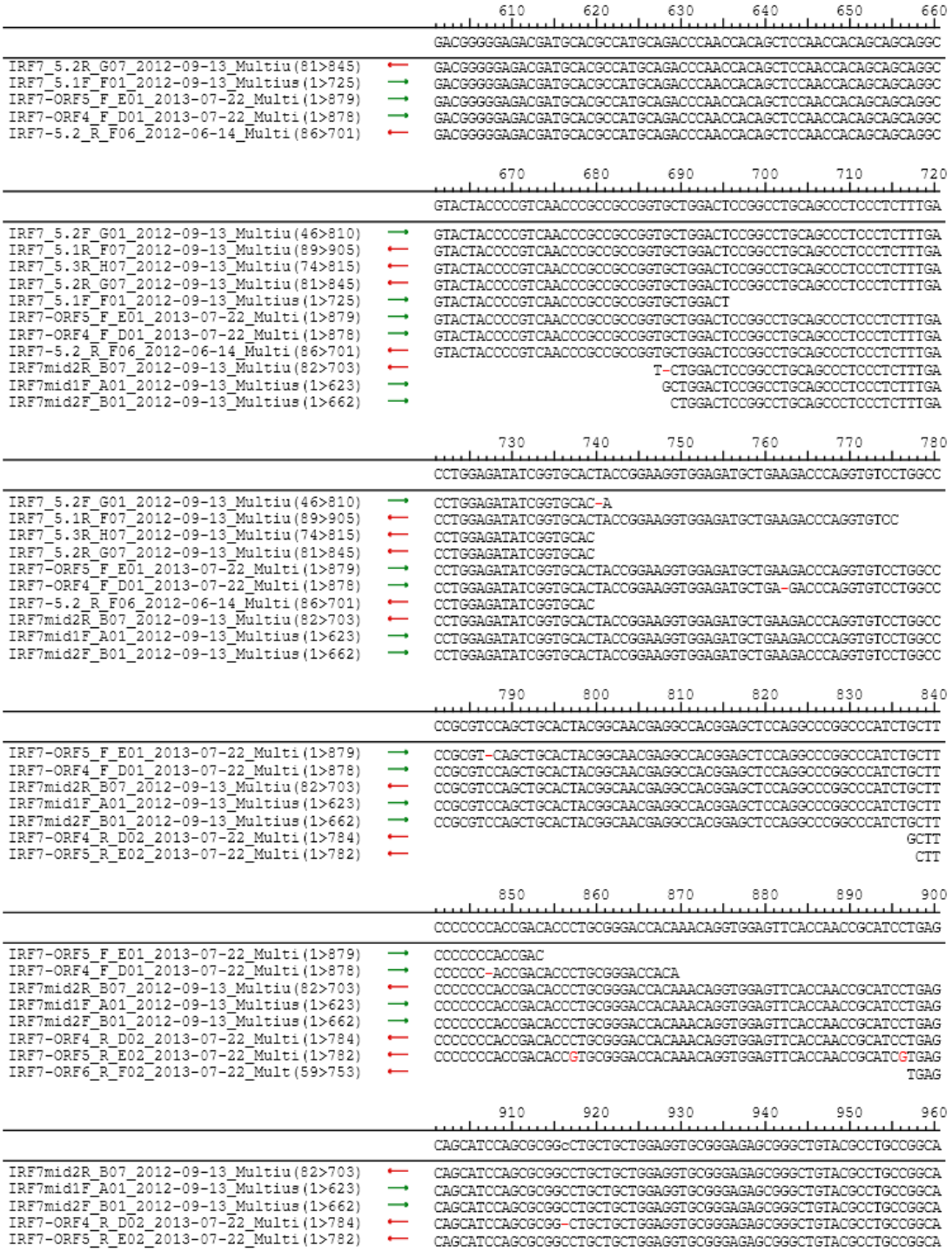

1150 1160 1170 1180 1190 1200
GGAGAGTGTGTACTGGAAGGAGGGGGTATCCCCTTACAAGGACAAACTCAACGAGATGGA
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) GGAGAGTGTGTACTGGAAGGAGGGGGTATCCCCTTACAAGGACAAACTCAACGAGATGGA
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) GGAGAGTGTGTACTGGAAGGAGGGGGTATCCCCTTACAAGGACAAACTCAACGAGATGGA
IRF4a-ORF3_R_G04_2013-07-22_Mul(86>951) GGAGAGTGTGTACTGGAAGGAGGGGGTATCCCCTTACAAGGACAAACTCAACGAGATGGA
IRF4a-ORF4_R_H04_2013-07-22_Mul(76>926) GGAGAGTGTGTACTGGAAGGAGGGGGTATCCCCTTACAAGGACAAACTCAACGAGATGGA
IRF4a-ORF6_R_B06_2013-07-22_Mul(70>933) GGAGAGTGTGTACTGGAAGGAGGGGGTATCCCCTTACAAGGACAAACTCAACGAGATGGA
IRF4a-ORF2_R_F04_2013-07-22_Mul(67>882) GGAGAGTGTGTACTGGAAGGAGAGGGTATCCCCTTACAAGGACAAACTCAACGAGATGGA
IRF4a-3.2R_B03_2013-03-28_Multiu(1>817) GGAGAGTGTGTACTGGAAGGAGGGGGTATCCCCTTACAAGGACAAACTCAACGAGATGGA
IRF4a-3.1R_A03_2013-03-28_Multiu(1>812) GGAGAGTGTGTACTGGAAGGAGGGGGTATCCCCTTACAAGGACAAACTCAACGAGATGGA
IRF4a-3.3R_C03_2013-03-28_Multiu(1>703) GGAGAGTGTGTACTGGAAGGAGGGGGTATCCCCTTACAAGGACAAACTCAACGAGATGGA
1210 1220 1230 1240 1250 1260
GAGAGAGGTCAACTGCAAAGTGCTTGACACCCAGGACTTCCTCACAGAAATCCAAAGTTA
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) GAGAGAGGTCAACTGCAAAGTGCTTGACACCCAGGACTTCCTCACAGAAATCCAAAGTTA
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) GAGAGAGGTCAACTGCAAAGTGCTTGACACCCAGGACTTCCTCACAGAAATCCAAAGTTA
IRF4a-ORF3_R_G04_2013-07-22_Mul(86>951) GAGAGAGGTCAACTGCAAAGTGCTTGACACCCAGGACTTCCTCACAGAAATCCAAAGTTA
IRF4a-ORF4_R_H04_2013-07-22_Mul(76>926) GAGAGAGGTCAACTGCAAAGTGCTTGACACCCAGGACTTCCTCACAGAAATCCAAAGTTA
IRF4a-ORF6_R_B06_2013-07-22_Mul(70>933) GAGAGAGGTCAACTGCAAAGTGCTTGACACCCAGGACTTCCTCACAGAAATCCAAAGTTA
IRF4a-ORF2_R_F04_2013-07-22_Mul(67>882) GAGAGAGGTCAACTGCAAAGTGCTTGACACCCAGGACTTCCTCACAGAAATCCAAAGTTA
IRF4a-3.2R_B03_2013-03-28_Multiu(1>817) GAGAGAGGTCAACTGCAAAGTGCTTGACACCCAGGACTTCCTCACAGAAATCCAAAGTTA
IRF4a-3.1R_A03_2013-03-28_Multiu(1>812) GAGAGAGGTCAACTGCAAAGTGCTTGACACCCAGGACTTCCTCACAGAAATCCAAAGTTA
IRF4a-3.3R_C03_2013-03-28_Multiu(1>703) GAGAGAGGTCAAATGCAAAGTGCTTGACACCCAGGACTTCCTCACAGAAATCCAAAGTTA
1270 1280 1290 1300 1310 1320
TGGGCTCCATGGCCGGCCCATACCTCCTTTCCAGGCCTTGCTGTGTTTTGGGGACGAGTG
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) TGGGCTCCATGGCCGGCCCATACCTCCTTTCCAGGCCTTGCTGTGTTTTGGGGACGAGTG
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) TGGGCTCCATGGCCGGCCCATACCTCCTTTCCAGGCCTTGCTGTGTTTTGGGGACGAGTG
IRF4a-ORF3_R_G04_2013-07-22_Mul(86>951) TGGGCTCCATGGCCGGCCCATACCTCCTTTCCAGGCCTTGCTGTGTTTTGGGGACGAGTG
IRF4a-ORF4_R_H04_2013-07-22_Mul(76>926) TGGGCTCCATGGCCGGCCCATACCTCCTTTCCAGGCCTTGCTGTGTTTTGGGGACGAGTG
IRF4a-ORF6_R_B06_2013-07-22_Mul(70>933) TGGGCTCCATGGCCGGCCCATACCTCCTTTCCAGGCCTTGCTGTGTTTTGGGGACGAGTG
IRF4a-ORF2_R_F04_2013-07-22_Mul(67>882) TGGGCTCCATGGCCGGCCCATACCTCCTTTCCAGGCCTTGCTGTGTTTTGGGGACGAGTG
IRF4a-3.2R_B03_2013-03-28_Multiu(1>817) TGGGCTCCATGGCCGGCCCATACCTCCTTTCCAGGCCTTGCTGTGTTTTGGGGACGAGTG
IRF4a-3.1R_A03_2013-03-28_Multiu(1>812) TGGGCTCCATGGCCGGCCCATACCTCCTTTCCAGGCCTTGCTGTGTTTTGGGGACGAGTG
IRF4a-3.3R_C03_2013-03-28_Multiu(1>703) TGGGCTCCATGGCCGGCCCATACCTCCTTTCCAGGCCTTGCTGTGTTTTGGGGACGAGTG
1330 1340 1350 1360 1370 1380
CGTCGACACAGAGAGACCAAGAAGGAGCCTCACCGTGCAGGTGGAACCCTTGTTTGCAAG
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) CGTCGACACAGAGAGACCAAGAAGGAGCCTCACCGTGCAGGTGGAACCCTTGTTTGCAAG
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) CGTCGACACAGAGAGACCAAGAAGGAGCCTCACCGTGCAGGTGGAACCCTTGTTTGCAAG
IRF4a-ORF3_R_G04_2013-07-22_Mul(86>951) CGTCGACACAGAGAGACCAAGAAGGAGCCTCACCGTGCAGGTGGAACCCTTGTTTGCAAG
IRF4a-ORF4_R_H04_2013-07-22_Mul(76>926) CGTCGACACAGAGAGACCAAGAAGGAGCCTCACCGTGCAGGTGGAACCCTTGTTTGCAAG
IRF4a-ORF6_R_B06_2013-07-22_Mul(70>933) CGTCGACACAGAGAGACCAAGAAGGAGCCTCACCGTGCAGGTGGAACCCTTGTTTGCAAG
IRF4a-ORF2_R_F04_2013-07-22_Mul(67>882) CGTCGACACAGAGAGACCAAGAAGGAGCCTCACCGTGCAGGTGGAACCCTTGTTTGCAAG
IRF4a-3.2R_B03_2013-03-28_Multiu(1>817) TGTCGACACAGAGAGACCAAGAAGGAGCCTCACCGTGCAGGTGGAACCCTTGTTTGCAAG
IRF4a-3.1R_A03_2013-03-28_Multiu(1>812) TGTCGACACAGAGAGACCAAGAAGGAGCCTCACCGTGCAGGTGGAACCCTTGTTTGCAAG
IRF4a-3.3R_C03_2013-03-28_Multiu(1>703) TGTCGACACAGAGAGACCAAGAAGGAGCCTCACCGTGCAGGTGGAACCCTTGTTTGCAAG
1390 1400 1410 1420 1430 1440
GCAGCTGTTTTACTATGCCCAGCAAACCGGCGGACACTATTACCGTGGCTACGAGCACCA
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) GCAGCTGTTTTACTATGCCCAGCAAACCGGCGGACACTATTACCGTGGCTACGAGCACCA
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) GCAGCTGTTTTACTATGCCCAGCAAACCGGCGGACACTATTACCGTGGCTACGAGCACCA
IRF4a-ORF3_R_G04_2013-07-22_Mul(86>951) GCAGCTGTTTTACTATGCCCAGCAAACCGGCGGACACTATTACCGTGGCTACGAGCACCA
IRF4a-ORF4_R_H04_2013-07-22_Mul(76>926) GCAGCTGTTTTACTATGCCCAGCAAACCGGCGGACACTATTACCGTGGCTACGAGCACCA
IRF4a-ORF6_R_B06_2013-07-22_Mul(70>933) GCAGCTGTTTTACTATGCCCAGCAAACCGGCGGACACTATTACCGTGGCTACGAGCACCA
IRF4a-ORF2_R_F04_2013-07-22_Mul(67>882) GCAGCTGTTTTACTATGCCCAGCAAACCGGCGGACACTATTACCGTGGCTACGAGCACCA
IRF4a-3.2R_B03_2013-03-28_Multiu(1>817) GCAGCTGTTTTATTATGCCCAGCAAACCGGCGGACACTATTACCGTGGTTACGAGCACCA
IRF4a-3.1R_A03_2013-03-28_Multiu(1>812) GCAGCTGTTTTATTATGCCCAGCAAACCGGCGGACACTATTACCGTGGCTACGAGCACCA
IRF4a-3.3R_C03_2013-03-28_Multiu(1>703) GCAGTTGTTTTACTATGCCCAGCAAACCGGCGGACATTATTACCGTGGGTACGAGCACCA

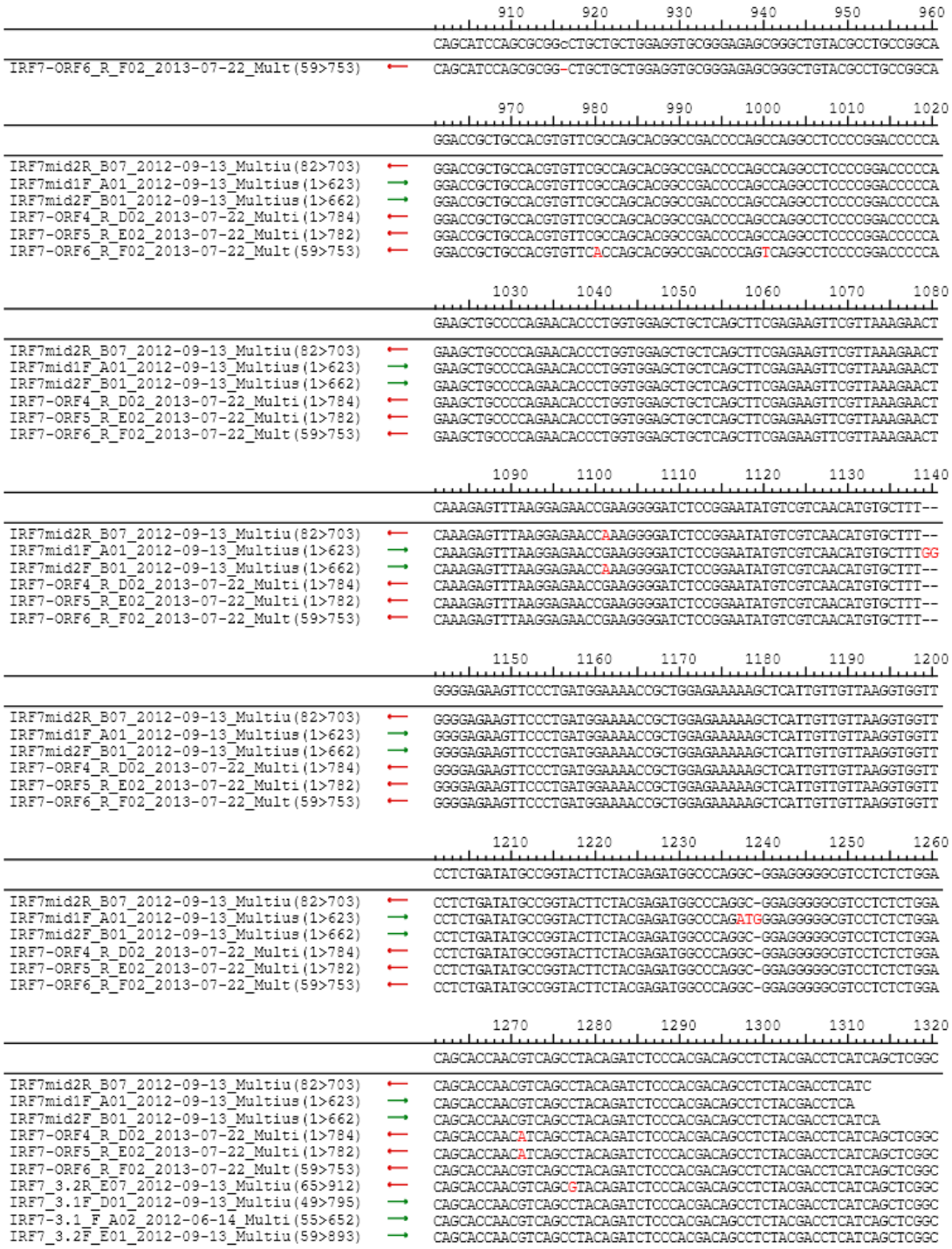

1450 1460 1470 1480 1490 1500
CGGTGTCCCAGAACACATAAGCCCTTCTGAGGACTATCAGCGGGCAATCTCACACCACCA
IRF4a-3.3F_C01_2013-03-28_Multi(25>879) CGGTGTCCCAGA-CACATAAGCCCTTCTGAGGACTATCAGCGGGCA-TCT
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) CGGTGTCCCAGAACACATAAGCCCTTCTGAGGACTATCAGCGGGCAATCTCACACCACCA
IRF4a-ORF3_R_G04_2013-07-22_Mul(86>951) CGGTGTCCCAGAACACATAAGCCCTTCTGAGGACTATCAGCGGGCAATCTCACACCACCA
IRF4a-ORF4_R_H04_2013-07-22_Mul(76>926) CGGTGTCCCAGAACACATAAGCCCTTCTGAGGACTATCAGCGGGCAATCTCACACCACCA
IRF4a-ORF6_R_B06_2013-07-22_Mul(70>933) CGGTGTCCCAGAACACATAAGCCCTTCTGAGGACTATCAGCGGGCAATCTCACACCACCA
IRF4a-ORF2_R_F04_2013-07-22_Mul(67>882) CGGTGTCCCAGAACACATAAGCCCTTCTGAGGACTATCAGCGGGCAATCTCACACCACCA
IRF4a-3.2R_B03_2013-03-28_Multiu(1>817) CGGTGTCCCAGAACACATAAGCCCTTTGAGGACTATCAGCGGGCAATTTCACACCACCA
IRF4a-3.1R_A03_2013-03-28_Multiu(1>812) CGGTGTCCCAGAACACATAAGCCCTTTGAGGACTATCAGCGGGCAATTTCACACCACCA
IRF4a-3.3R_C03_2013-03-28_Multiu(1>703) CGGTGTCCCAGAACACATAAGCCCTTCTGAGGACTATCAGCGGGCAATCTCCCCCACCA
1510 1520 1530 1540 1550 1560
TCACCACCACGGCAGTATGATGCAGGAGTGATCGCGACCGATGTTGGATACTTGGATGGA
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) TCACCACCACGGCAGTATGATGCAGGAGTGATCGCGACCGATGTTGGATACTTGGATGGA
IRF4a-ORF3_R_G04_2013-07-22_Mul(86>951) TCACCACCACGGCAGTATGATGCAGGAGTGATCGCGACCGATGTTGGATACTTGGATGGA
IRF4a-ORF4_R_H04_2013-07-22_Mul(76>926) TCACCACCACGGCAGTATGATGCAGGAGTGATCGCGACCGATGTTGGATACTTGGATGGA
IRF4a-ORF6_R_B06_2013-07-22_Mul(70>933) TCACCACCACGGCAGTATGATGCAGGAGTGATCGCGACCGATGTTGGATACTTGGATGGA
IRF4a-ORF2_R_F04_2013-07-22_Mul(67>882) TCACCACCACGGCAGTATGATGCAGGAGTGATCGCGACCGATGTTGGATACTTGGATGGA
IRF4a-3.2R_B03_2013-03-28_Multiu(1>817) TCACCACCACGGCAGTATGATGCAGGAGTGATCACGACCGATGTTGGATGCTTGGATGGA
IRF4a-3.1R_A03_2013-03-28_Multiu(1>812) TCACCACCACGGCAGTATGATGCAGGAGTGATCACGACCGATGTTGGATGCTTGGATGGA
IRF4a-3.3R_C03_2013-03-28_Multiu(1>703) TCACCACCCCGGCAGTATGATGCAGGAGTGATCGCGACCGATGTTGGATACTTGGATGGA
IRF4a-3.4R_D03_2013-03-28_Multiu(1>245) ACCGATGTTGGATACTTGGATGGA
1570 1580 1590 1600 1610 1620
TAAAGGAGCAGACACGTTTGGATGTCACTAACATGAAATCCCGATTGGTTGAGCTGCAAA
IRF4a-ORF5_R_A06_2013-07-22_Mul(76>950) TAAAGGAGCAGACACGTTTGGATGTCACTAACATGAAATCCCGATTGGTTGAGC
IRF4a-ORF3_R_G04_2013-07-22_Mul(86>951) TAAAGGAGCAGACACGTTTGGATGTCACTAACATGAAATCCCGATTGGTTGAGC
IRF4a-ORF4_R_H04_2013-07-22_Mul(76>926) TAAAGGAGCAGACACGTTTGGATGTCACTAACATGAAATCCCGATTGGTTGAGC
IRF4a-ORF6_R_B06_2013-07-22_Mul(70>933) TAAAGGAGCAGACACGTTTGGATGTCACTAACATGAAATCCCGATTGGTTGAGC
IRF4a-ORF2_R_F04_2013-07-22_Mul(67>882) TAAAGGAGCAGACACGTTTGGATGTCACTAACATGAAATCCCGATTGGTTGAGC
IRF4a-3.2R_B03_2013-03-28_Multiu(1>817) TAAAGGAGCAGACACGTTTGGATGTCACTAACATGAAATCCCGATTGGTTGAGCTGCAAA
IRF4a-3.1R_A03_2013-03-28_Multiu(1>812) TAAAGGAGCAGACACGTTTGGATGTCACTAACATGAAATCCCGATTGGTTGAGCTGCAAA
IRF4a-3.3R_C03_2013-03-28_Multiu(1>703) TAAAGGAGCAGACACGTTTGGATGTCACTAACAAGAAATCCCGATTGGTTGAGATGCAAA
IRF4a-3.4R_D03_2013-03-28_Multiu(1>245) TAAAGGAGCAGACACGTTTGGATGTCACTAACAAGAAATCCCGATTGGTTGAGATGCAAA
1630 1640 1650 1660 1670 1680
ATCGGTGAAACTGTTGGTAACAGCGAAATCAACAAACATGGACCAAACATCTTGGACTAA
IRF4a-3.2R_B03_2013-03-28_Multiu(1>817) ATCGGTGAAACTGTTGGTAACAGCGAAATCAACAAACATGGACCAAACATCTTGGACTAA
IRF4a-3.1R_A03_2013-03-28_Multiu(1>812) ATCGGTGAAACTGTTGGTAACAGCGAAATCAACAAACATGGACCAAACATCTTGGACTAA
IRF4a-3.3R_C03_2013-03-28_Multiu(1>703) ATCGGTGAAACTGTTGGTAACAGCGAAATCAACAAACATGGACCAAACATCTTGGAATAA
IRF4a-3.4R_D03_2013-03-28_Multiu(1>245) ATCGGTGAAACTGTTGGTAACAGCGAAATCAACAAACATGGACCAAACATCTTGGACTAA
1690 1700 1710 1720
AAACAGTTATGTTAAATATGTAAAAAAAAAAAAAAAAAAAAAAAAA
IRF4a-3.2R_B03_2013-03-28_Multiu(1>817) AAACAGTTATGTTAAATATGTAAAAAAAAAAAAAAAAAAAAAAAAAAA
IRF4a-3.1R_A03_2013-03-28_Multiu(1>812) AAACAGTTATGTTAAATATGTAAAAAAAAAAAAAAAAAAAAAAAAAAA
IRF4a-3.3R_C03_2013-03-28_Multiu(1>703) AAACAGTTATGTTAAATATGTAAAAAAAAAAAAAAAAAAAA
IRF4a-3.4R_D03_2013-03-28_Multiu(1>245) AAACAGTTATGTTAAATATGTAAAAAAAAAAAAAAAAAAAAAA

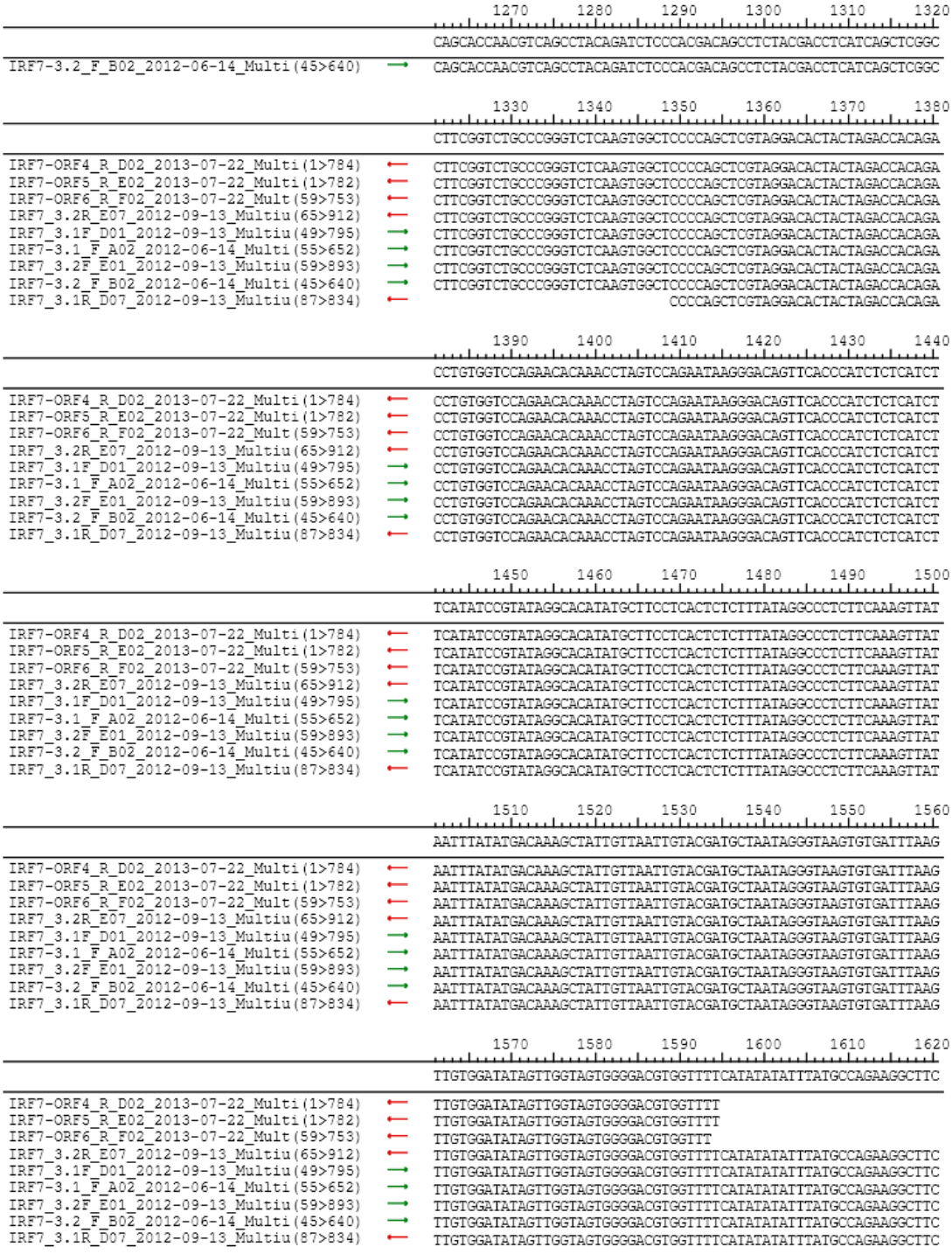
Appendix 4: Assembly of Atlantic cod *Irf7* RACE and ORF PCR sequencing reads. Sequencing methods are described in section 2.1.2. Sequence data was assembled using Lasergene SeqMan Pro software (DNASTAR). Consensus sequence is indicated between horizontal lines.

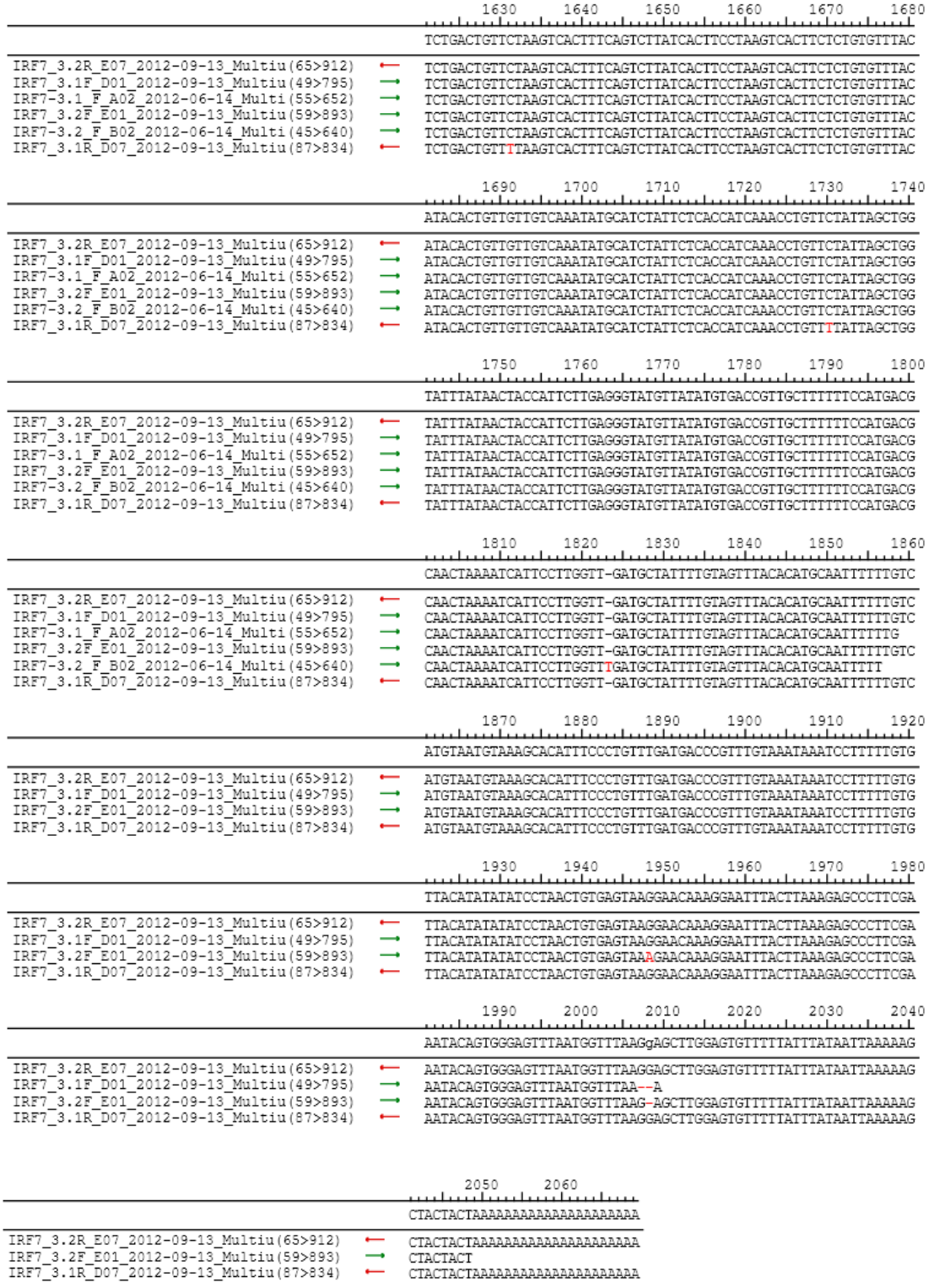

10 20 30 40 50 60
GAAAACTTCGTCCGGGACGACACAACGAGGTACACTGCAAACATGCAAAGCAGTCACAAG
IRF7_5.2F_G01_2012-09-13_Multiu(46>810)
IRF7_5.3F_H01_2012-09-13_Multius(1>656)
IRF7_5.1R_F07_2012-09-13_Multiu(89>905)
IRF7_5.3R_H07_2012-09-13_Multiu(74>815)
IRF7_5.2R_G07_2012-09-13_Multiu(81>845)
IRF7_5.1F_F01_2012-09-13_Multius(1>725)
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879)
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878)
70 80 90 100 110 120
CCGCTGTTCGCCAACTGGCTAATCGAGCAAGTGGAAACTGGGAACTATCCAGGTTTGTCC
IRF7_5.2F_G01_2012-09-13_Multiu(46>810)
IRF7_5.3F_H01_2012-09-13_Multius(1>656)
IRF7_5.1R_F07_2012-09-13_Multiu(89>905)
IRF7_5.3R_H07_2012-09-13_Multiu(74>815)
IRF7_5.2R_G07_2012-09-13_Multiu(81>845)
IRF7_5.1F_F01_2012-09-13_Multius(1>725)
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879)
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878)
130 140 150 160 170 180
TACATCAGCACGAATCTATTCAGAGTCCCCTGGAAACACAACTCCCGAAAGGACTGCAAC
IRF7_5.2F_G01_2012-09-13_Multiu(46>810)
IRF7_5.3F_H01_2012-09-13_Multius(1>656)
IRF7_5.1R_F07_2012-09-13_Multiu(89>905)
IRF7_5.3R_H07_2012-09-13_Multiu(74>815)
IRF7_5.2R_G07_2012-09-13_Multiu(81>845)
IRF7_5.1F_F01_2012-09-13_Multius(1>725)
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879)
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878)
IRF7-5.2_R_F06_2012-06-14_Multi(86>701)
190 200 210 220 230 240
GACGAGGACTGTAAAATATTTCGGGCATGGGCCGTCGCCAGTGGTAAAATCCACGAGTTT
IRF7_5.2F_G01_2012-09-13_Multiu(46>810)
IRF7_5.3F_H01_2012-09-13_Multius(1>656)
IRF7_5.1R_F07_2012-09-13_Multiu(89>905)
IRF7_5.3R_H07_2012-09-13_Multiu(74>815)
IRF7_5.2R_G07_2012-09-13_Multiu(81>845)
IRF7_5.1F_F01_2012-09-13_Multius(1>725)
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879)
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878)
IRF7-5.2_R_F06_2012-06-14_Multi(86>701)
250 260 270 280 290 300
CCAAACGACAAGGCCAAATGGAAGACCAACTTCCGCTGCGCTCTGAAGAACCTCAACAAA
IRF7_5.2F_G01_2012-09-13_Multiu(46>810)
IRF7_5.3F_H01_2012-09-13_Multius(1>656)
IRF7_5.1R_F07_2012-09-13_Multiu(89>905)
IRF7_5.3R_H07_2012-09-13_Multiu(74>815)
IRF7_5.2R_G07_2012-09-13_Multiu(81>845)
IRF7_5.1F_F01_2012-09-13_Multius(1>725)
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879)
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878)
IRF7-5.2_R_F06_2012-06-14_Multi(86>701)
310 320 330 340 350 360
CGCTTCAGGATGTCCAAGGACAACTCCAAGAACTCCGACGACCCGCACAAGATCTACGAG
IRF7_5.2F_G01_2012-09-13_Multiu(46>810)
IRF7_5.3F_H01_2012-09-13_Multius(1>656)


310 320 330 340 350 360
CGCTTCAGGATGTCCAAGGACAACTCCAAGAACTCCGACGACCCGCACAAGATCTACGAG
IRF7_5.1R_F07_2012-09-13_Multiu(89>905) CGCTTCAGGATGTCCAAGGACAACTCCAAGAACTCCGACGACCCGCACAAGATCTACGAG
IRF7_5.3R_H07_2012-09-13_Multiu(74>815) CGCTTCAGGATGTCCAAGGACAACTCCAAGAACTCCGACGACCCGCACAAGATCTACGAG
IRF7_5.2R_G07_2012-09-13_Multiu(81>845) CGCTTCAGGATGTCCAAGGACAACTCCAAGAACTCCGGCGACCCGCACAAGATCTACGAG
IRF7_5.1F_F01_2012-09-13_Multius(1>725) CGCTTCAGGATGTCCAAGGACAACTCCAAGAACTCCGACGACCCGCACAAGATCTACGAG
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879) CGCTTCAGGATGTCCAAGGACAACTCCAAGAACTCCGACGACCCGCACAAGATCTACGAG
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878) CGCTTCAGGATGTCCAAGGACAACTCCAAGAACTCCGACGACCCGCACAAGATCTACGAG
IRF7-5.2_R_F06_2012-06-14_Multi(86>701) CGCTTCAGGATGTCCAAGGACAACTCCAAGAACTCCGACGACCCGCACAAGATCTACGAG
370 380 390 400 410 420
ATCATCAATAGGGAGGCTGCCTACCAGCCTTCGCCCCCGGAGGAGGACATGGTACCTGTG
IRF7_5.2F_G01_2012-09-13_Multiu(46>810) ATCATCAATAGGGAGGCTGCCTACCAGCCTTCGCCCCCGGAGGAGGACATGGTACCTGTG
IRF7_5.3F_H01_2012-09-13_Multius(1>656) ATCATCAATAGGGAGGCTGCCTACCAGCCTTCGCCCCCGGAGGAGGACATGGTACCTGTG
IRF7_5.1R_F07_2012-09-13_Multiu(89>905) ATCATCAATAGGGAGGCTGCCTACCAGCCTTCGCCCCCGGAGGAGGACATGGTACCTGTG
IRF7_5.3R_H07_2012-09-13_Multiu(74>815) ATCATCAATAGGGAGGCTGCCTACCAGCCTTCGCCCCCGGAGGAGGACATGGTACCTGTG
IRF7_5.2R_G07_2012-09-13_Multiu(81>845) ATCATCAATAGGGAGGCTGCCTACCAGCCTTCGCCCCCGGAGGAGGACATGGTACCTGTG
IRF7_5.1F_F01_2012-09-13_Multius(1>725) ATCATCAATAGGGAGGCTGCCTACCAGCCTTCGCCCCCGGAGGAGGACATGGTACCTGTG
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879) ATCATCAATAGGGAGGCTGCCTACCAGCCTTCGCCCCCGGAGGAGGACATGGTACCTGTG
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878) ATCATCAATAGGGAGGCTGCCTACCAGCCTTCGCCCCCGGAGGAGGACATGGTACCTGTG
IRF7-5.2_R_F06_2012-06-14_Multi(86>701) ATCATCAATAGGGAGGCTGCCTACCAGCCTTCGCCCCCGGAGGAGGACATGGTACCTGTG
430 440 450 460 470 480
ATCTACAGTTCCCCGACGGAGAGCTACCCACCTGGGCATGAGCAGAATATCCTGGAACAA
IRF7_5.2F_G01_2012-09-13_Multiu(46>810) ATCTACAGTTCCCCGACGGAGAGCTACCCACCTGGGCATGAGCAGAATATCCTGGAACAA
IRF7_5.3F_H01_2012-09-13_Multius(1>656) ATCTACAGTTCCCCGACGGAGAGCTACCCACCTGGGCATGAGCAGAATATCCTGGAACAA
IRF7_5.1R_F07_2012-09-13_Multiu(89>905) ATCTACAGTTCCCCGACGGAGAGCTACCCACCTGGGCATGAGCAGAATATCCTGGAACAA
IRF7_5.3R_H07_2012-09-13_Multiu(74>815) ATCTACAGTTCCCCGACGGAGAGCTACCCACCTGGGCATGAGCAGAATATCCTGGAACAA
IRF7_5.2R_G07_2012-09-13_Multiu(81>845) ATCTACAGTTCCCCGACGGAGAGCTACCCACCTGGGCATGAGCAGAATATCCTGGAACAA
IRF7_5.1F_F01_2012-09-13_Multius(1>725) ATCTACAGTTCCCCGACGGAGAGCTACCCACCTGGGCATGAGCAGAATATCCTGGAACAA
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879) ATCTACAGTTCCCCGACGGAGAGCTACCCACCTGGGCATGAGCAGAATATCCTGGAACAA
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878) ATCTACAGTTCCCCGACGGAGAGCTACCCACCTGGGCATGAGCAGAATATCCTGGAACAA
IRF7-5.2_R_F06_2012-06-14_Multi(86>701) ATCTACAGTTCCCCGACGGAGAGCTACCCACCTGGGCATGAGCAGAATATCCTGGAACAA
490 500 510 520 530 540
CTCATGACCTTGGATTTACTGGATGAACCCTGTCAACAAACAGTAGGCGAGCAGTGGGCG
IRF7_5.2F_G01_2012-09-13_Multiu(46>810) CTCATGACCTTGGATTTACTGGATGAACCCTGTCAACAAACAGTAGGCGAGCAGTGGGCG
IRF7_5.3F_H01_2012-09-13_Multius(1>656) CTCATGACCTTGGATTTACTGGATGAACCCTGTCAACAAACAGTACGCGAGCAGTGGGCG
IRF7_5.1R_F07_2012-09-13_Multiu(89>905) CTCATGACCTTGGATTTACTGGATGAACCCTGTCAACAAACAGTAGGCGAGCAGTGGGCG
IRF7_5.3R_H07_2012-09-13_Multiu(74>815) CTCATGACCTTGGATTTACTGGATGAACCCTGTCAACAAACAGTAGGCGAGCAGTGGGCG
IRF7_5.2R_G07_2012-09-13_Multiu(81>845) CTCATGACCTTGGATTTACTGGATGAACCCTGTCAACAAACAGTAGGCGAGCAGTGGGCG
IRF7_5.1F_F01_2012-09-13_Multius(1>725) CTCATGACCTTGGATTTACTGGATGAACCCTGTCAACAAACAGTAGGCGAGCAGTGGGCG
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879) CTCATGACCTTGGATTTACTGGATGAACCCTGTCAACAAACAGTAGGCGAGCAGTGGGCG
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878) CTCATGACCTTGGATTTACTGGATGAACCCTGTCAACAAACAGTAGGCGAGCAGTGGGCG
IRF7-5.2_R_F06_2012-06-14_Multi(86>701) CTCATGACCTTGGATTTACTGGATGAACCCTGTCAACAAACAGTAGGCGAGCAGTGGGCG
550 560 570 580 590 600
GAAAGCTACGGCCAGCAGAGCGCCATTGGGCTGGGGG-TGTACGCCACAAACCAGCAGGC
IRF7_5.2F_G01_2012-09-13_Multiu(46>810) GAAAGCTACGGCCAGCAGAGCGCCATTGGGCTGGGGG-TGTACGCCACAAACCAGCAGGC
IRF7_5.3F_H01_2012-09-13_Multius(1>656) GAAAGCTACGGCCAGCAGAGCGCCATTGGGCTGGGGG-TGTACGCCACAAA
IRF7_5.1R_F07_2012-09-13_Multiu(89>905) GAAAGCTACGGCCAGCAGAGCGCCATTGGGCTGGGGG-TGTACGCCACAAACCAGCAGGC
IRF7_5.3R_H07_2012-09-13_Multiu(74>815) GAAAGCTACGGCCAGCAGAGCGCCATTGGGCTGGGGG-TGTACGCCACAAACCAGCAGGC
IRF7_5.2R_G07_2012-09-13_Multiu(81>845) GAAAGCTACGGCCAGCAGAGCGCCATTGGGCTGGGGG-TGTACGCCACAAACCAGCAGGC
IRF7_5.1F_F01_2012-09-13_Multius(1>725) GAAAGCTACGGCCAGCAGAGCGCCATTGGGCTGGGGGGTGTACGCCACAAACCAGCAGGC
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879) GAAAGCTACAGCCAGCAGAGCGCCATTGGGCTGGGGG-TGTACGCCACAAACCAGCAGGC
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878) GAAAGCTACAGCCAGCAGAGCGCCATTGGGCTGGGGG-TGTACGCCACAAACCAGCAGGC
IRF7-5.2_R_F06_2012-06-14_Multi(86>701) GAAAGCTACGGCCAGCAGAGCGCCATTGGGCTGGGGG-TGTACGCCACAAACCAGCAGGC
610 620 630 640 650 660
GACGGGGGAGACGATGCACGCCATGCAGACCCAACCACAGCTCCAACCACAGCAGCAGGC
IRF7_5.2F_G01_2012-09-13_Multiu(46>810) GACGGGGGAGACGATGCACGCCATGCAGACCCAACCACAGCTCCAACCACAGCAGCAGGC
IRF7_5.1R_F07_2012-09-13_Multiu(89>905) GACGGGGGAGACGATGCACGCCATGCAGACCCAACCACAGCTCCAACCACAGCAGCAGGC
IRF7_5.3R_H07_2012-09-13_Multiu(74>815) GACGGGGGAGACGATGCACGCCATGCAGACCCAACCACAGCTCCAACCACAGCAGCAGGC

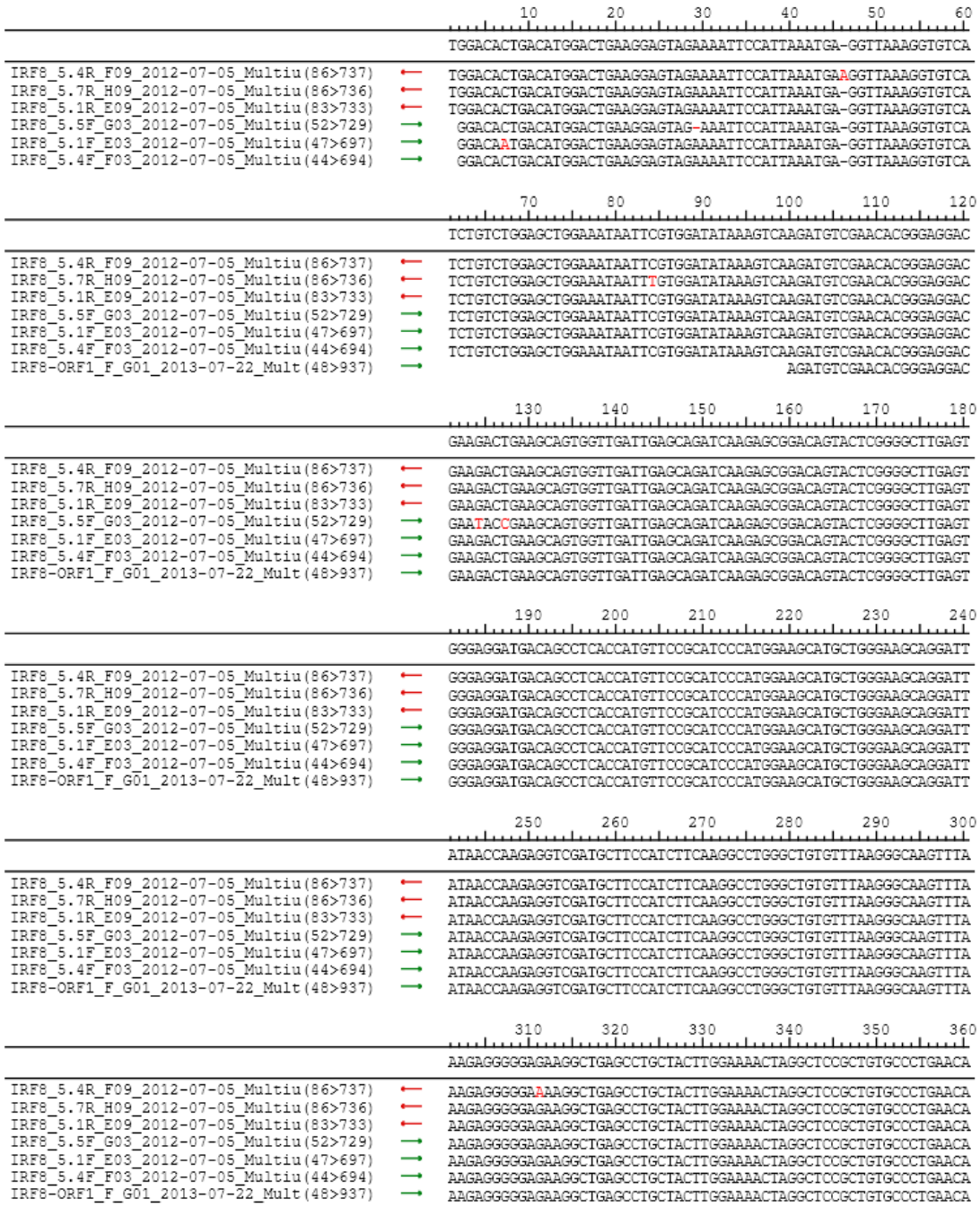

610 620 630 640 650 660
GACGGGGGAGACGATGCACGCCATGCAGACCCAACCACAGCTCCAACCACAGCAGCAGGC
IRF7_5.2R_G07_2012-09-13_Multiu(81>845)
IRF7_5.1F_F01_2012-09-13_Multius(1>725)
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879)
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878)
IRF7-5.2_R_F06_2012-06-14_Multi(86>701)
670 680 690 700 710 720
GTACTACCCCGTCAACCCGCCGCCGGTGCTGGACTCCGGCCTGCAGCCCTCCCTCTTTGA
IRF7_5.2F_G01_2012-09-13_Multiu(46>810)
IRF7_5.1R_F07_2012-09-13_Multiu(89>905)
IRF7_5.3R_H07_2012-09-13_Multiu(74>815)
IRF7_5.2R_G07_2012-09-13_Multiu(81>845)
IRF7_5.1F_F01_2012-09-13_Multius(1>725)
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879)
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878)
IRF7-5.2_R_F06_2012-06-14_Multi(86>701)
IRF7mid2R_B07_2012-09-13_Multiu(82>703)
IRF7mid1F_A01_2012-09-13_Multius(1>623)
IRF7mid2F_B01_2012-09-13_Multius(1>662)
730 740 750 760 770 780
CCTGGAGATATCGGTGCACTACCGGAAGGTGGAGATGCTGAAGACCCAGGTGTCCTGGCC
IRF7_5.2F_G01_2012-09-13_Multiu(46>810)
IRF7_5.1R_F07_2012-09-13_Multiu(89>905)
IRF7_5.3R_H07_2012-09-13_Multiu(74>815)
IRF7_5.2R_G07_2012-09-13_Multiu(81>845)
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879)
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878)
IRF7-5.2_R_F06_2012-06-14_Multi(86>701)
IRF7mid2R_B07_2012-09-13_Multiu(82>703)
IRF7mid1F_A01_2012-09-13_Multius(1>623)
IRF7mid2F_B01_2012-09-13_Multius(1>662)
790 800 810 820 830 840
CCGCGTCCAGCTGCACTACGGCAACGAGGCCACGGAGCTCCAGGCCCGGCCCATCTGCTT
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879)
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878)
IRF7mid2R_B07_2012-09-13_Multiu(82>703)
IRF7mid1F_A01_2012-09-13_Multius(1>623)
IRF7mid2F_B01_2012-09-13_Multius(1>662)
IRF7-ORF4_R_D02_2013-07-22_Multi(1>784)
IRF7-ORF5_R_E02_2013-07-22_Multi(1>782)
850 860 870 880 890 900
CCCCCCCACCGACACCCTGCGGGACCACAAACAGGTGGAGTTCACCAACCGCATCCTGAG
IRF7-ORF5_F_E01_2013-07-22_Multi(1>879)
IRF7-ORF4_F_D01_2013-07-22_Multi(1>878)
IRF7mid2R_B07_2012-09-13_Multiu(82>703)
IRF7mid1F_A01_2012-09-13_Multius(1>623)
IRF7mid2F_B01_2012-09-13_Multius(1>662)
IRF7-ORF4_R_D02_2013-07-22_Multi(1>784)
IRF7-ORF5_R_E02_2013-07-22_Multi(1>782)
IRF7-ORF6_R_F02_2013-07-22_Mult(59>753)
910 920 930 940 950 960
CAGCATCCAGCGCGGcCTGCTGCTGGAGGTGCGGGAGAGCGGGCTGTACGCCTGCCGGCA
IRF7mid2R_B07_2012-09-13_Multiu(82>703)
IRF7mid1F_A01_2012-09-13_Multius(1>623)
IRF7mid2F_B01_2012-09-13_Multius(1>662)
IRF7-ORF4_R_D02_2013-07-22_Multi(1>784)
IRF7-ORF5_R_E02_2013-07-22_Multi(1>782)

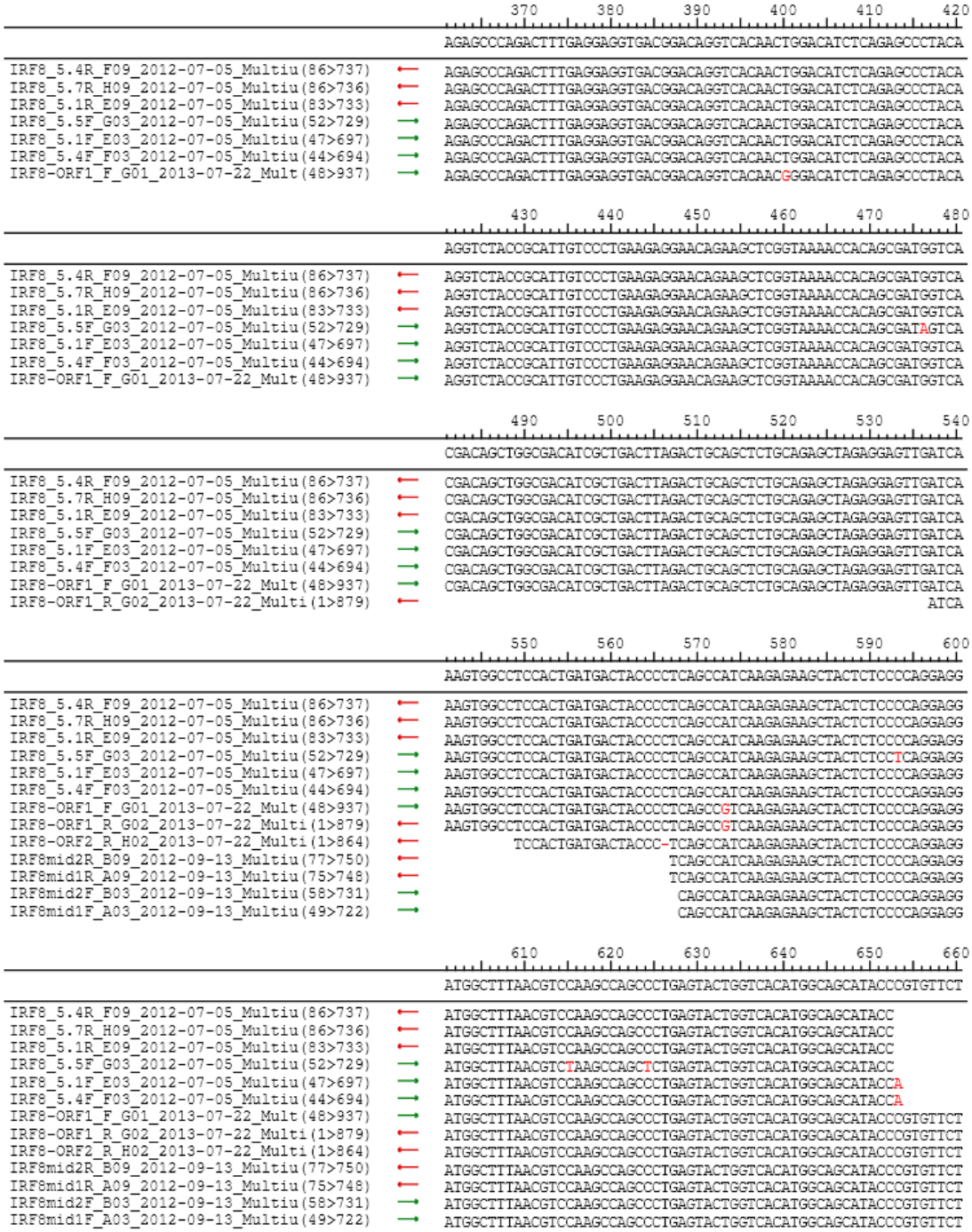

910 920 930 940 950 960
CAGCATCCAGCGCGGcCTGCTGCTGGAGGTGCGGGAGAGCGGGCTGTACGCCTGCCGGCA
IRF7-ORF6_R_F02_2013-07-22_Mult(59>753) CAGCATCCAGCGCGG-CTGCTGCTGGAGGTGCGGGAGAGCGGGCTGTACGCCTGCCGGCA
970 980 990 1000 1010 1020
GGACCGCTGCCACGTGTTCGCCAGCACGGCCGACCCCAGCCAGGCCTCCCCGGACCCCCA
IRF7mid2R_B07_2012-09-13_Multiu(82>703) GGACCGCTGCCACGTGTTCGCCAGCACGGCCGACCCCAGCCAGGCCTCCCCGGACCCCCA
IRF7mid1F_A01_2012-09-13_Multius(1>623) GGACCGCTGCCACGTGTTCGCCAGCACGGCCGACCCCAGCCAGGCCTCCCCGGACCCCCA
IRF7mid2F_B01_2012-09-13_Multius(1>662) GGACCGCTGCCACGTGTTCGCCAGCACGGCCGACCCCAGCCAGGCCTCCCCGGACCCCCA
IRF7-ORF4_R_D02_2013-07-22_Multi(1>784) GGACCGCTGCCACGTGTTCGCCAGCACGGCCGACCCCAGCCAGGCCTCCCCGGACCCCCA
IRF7-ORF5_R_E02_2013-07-22_Multi(1>782) GGACCGCTGCCACGTGTTCGCCAGCACGGCCGACCCCAGCCAGGCCTCCCCGGACCCCCA
IRF7-ORF6_R_F02_2013-07-22_Mult(59>753) GGACCGCTGCCACGTGTTCACCAGCACGGCCGACCCCAGTCAGGCCTCCCCGGACCCCCA
1030 1040 1050 1060 1070 1080
GAAGCTGCCCCAGAACACCCTGGTGGAGCTGCTCAGCTTCGAGAAGTTCGTTAAAGAACT
IRF7mid2R_B07_2012-09-13_Multiu(82>703) GAAGCTGCCCCAGAACACCCTGGTGGAGCTGCTCAGCTTCGAGAAGTTCGTTAAAGAACT
IRF7mid1F_A01_2012-09-13_Multius(1>623) GAAGCTGCCCCAGAACACCCTGGTGGAGCTGCTCAGCTTCGAGAAGTTCGTTAAAGAACT
IRF7mid2F_B01_2012-09-13_Multius(1>662) GAAGCTGCCCCAGAACACCCTGGTGGAGCTGCTCAGCTTCGAGAAGTTCGTTAAAGAACT
IRF7-ORF4_R_D02_2013-07-22_Multi(1>784) GAAGCTGCCCCAGAACACCCTGGTGGAGCTGCTCAGCTTCGAGAAGTTCGTTAAAGAACT
IRF7-ORF5_R_E02_2013-07-22_Multi(1>782) GAAGCTGCCCCAGAACACCCTGGTGGAGCTGCTCAGCTTCGAGAAGTTCGTTAAAGAACT
IRF7-ORF6_R_F02_2013-07-22_Mult(59>753) GAAGCTGCCCCAGAACACCCTGGTGGAGCTGCTCAGCTTCGAGAAGTTCGTTAAAGAACT
1090 1100 1110 1120 1130 1140
CAAAGAGTTTAAGGAGAACCGAAGGGGATCTCCGGAATATGTCGTCAACATGTGCTTT--
IRF7mid2R_B07_2012-09-13_Multiu(82>703) CAAAGAGTTTAAGGAGAACCAAAGGGGATCTCCGGAATATGTCGTCAACATGTGCTTT--
IRF7mid1F_A01_2012-09-13_Multius(1>623) CAAAGAGTTTAAGGAGAACCGAAGGGGATCTCCGGAATATGTCGTCAACATGTGCTTTGG
IRF7mid2F_B01_2012-09-13_Multius(1>662) CAAAGAGTTTAAGGAGAACCAAAGGGGATCTCCGGAATATGTCGTCAACATGTGCTTT--
IRF7-ORF4_R_D02_2013-07-22_Multi(1>784) CAAAGAGTTTAAGGAGAACCGAAGGGGATCTCCGGAATATGTCGTCAACATGTGCTTT--
IRF7-ORF5_R_E02_2013-07-22_Multi(1>782) CAAAGAGTTTAAGGAGAACCGAAGGGGATCTCCGGAATATGTCGTCAACATGTGCTTT--
IRF7-ORF6_R_F02_2013-07-22_Mult(59>753) CAAAGAGTTTAAGGAGAACCGAAGGGGATCTCCGGAATATGTCGTCAACATGTGCTTT--
1150 1160 1170 1180 1190 1200
GGGGAGAAGTTCCCTGATGGAAAACCGCTGGAGAAAAAGCTCATTGTTGTTAAGGTGGTT
IRF7mid2R_B07_2012-09-13_Multiu(82>703) GGGGAGAAGTTCCCTGATGGAAAACCGCTGGAGAAAAAGCTCATTGTTGTTAAGGTGGTT
IRF7mid1F_A01_2012-09-13_Multius(1>623) GGGGAGAAGTTCCCTGATGGAAAACCGCTGGAGAAAAAGCTCATTGTTGTTAAGGTGGTT
IRF7mid2F_B01_2012-09-13_Multius(1>662) GGGGAGAAGTTCCCTGATGGAAAACCGCTGGAGAAAAAGCTCATTGTTGTTAAGGTGGTT
IRF7-ORF4_R_D02_2013-07-22_Multi(1>784) GGGGAGAAGTTCCCTGATGGAAAACCGCTGGAGAAAAAGCTCATTGTTGTTAAGGTGGTT
IRF7-ORF5_R_E02_2013-07-22_Multi(1>782) GGGGAGAAGTTCCCTGATGGAAAACCGCTGGAGAAAAAGCTCATTGTTGTTAAGGTGGTT
IRF7-ORF6_R_F02_2013-07-22_Mult(59>753) GGGGAGAAGTTCCCTGATGGAAAACCGCTGGAGAAAAAGCTCATTGTTGTTAAGGTGGTT
1210 1220 1230 1240 1250 1260
CCTCTGATATGCCGGTACTTCTACGAGATGGCCCAGGC-GGAGGGGGCGTCCTCTCTGGA
IRF7mid2R_B07_2012-09-13_Multiu(82>703) CCTCTGATATGCCGGTACTTCTACGAGATGGCCCAGGC-GGAGGGGGCGTCCTCTCTGGA
IRF7mid1F_A01_2012-09-13_Multius(1>623) CCTCTGATATGCCGGTACTTCTACGAGATGGCCCAGATGGGAGGGGGCGTCCTCTCTGGA
IRF7mid2F_B01_2012-09-13_Multius(1>662) CCTCTGATATGCCGGTACTTCTACGAGATGGCCCAGGC-GGAGGGGGCGTCCTCTCTGGA
IRF7-ORF4_R_D02_2013-07-22_Multi(1>784) CCTCTGATATGCCGGTACTTCTACGAGATGGCCCAGGC-GGAGGGGGCGTCCTCTCTGGA
IRF7-ORF5_R_E02_2013-07-22_Multi(1>782) CCTCTGATATGCCGGTACTTCTACGAGATGGCCCAGGC-GGAGGGGGCGTCCTCTCTGGA
IRF7-ORF6_R_F02_2013-07-22_Mult(59>753) CCTCTGATATGCCGGTACTTCTACGAGATGGCCCAGGC-GGAGGGGGCGTCCTCTCTGGA
1270 1280 1290 1300 1310 1320
CAGCACCAACGTCAGCCTACAGATCTCCCACGACAGCCTCTACGACCTCATCAGCTCGGC
IRF7mid2R_B07_2012-09-13_Multiu(82>703) CAGCACCAACGTCAGCCTACAGATCTCCCACGACAGCCTCTACGACCTCATC
IRF7mid1F_A01_2012-09-13_Multius(1>623) CAGCACCAACGTCAGCCTACAGATCTCCCACGACAGCCTCTACGACCTCA
IRF7mid2F_B01_2012-09-13_Multius(1>662) CAGCACCAACGTCAGCCTACAGATCTCCCACGACAGCCTCTACGACCTCATCA
IRF7-ORF4_R_D02_2013-07-22_Multi(1>784) CAGCACCAACATCAGCCTACAGATCTCCCACGACAGCCTCTACGACCTCATCAGCTCGGC
IRF7-ORF5_R_E02_2013-07-22_Multi(1>782) CAGCACCAACATCAGCCTACAGATCTCCCACGACAGCCTCTACGACCTCATCAGCTCGGC
IRF7-ORF6_R_F02_2013-07-22_Mult(59>753) CAGCACCAACGTCAGCCTACAGATCTCCCACGACAGCCTCTACGACCTCATCAGCTCGGC
IRF7_3.2R_E07_2012-09-13_Multiu(65>912) CAGCACCAACGTCAGCGTACAGATCTCCCACGACAGCCTCTACGACCTCATCAGCTCGGC
IRF7_3.1F_D01_2012-09-13_Multiu(49>795) CAGCACCAACGTCAGCCTACAGATCTCCCACGACAGCCTCTACGACCTCATCAGCTCGGC
IRF7-3.1_F_A02_2012-06-14_Multi(55>652) CAGCACCAACGTCAGCCTACAGATCTCCCACGACAGCCTCTACGACCTCATCAGCTCGGC
IRF7_3.2F_E01_2012-09-13_Multiu(59>893) CAGCACCAACGTCAGCCTACAGATCTCCCACGACAGCCTCTACGACCTCATCAGCTCGGC

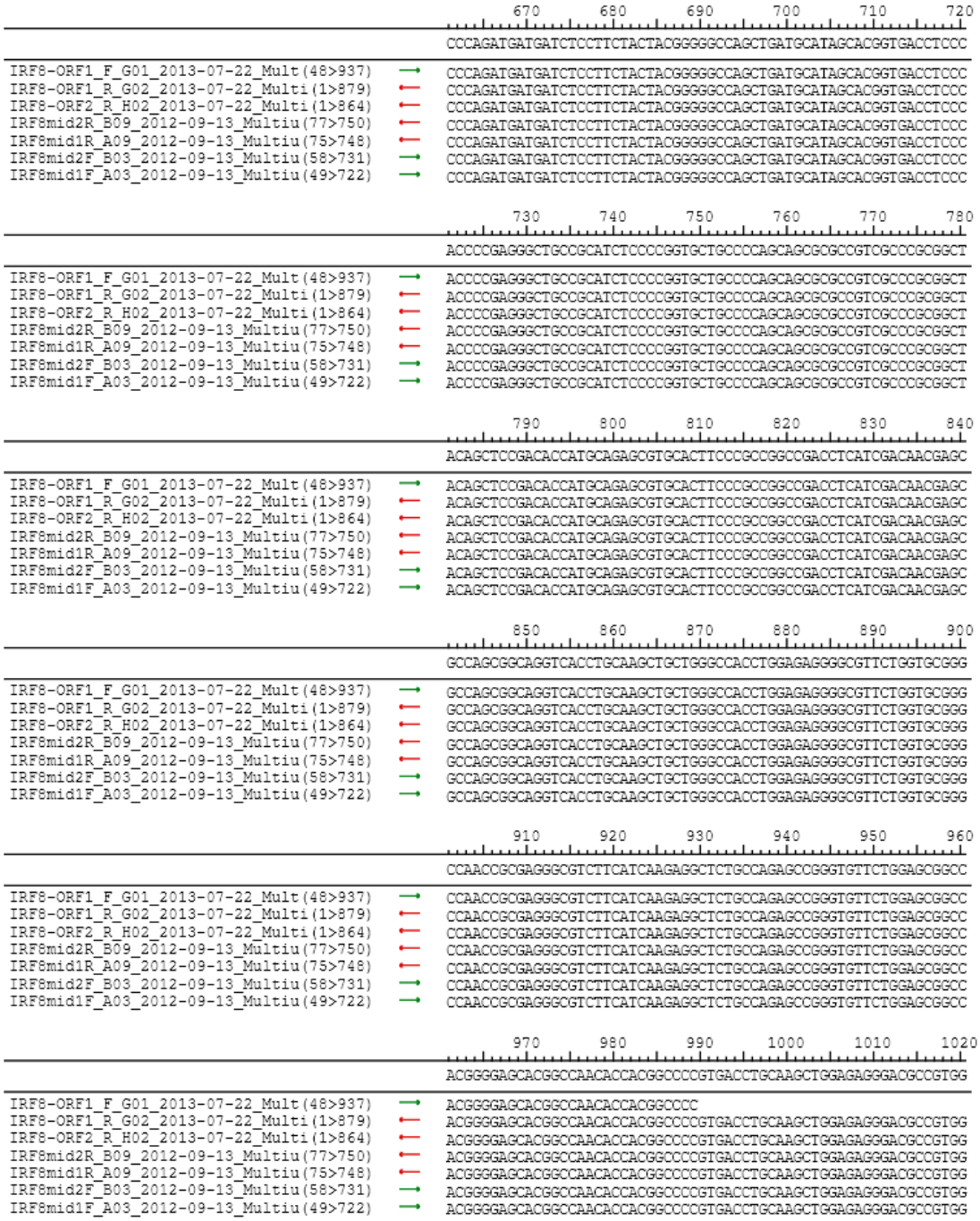

1270 1280 1290 1300 1310 1320
CAGCACCAACGTCAGCCTACAGATCTCCCACGACAGCCTCTACGACCTCATCAGCTCGGC
IRF7-3.2_F_B02_2012-06-14_Multi(45>640) CAGCACCAACGTCAGCCTACAGATCTCCCACGACAGCCTCTACGACCTCATCAGCTCGGC
1330 1340 1350 1360 1370 1380
CTTCGGTCTGCCCGGGTCTCAAGTGGCTCCCCAGCTCGTAGGACACTACTAGACCACAGA
IRF7-ORF4_R_D02_2013-07-22_Multi(1>784) CTTCGGTCTGCCCGGGTCTCAAGTGGCTCCCCAGCTCGTAGGACACTACTAGACCACAGA
IRF7-ORF5_R_E02_2013-07-22_Multi(1>782) CTTCGGTCTGCCCGGGTCTCAAGTGGCTCCCCAGCTCGTAGGACACTACTAGACCACAGA
IRF7-ORF6_R_F02_2013-07-22_Mult(59>753) CTTCGGTCTGCCCGGGTCTCAAGTGGCTCCCCAGCTCGTAGGACACTACTAGACCACAGA
IRF7_3.2R_E07_2012-09-13_Multiu(65>912) CTTCGGTCTGCCCGGGTCTCAAGTGGCTCCCCAGCTCGTAGGACACTACTAGACCACAGA
IRF7_3.1F_D01_2012-09-13_Multiu(49>795) CTTCGGTCTGCCCGGGTCTCAAGTGGCTCCCCAGCTCGTAGGACACTACTAGACCACAGA
IRF7-3.1_F_A02_2012-06-14_Multi(55>652) CTTCGGTCTGCCCGGGTCTCAAGTGGCTCCCCAGCTCGTAGGACACTACTAGACCACAGA
IRF7_3.2F_E01_2012-09-13_Multiu(59>893) CTTCGGTCTGCCCGGGTCTCAAGTGGCTCCCCAGCTCGTAGGACACTACTAGACCACAGA
IRF7-3.2_F_B02_2012-06-14_Multi(45>640) CTTCGGTCTGCCCGGGTCTCAAGTGGCTCCCCAGCTCGTAGGACACTACTAGACCACAGA
IRF7_3.1R_D07_2012-09-13_Multiu(87>834) CCCCAGCTCGTAGGACACTACTAGACCACAGA
1390 1400 1410 1420 1430 1440
CCTGTGGTCCAGAACACAAACCTAGTCCAGAATAAGGGACAGTTCACCCATCTCTCATCT
IRF7-ORF4_R_D02_2013-07-22_Multi(1>784) CCTGTGGTCCAGAACACAAACCTAGTCCAGAATAAGGGACAGTTCACCCATCTCTCATCT
IRF7-ORF5_R_E02_2013-07-22_Multi(1>782) CCTGTGGTCCAGAACACAAACCTAGTCCAGAATAAGGGACAGTTCACCCATCTCTCATCT
IRF7-ORF6_R_F02_2013-07-22_Mult(59>753) CCTGTGGTCCAGAACACAAACCTAGTCCAGAATAAGGGACAGTTCACCCATCTCTCATCT
IRF7_3.2R_E07_2012-09-13_Multiu(65>912) CCTGTGGTCCAGAACACAAACCTAGTCCAGAATAAGGGACAGTTCACCCATCTCTCATCT
IRF7_3.1F_D01_2012-09-13_Multiu(49>795) CCTGTGGTCCAGAACACAAACCTAGTCCAGAATAAGGGACAGTTCACCCATCTCTCATCT
IRF7-3.1_F_A02_2012-06-14_Multi(55>652) CCTGTGGTCCAGAACACAAACCTAGTCCAGAATAAGGGACAGTTCACCCATCTCTCATCT
IRF7_3.2F_E01_2012-09-13_Multiu(59>893) CCTGTGGTCCAGAACACAAACCTAGTCCAGAATAAGGGACAGTTCACCCATCTCTCATCT
IRF7-3.2_F_B02_2012-06-14_Multi(45>640) CCTGTGGTCCAGAACACAAACCTAGTCCAGAATAAGGGACAGTTCACCCATCTCTCATCT
IRF7_3.1R_D07_2012-09-13_Multiu(87>834) CCTGTGGTCCAGAACACAAACCTAGTCCAGAATAAGGGACAGTTCACCCATCTCTCATCT
1450 1460 1470 1480 1490 1500
TCATATCCGTATAGGCACATATGCTTCCTCACTCTCTTTATAGGCCCTCTTCAAAGTTAT
IRF7-ORF4_R_D02_2013-07-22_Multi(1>784) TCATATCCGTATAGGCACATATGCTTCCTCACTCTCTTTATAGGCCCTCTTCAAAGTTAT
IRF7-ORF5_R_E02_2013-07-22_Multi(1>782) TCATATCCGTATAGGCACATATGCTTCCTCACTCTCTTTATAGGCCCTCTTCAAAGTTAT
IRF7-ORF6_R_F02_2013-07-22_Mult(59>753) TCATATCCGTATAGGCACATATGCTTCCTCACTCTCTTTATAGGCCCTCTTCAAAGTTAT
IRF7_3.2R_E07_2012-09-13_Multiu(65>912) TCATATCCGTATAGGCACATATGCTTCCTCACTCTCTTTATAGGCCCTCTTCAAAGTTAT
IRF7_3.1F_D01_2012-09-13_Multiu(49>795) TCATATCCGTATAGGCACATATGCTTCCTCACTCTCTTTATAGGCCCTCTTCAAAGTTAT
IRF7-3.1_F_A02_2012-06-14_Multi(55>652) TCATATCCGTATAGGCACATATGCTTCCTCACTCTCTTTATAGGCCCTCTTCAAAGTTAT
IRF7_3.2F_E01_2012-09-13_Multiu(59>893) TCATATCCGTATAGGCACATATGCTTCCTCACTCTCTTTATAGGCCCTCTTCAAAGTTAT
IRF7-3.2_F_B02_2012-06-14_Multi(45>640) TCATATCCGTATAGGCACATATGCTTCCTCACTCTCTTTATAGGCCCTCTTCAAAGTTAT
IRF7_3.1R_D07_2012-09-13_Multiu(87>834) TCATATCCGTATAGGCACATATGCTTCCTCACTCTCTTTATAGGCCCTCTTCAAAGTTAT
1510 1520 1530 1540 1550 1560
AATTTATATGACAAAGCTATTGTTAATTGTACGATGCTAATAGGGTAAGTGTGATTTAAG
IRF7-ORF4_R_D02_2013-07-22_Multi(1>784) AATTTATATGACAAAGCTATTGTTAATTGTACGATGCTAATAGGGTAAGTGTGATTTAAG
IRF7-ORF5_R_E02_2013-07-22_Multi(1>782) AATTTATATGACAAAGCTATTGTTAATTGTACGATGCTAATAGGGTAAGTGTGATTTAAG
IRF7-ORF6_R_F02_2013-07-22_Mult(59>753) AATTTATATGACAAAGCTATTGTTAATTGTACGATGCTAATAGGGTAAGTGTGATTTAAG
IRF7_3.2R_E07_2012-09-13_Multiu(65>912) AATTTATATGACAAAGCTATTGTTAATTGTACGATGCTAATAGGGTAAGTGTGATTTAAG
IRF7_3.1F_D01_2012-09-13_Multiu(49>795) AATTTATATGACAAAGCTATTGTTAATTGTACGATGCTAATAGGGTAAGTGTGATTTAAG
IRF7-3.1_F_A02_2012-06-14_Multi(55>652) AATTTATATGACAAAGCTATTGTTAATTGTACGATGCTAATAGGGTAAGTGTGATTTAAG
IRF7_3.2F_E01_2012-09-13_Multiu(59>893) AATTTATATGACAAAGCTATTGTTAATTGTACGATGCTAATAGGGTAAGTGTGATTTAAG
IRF7-3.2_F_B02_2012-06-14_Multi(45>640) AATTTATATGACAAAGCTATTGTTAATTGTACGATGCTAATAGGGTAAGTGTGATTTAAG
IRF7_3.1R_D07_2012-09-13_Multiu(87>834) AATTTATATGACAAAGCTATTGTTAATTGTACGATGCTAATAGGGTAAGTGTGATTTAAG
1570 1580 1590 1600 1610 1620
TTGTGGATATAGTTGGTAGTGGGGACGTGGTTTTCATATATATTTATGCCAGAAGGCTTC
IRF7-ORF4_R_D02_2013-07-22_Multi(1>784) TTGTGGATATAGTTGGTAGTGGGGACGTGGTTTT
IRF7-ORF5_R_E02_2013-07-22_Multi(1>782) TTGTGGATATAGTTGGTAGTGGGGACGTGGTTTT
IRF7-ORF6_R_F02_2013-07-22_Mult(59>753) TTGTGGATATAGTTGGTAGTGGGGACGTGGTTT
IRF7_3.2R_E07_2012-09-13_Multiu(65>912) TTGTGGATATAGTTGGTAGTGGGGACGTGGTTTTCATATATATTTATGCCAGAAGGCTTC
IRF7_3.1F_D01_2012-09-13_Multiu(49>795) TTGTGGATATAGTTGGTAGTGGGGACGTGGTTTTCATATATATTTATGCCAGAAGGCTTC
IRF7-3.1_F_A02_2012-06-14_Multi(55>652) TTGTGGATATAGTTGGTAGTGGGGACGTGGTTTTCATATATATTTATGCCAGAAGGCTTC
IRF7_3.2F_E01_2012-09-13_Multiu(59>893) TTGTGGATATAGTTGGTAGTGGGGACGTGGTTTTCATATATATTTATGCCAGAAGGCTTC
IRF7-3.2_F_B02_2012-06-14_Multi(45>640) TTGTGGATATAGTTGGTAGTGGGGACGTGGTTTTCATATATATTTATGCCAGAAGGCTTC
IRF7_3.1R_D07_2012-09-13_Multiu(87>834) TTGTGGATATAGTTGGTAGTGGGGACGTGGTTTTCATATATATTTATGCCAGAAGGCTTC

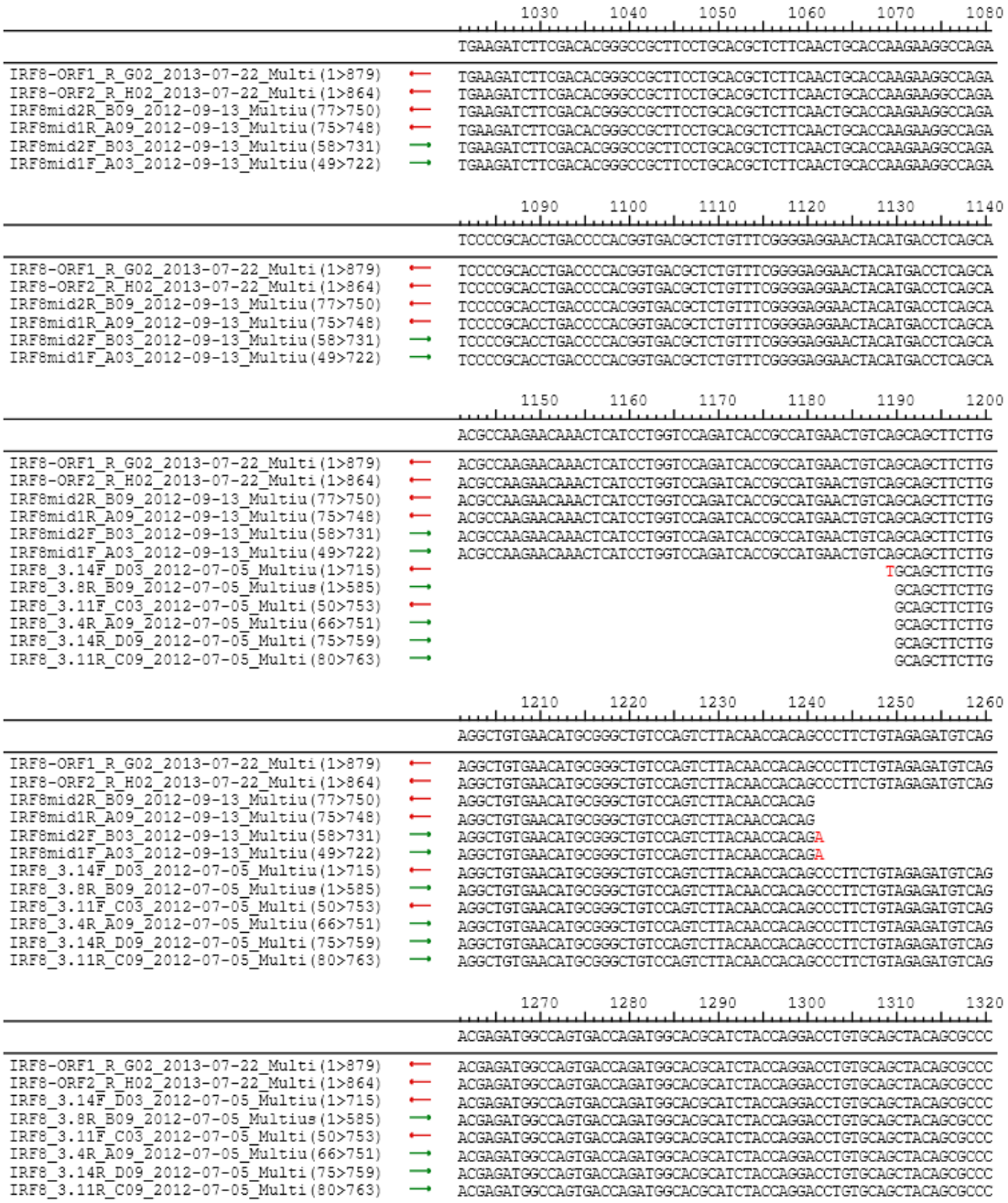

1630 1640 1650 1660 1670 1680
TCTGACTGTTCTAAGTCACTTTCAGTCTTATCACTTCCTAAGTCACTTCTCTGTGTTTAC
IRF7_3.2R_E07_2012-09-13_Multiu(65>912) TCTGACTGTTCTAAGTCACTTTCAGTCTTATCACTTCCTAAGTCACTTCTCTGTGTTTAC
IRF7_3.1F_D01_2012-09-13_Multiu(49>795) TCTGACTGTTCTAAGTCACTTTCAGTCTTATCACTTCCTAAGTCACTTCTCTGTGTTTAC
IRF7-3.1_F_A02_2012-06-14_Multi(55>652) TCTGACTGTTCTAAGTCACTTTCAGTCTTATCACTTCCTAAGTCACTTCTCTGTGTTTAC
IRF7_3.2F_E01_2012-09-13_Multiu(59>893) TCTGACTGTTCTAAGTCACTTTCAGTCTTATCACTTCCTAAGTCACTTCTCTGTGTTTAC
IRF7-3.2_F_B02_2012-06-14_Multi(45>640) TCTGACTGTTCTAAGTCACTTTCAGTCTTATCACTTCCTAAGTCACTTCTCTGTGTTTAC
IRF7_3.1R_D07_2012-09-13_Multiu(87>834) TCTGACTGTTTTAAGTCACTTTCAGTCTTATCACTTCCTAAGTCACTTCTCTGTGTTTAC
1690 1700 1710 1720 1730 1740
ATACACTGTTGTTGTCAAATATGCATCTATTCTCACCATCAAACCTGTTCTATTAGCTGG
IRF7_3.2R_E07_2012-09-13_Multiu(65>912) ATACACTGTTGTTGTCAAATATGCATCTATTCTCACCATCAAACCTGTTCTATTAGCTGG
IRF7_3.1F_D01_2012-09-13_Multiu(49>795) ATACACTGTTGTTGTCAAATATGCATCTATTCTCACCATCAAACCTGTTCTATTAGCTGG
IRF7-3.1_F_A02_2012-06-14_Multi(55>652) ATACACTGTTGTTGTCAAATATGCATCTATTCTCACCATCAAACCTGTTCTATTAGCTGG
IRF7_3.2F_E01_2012-09-13_Multiu(59>893) ATACACTGTTGTTGTCAAATATGCATCTATTCTCACCATCAAACCTGTTCTATTAGCTGG
IRF7-3.2_F_B02_2012-06-14_Multi(45>640) ATACACTGTTGTTGTCAAATATGCATCTATTCTCACCATCAAACCTGTTCTATTAGCTGG
IRF7_3.1R_D07_2012-09-13_Multiu(87>834) ATACACTGTTGTTGTCAAATATGCATCTATTCTCACCATCAAACCTGTTTTATTAGCTGG
1750 1760 1770 1780 1790 1800
TATTTATAACTACCATTCTTGAGGGTATGTTATATGTGACCGTTGCTTTTTTCCATGACG
IRF7_3.2R_E07_2012-09-13_Multiu(65>912) TATTTATAACTACCATTCTTGAGGGTATGTTATATGTGACCGTTGCTTTTTTCCATGACG
IRF7_3.1F_D01_2012-09-13_Multiu(49>795) TATTTATAACTACCATTCTTGAGGGTATGTTATATGTGACCGTTGCTTTTTTCCATGACG
IRF7-3.1_F_A02_2012-06-14_Multi(55>652) TATTTATAACTACCATTCTTGAGGGTATGTTATATGTGACCGTTGCTTTTTTCCATGACG
IRF7_3.2F_E01_2012-09-13_Multiu(59>893) TATTTATAACTACCATTCTTGAGGGTATGTTATATGTGACCGTTGCTTTTTTCCATGACG
IRF7-3.2_F_B02_2012-06-14_Multi(45>640) TATTTATAACTACCATTCTTGAGGGTATGTTATATGTGACCGTTGCTTTTTTCCATGACG
IRF7_3.1R_D07_2012-09-13_Multiu(87>834) TATTTATAACTACCATTCTTGAGGGTATGTTATATGTGACCGTTGCTTTTTTCCATGACG
1810 1820 1830 1840 1850 1860
CAACTAAAATCATTCCTTGGTT-GATGCTATTTTGTAGTTTACACATGCAATTTTTTGTC
IRF7_3.2R_E07_2012-09-13_Multiu(65>912) CAACTAAAATCATTCCTTGGTT-GATGCTATTTTGTAGTTTACACATGCAATTTTTTGTC
IRF7_3.1F_D01_2012-09-13_Multiu(49>795) CAACTAAAATCATTCCTTGGTT-GATGCTATTTTGTAGTTTACACATGCAATTTTTTGTC
IRF7-3.1_F_A02_2012-06-14_Multi(55>652) CAACTAAAATCATTCCTTGGTT-GATGCTATTTTGTAGTTTACACATGCAATTTTTTG
IRF7_3.2F_E01_2012-09-13_Multiu(59>893) CAACTAAAATCATTCCTTGGTT-GATGCTATTTTGTAGTTTACACATGCAATTTTTGTC
IRF7-3.2_F_B02_2012-06-14_Multi(45>640) CAACTAAAATCATTCCTTGGTTTGATGCTATTTTGTAGTTTACACATGCAATTTTT
IRF7_3.1R_D07_2012-09-13_Multiu(87>834) CAACTAAAATCATTCCTTGGTT-GATGCTATTTTGTAGTTTACACATGCAATTTTTTGTC
1870 1880 1890 1900 1910 1920
ATGTAATGTAAAGCACATTTCCCTGTTTGATGACCCGTTTGTAAATAAATCCTTTTTGTG
IRF7_3.2R_E07_2012-09-13_Multiu(65>912) ATGTAATGTAAAGCACATTTCCCTGTTTGATGACCCGTTTGTAAATAAATCCTTTTTGTG
IRF7_3.1F_D01_2012-09-13_Multiu(49>795) ATGTAATGTAAAGCACATTTCCCTGTTTGATGACCCGTTTGTAAATAAATCCTTTTTGTG
IRF7_3.2F_E01_2012-09-13_Multiu(59>893) ATGTAATGTAAAGCACATTTCCCTGTTTGATGACCCGTTTGTAAATAAATCCTTTTTGTG
IRF7_3.1R_D07_2012-09-13_Multiu(87>834) ATGTAATGTAAAGCACATTTCCCTGTTTGATGACCCGTTTGTAAATAAATCCTTTTTGTG
1930 1940 1950 1960 1970 1980
TTACATATATATCCTAACTGTGAGTAAGGAACAAAGGAATTTACTTAAAGAGCCCTTCGA
IRF7_3.2R_E07_2012-09-13_Multiu(65>912) TTACATATATATCCTAACTGTGAGTAAGGAACAAAGGAATTTACTTAAAGAGCCCTTCGA
IRF7_3.1F_D01_2012-09-13_Multiu(49>795) TTACATATATATCCTAACTGTGAGTAAGGAACAAAGGAATTTACTTAAAGAGCCCTTCGA
IRF7_3.2F_E01_2012-09-13_Multiu(59>893) TTACATATATATCCTAACTGTGAGTAAAGAACAAAGGAATTTACTTAAAGAGCCCTTCGA
IRF7_3.1R_D07_2012-09-13_Multiu(87>834) TTACATATATATCCTAACTGTGAGTAAGGAACAAAGGAATTTACTTAAAGAGCCCTTCGA
1990 2000 2010 2020 2030 2040
AATACAGTGGGAGTTTAATGGTTTAAGgAGCTTGGAGTGTTTTTATTTATAATTAAAAAG
IRF7_3.2R_E07_2012-09-13_Multiu(65>912) AATACAGTGGGAGTTTAATGGTTTAAGGAGCTTGGAGTGTTTTTATTTATAATTAAAAAG
IRF7_3.1F_D01_2012-09-13_Multiu(49>795) AATACAGTGGGAGTTTAATGGTTTAA--A
IRF7_3.2F_E01_2012-09-13_Multiu(59>893) AATACAGTGGGAGTTTAATGGTTTAAG-AGCTTGGAGTGTTTTTATTTATAATTAAAAAG
IRF7_3.1R_D07_2012-09-13_Multiu(87>834) AATACAGTGGGAGTTTAATGGTTTAAGGAGCTTGGAGTGTTTTTATTTATAATTAAAAAG
2050 2060
CTACTACTAAAAAAAAAAAAAAAAAAAA
IRF7_3.2R_E07_2012-09-13_Multiu(65>912) CTACTACTAAAAAAAAAAAAAAAAAAAA
IRF7_3.2F_E01_2012-09-13_Multiu(59>893) CTACTACT
IRF7_3.1R_D07_2012-09-13_Multiu(87>834) CTACTACTAAAAAAAAAAAAAAAAAAAA

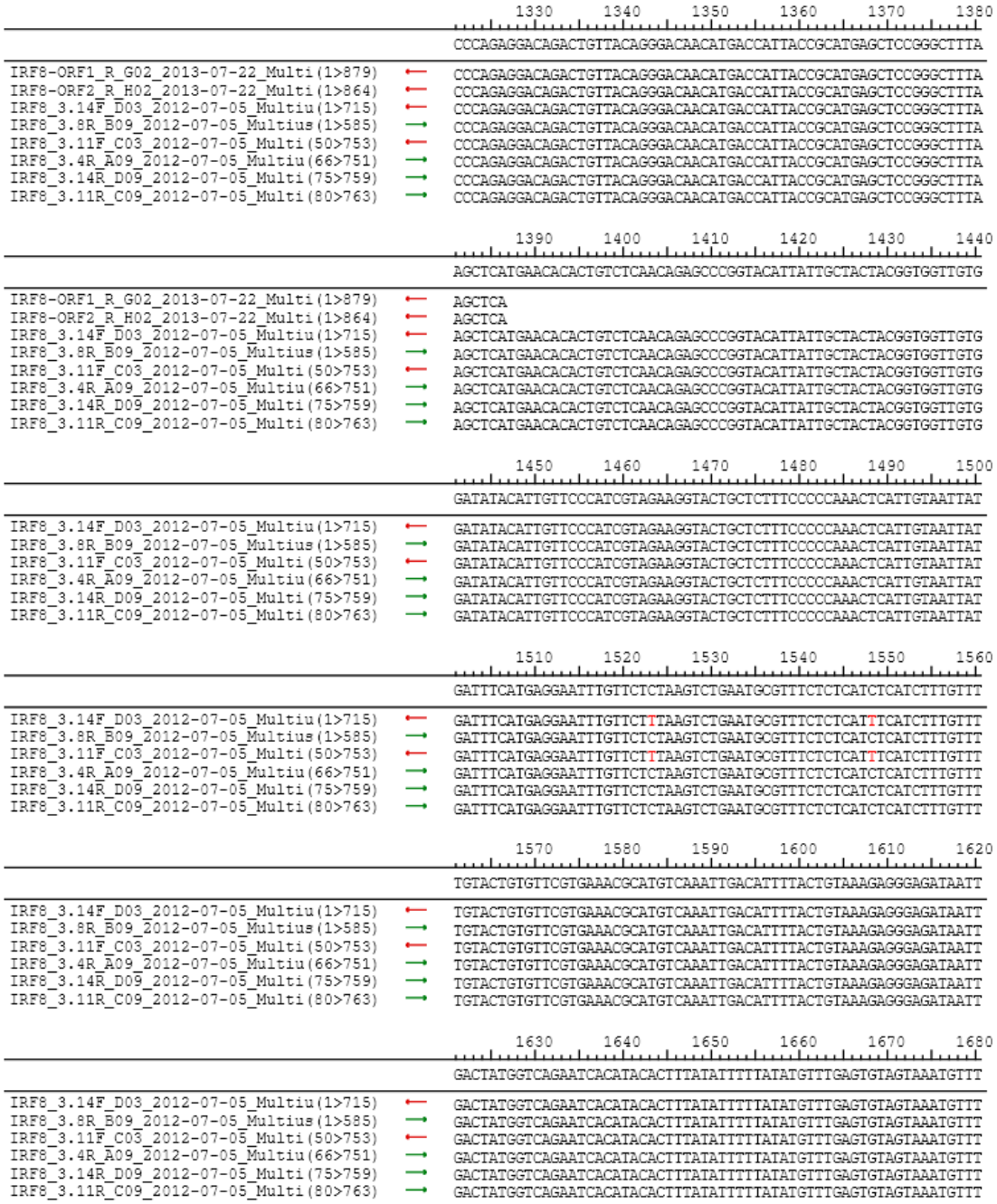
Appendix 5: Assembly of Atlantic cod *Irf8* RACE and ORF PCR sequencing reads. Sequencing methods are described in section 2.1.2. Sequence data was assembled using Lasergene SeqMan Pro software (DNASTAR). Consensus sequence is indicated between horizontal lines.

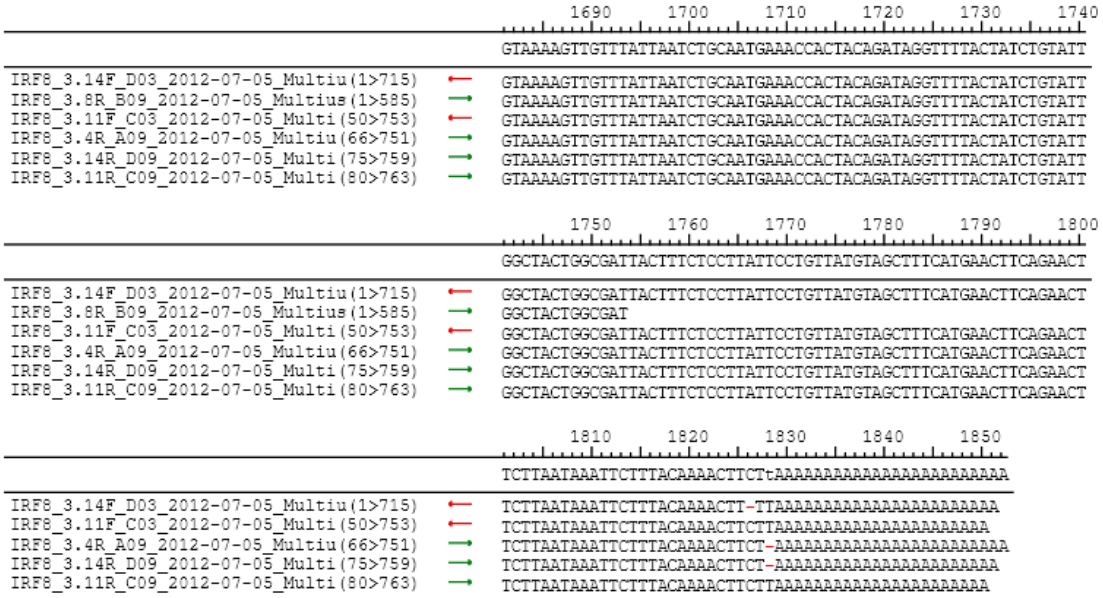

10 20 30 40 50 60
TGGACACTGACATGGACTGAAGGAGTAGAAAATTCCATTAAATGA-GGTTAAAGGTGTCA
IRF8_5.4R_F09_2012-07-05_Multiu(86>737)
IRF8_5.7R_H09_2012-07-05_Multiu(86>736)
IRF8_5.1R_E09_2012-07-05_Multiu(83>733)
IRF8_5.5F_G03_2012-07-05_Multiu(52>729)
IRF8_5.1F_E03_2012-07-05_Multiu(47>697)
IRF8_5.4F_F03_2012-07-05_Multiu(44>694)
70 80 90 100 110 120
TCTGTCTGGAGCTGGAAATAATTCGTGGATATAAAGTCAAGATGTCGAACACGGGAGGAC
IRF8_5.4R_F09_2012-07-05_Multiu(86>737)
IRF8_5.7R_H09_2012-07-05_Multiu(86>736)
IRF8_5.1R_E09_2012-07-05_Multiu(83>733)
IRF8_5.5F_G03_2012-07-05_Multiu(52>729)
IRF8_5.1F_E03_2012-07-05_Multiu(47>697)
IRF8_5.4F_F03_2012-07-05_Multiu(44>694)
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937)
130 140 150 160 170 180
GAAGACTGAAGCAGTGGTTGATTGAGCAGATCAAGAGCGGACAGTACTCGGGGCTTGAGT
IRF8_5.4R_F09_2012-07-05_Multiu(86>737)
IRF8_5.7R_H09_2012-07-05_Multiu(86>736)
IRF8_5.1R_E09_2012-07-05_Multiu(83>733)
IRF8_5.5F_G03_2012-07-05_Multiu(52>729)
IRF8_5.1F_E03_2012-07-05_Multiu(47>697)
IRF8_5.4F_F03_2012-07-05_Multiu(44>694)
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937)
190 200 210 220 230 240
GGGAGGATGACAGCCTCACCATGTTCCGCATCCCATGGAAGCATGCTGGGAAGCAGGATT
IRF8_5.4R_F09_2012-07-05_Multiu(86>737)
IRF8_5.7R_H09_2012-07-05_Multiu(86>736)
IRF8_5.1R_E09_2012-07-05_Multiu(83>733)
IRF8_5.5F_G03_2012-07-05_Multiu(52>729)
IRF8_5.1F_E03_2012-07-05_Multiu(47>697)
IRF8_5.4F_F03_2012-07-05_Multiu(44>694)
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937)
250 260 270 280 290 300
ATAACCAAGAGGTCGATGCTTCCATCTTCAAGGCCTGGGCTGTGTTTAAGGGCAAGTTTA
IRF8_5.4R_F09_2012-07-05_Multiu(86>737)
IRF8_5.7R_H09_2012-07-05_Multiu(86>736)
IRF8_5.1R_E09_2012-07-05_Multiu(83>733)
IRF8_5.5F_G03_2012-07-05_Multiu(52>729)
IRF8_5.1F_E03_2012-07-05_Multiu(47>697)
IRF8_5.4F_F03_2012-07-05_Multiu(44>694)
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937)
310 320 330 340 350 360
AAGAGGGGGAGAAGGCTGAGCCTGCTACTTGGAAAACTAGGCTCCGCTGTGCCCTGAACA
IRF8_5.4R_F09_2012-07-05_Multiu(86>737)
IRF8_5.7R_H09_2012-07-05_Multiu(86>736)
IRF8_5.1R_E09_2012-07-05_Multiu(83>733)
IRF8_5.5F_G03_2012-07-05_Multiu(52>729)
IRF8_5.1F_E03_2012-07-05_Multiu(47>697)
IRF8_5.4F_F03_2012-07-05_Multiu(44>694)
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937)


370 380 390 400 410 420
AGAGCCCAGACTTTGAGGAGGTGACGGACAGGTCACAACTGGACATCTCAGAGCCCTACA
IRF8_5.4R_F09_2012-07-05_Multiu(86>737)
IRF8_5.7R_H09_2012-07-05_Multiu(86>736)
IRF8_5.1R_E09_2012-07-05_Multiu(83>733)
IRF8_5.5F_G03_2012-07-05_Multiu(52>729)
IRF8_5.1F_E03_2012-07-05_Multiu(47>697)
IRF8_5.4F_F03_2012-07-05_Multiu(44>694)
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937)
430 440 450 460 470 480
AGGTCTACCGCATTGTCCCTGAAGAGGAACAGAAGCTCGGTAAAACCACAGCGATGGTCA
IRF8_5.4R_F09_2012-07-05_Multiu(86>737)
IRF8_5.7R_H09_2012-07-05_Multiu(86>736)
IRF8_5.1R_E09_2012-07-05_Multiu(83>733)
IRF8_5.5F_G03_2012-07-05_Multiu(52>729)
IRF8_5.1F_E03_2012-07-05_Multiu(47>697)
IRF8_5.4F_F03_2012-07-05_Multiu(44>694)
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937)
490 500 510 520 530 540
CGACAGCTGGCGACATCGCTGACTTAGACTGCAGCTCTGCAGAGCTAGAGGAGTTGATCA
IRF8_5.4R_F09_2012-07-05_Multiu(86>737)
IRF8_5.7R_H09_2012-07-05_Multiu(86>736)
IRF8_5.1R_E09_2012-07-05_Multiu(83>733)
IRF8_5.5F_G03_2012-07-05_Multiu(52>729)
IRF8_5.1F_E03_2012-07-05_Multiu(47>697)
IRF8_5.4F_F03_2012-07-05_Multiu(44>694)
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937)
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879)
550 560 570 580 590 600
AAGTGGCCTCCACTGATGACTACCCCTCAGCCATCAAGAGAAGCTACTCTCCCCAGGAGG
IRF8_5.4R_F09_2012-07-05_Multiu(86>737)
IRF8_5.7R_H09_2012-07-05_Multiu(86>736)
IRF8_5.1R_E09_2012-07-05_Multiu(83>733)
IRF8_5.5F_G03_2012-07-05_Multiu(52>729)
IRF8_5.1F_E03_2012-07-05_Multiu(47>697)
IRF8_5.4F_F03_2012-07-05_Multiu(44>694)
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937)
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879)
IRF8-ORF2_R_H02_2013-07-22_Multi(1>864)
IRF8mid2R_B09_2012-09-13_Multiu(77>750)
IRF8mid1R_A09_2012-09-13_Multiu(75>748)
IRF8mid2F_B03_2012-09-13_Multiu(58>731)
IRF8mid1F_A03_2012-09-13_Multiu(49>722)
610 620 630 640 650 660
ATGGCTTTAACGTCCAAGCCAGCCCTGAGTACTGGTCACATGGCAGCATACCCGTGTTCT
IRF8_5.4R_F09_2012-07-05_Multiu(86>737)
IRF8_5.7R_H09_2012-07-05_Multiu(86>736)
IRF8_5.1R_E09_2012-07-05_Multiu(83>733)
IRF8_5.5F_G03_2012-07-05_Multiu(52>729)
IRF8_5.1F_E03_2012-07-05_Multiu(47>697)
IRF8_5.4F_F03_2012-07-05_Multiu(44>694)
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937)
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879)
IRF8-ORF2_R_H02_2013-07-22_Multi(1>864)
IRF8mid2R_B09_2012-09-13_Multiu(77>750)
IRF8mid1R_A09_2012-09-13_Multiu(75>748)
IRF8mid2F_B03_2012-09-13_Multiu(58>731)
IRF8mid1F_A03_2012-09-13_Multiu(49>722)

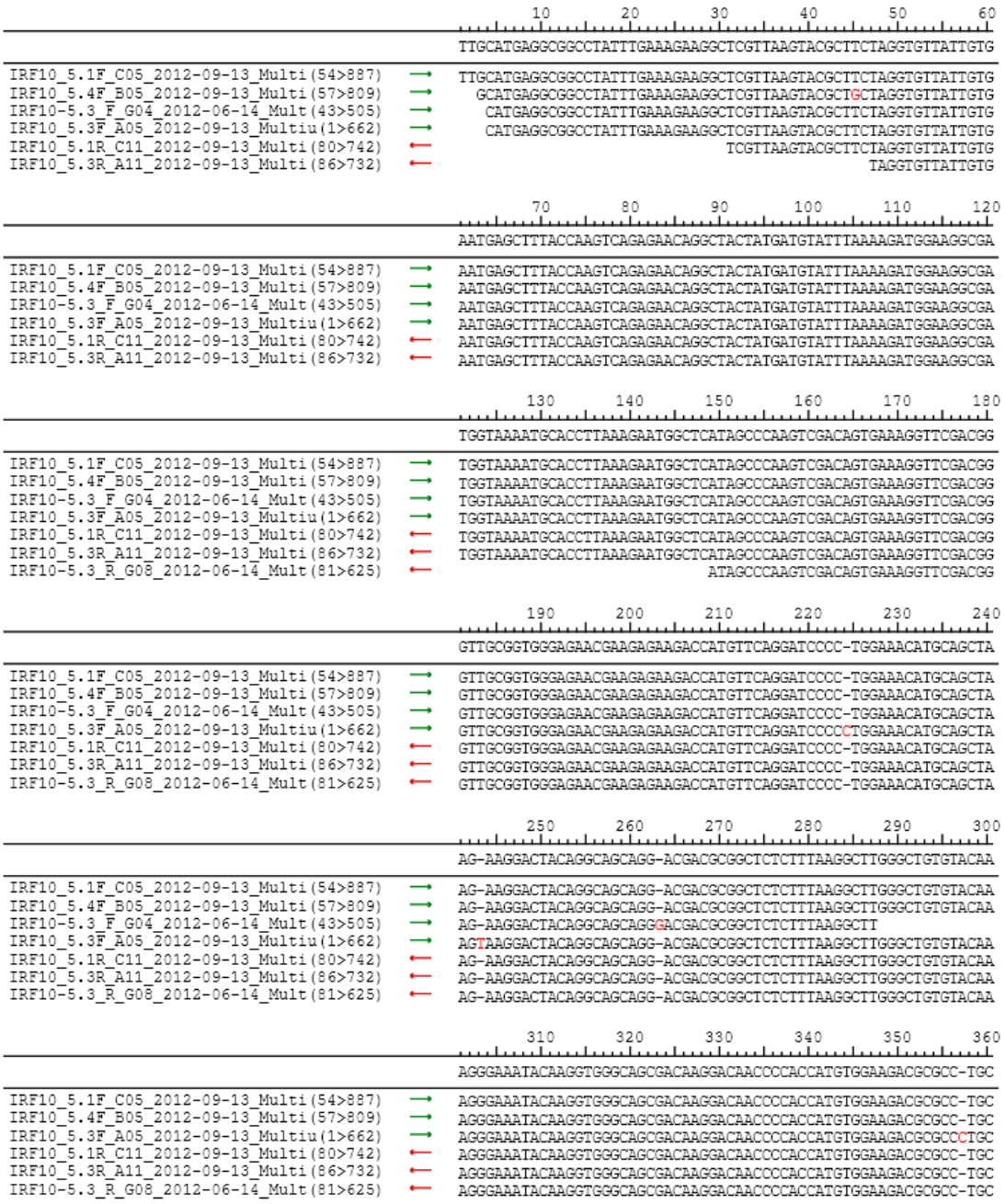

670 680 690 700 710 720
CCCAGATGATGATCTCCTTCTACTACGGGGGCCAGCTGATGCATAGCACGGTGACCTCCC
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937) CCCAGATGATGATCTCCTTCTACTACGGGGGCCAGCTGATGCATAGCACGGTGACCTCCC
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879) CCCAGATGATGATCTCCTTCTACTACGGGGGCCAGCTGATGCATAGCACGGTGACCTCCC
IRF8-ORF2_R_H02_2013-07-22_Multi(1>864) CCCAGATGATGATCTCCTTCTACTACGGGGGCCAGCTGATGCATAGCACGGTGACCTCCC
IRF8mid2R_B09_2012-09-13_Multiu(77>750) CCCAGATGATGATCTCCTTCTACTACGGGGGCCAGCTGATGCATAGCACGGTGACCTCCC
IRF8mid1R_A09_2012-09-13_Multiu(75>748) CCCAGATGATGATCTCCTTCTACTACGGGGGCCAGCTGATGCATAGCACGGTGACCTCCC
IRF8mid2F_B03_2012-09-13_Multiu(58>731) CCCAGATGATGATCTCCTTCTACTACGGGGGCCAGCTGATGCATAGCACGGTGACCTCCC
IRF8mid1F_A03_2012-09-13_Multiu(49>722) CCCAGATGATGATCTCCTTCTACTACGGGGGCCAGCTGATGCATAGCACGGTGACCTCCC
730 740 750 760 770 780
ACCCCGAGGGCTGCCGCATCTCCCCGGTGCTGCCCCAGCAGCGCGCCGTCGCCCGCGGCT
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937) ACCCCGAGGGCTGCCGCATCTCCCCGGTGCTGCCCCAGCAGCGCGCCGTCGCCCGCGGCT
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879) ACCCCGAGGGCTGCCGCATCTCCCCGGTGCTGCCCCAGCAGCGCGCCGTCGCCCGCGGCT
IRF8-ORF2_R_H02_2013-07-22_Multi(1>864) ACCCCGAGGGCTGCCGCATCTCCCCGGTGCTGCCCCAGCAGCGCGCCGTCGCCCGCGGCT
IRF8mid2R_B09_2012-09-13_Multiu(77>750) ACCCCGAGGGCTGCCGCATCTCCCCGGTGCTGCCCCAGCAGCGCGCCGTCGCCCGCGGCT
IRF8mid1R_A09_2012-09-13_Multiu(75>748) ACCCCGAGGGCTGCCGCATCTCCCCGGTGCTGCCCCAGCAGCGCGCCGTCGCCCGCGGCT
IRF8mid2F_B03_2012-09-13_Multiu(58>731) ACCCCGAGGGCTGCCGCATCTCCCCGGTGCTGCCCCAGCAGCGCGCCGTCGCCCGCGGCT
IRF8mid1F_A03_2012-09-13_Multiu(49>722) ACCCCGAGGGCTGCCGCATCTCCCCGGTGCTGCCCCAGCAGCGCGCCGTCGCCCGCGGCT
790 800 810 820 830 840
ACAGCTCCGACACCATGCAGAGCGTGCACTTCCCGCCGGCCGACCTCATCGACAACGAGC
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937) ACAGCTCCGACACCATGCAGAGCGTGCACTTCCCGCCGGCCGACCTCATCGACAACGAGC
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879) ACAGCTCCGACACCATGCAGAGCGTGCACTTCCCGCCGGCCGACCTCATCGACAACGAGC
IRF8-ORF2_R_H02_2013-07-22_Multi(1>864) ACAGCTCCGACACCATGCAGAGCGTGCACTTCCCGCCGGCCGACCTCATCGACAACGAGC
IRF8mid2R_B09_2012-09-13_Multiu(77>750) ACAGCTCCGACACCATGCAGAGCGTGCACTTCCCGCCGGCCGACCTCATCGACAACGAGC
IRF8mid1R_A09_2012-09-13_Multiu(75>748) ACAGCTCCGACACCATGCAGAGCGTGCACTTCCCGCCGGCCGACCTCATCGACAACGAGC
IRF8mid2F_B03_2012-09-13_Multiu(58>731) ACAGCTCCGACACCATGCAGAGCGTGCACTTCCCGCCGGCCGACCTCATCGACAACGAGC
IRF8mid1F_A03_2012-09-13_Multiu(49>722) ACAGCTCCGACACCATGCAGAGCGTGCACTTCCCGCCGGCCGACCTCATCGACAACGAGC
850 860 870 880 890 900
GCCAGCGGCAGGTCACCTGCAAGCTGCTGGGCCACCTGGAGAGGGGCGTTCTGGTGCGGG
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937) GCCAGCGGCAGGTCACCTGCAAGCTGCTGGGCCACCTGGAGAGGGGCGTTCTGGTGCGGG
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879) GCCAGCGGCAGGTCACCTGCAAGCTGCTGGGCCACCTGGAGAGGGGCGTTCTGGTGCGGG
IRF8-ORF2_R_H02_2013-07-22_Multi(1>864) GCCAGCGGCAGGTCACCTGCAAGCTGCTGGGCCACCTGGAGAGGGGCGTTCTGGTGCGGG
IRF8mid2R_B09_2012-09-13_Multiu(77>750) GCCAGCGGCAGGTCACCTGCAAGCTGCTGGGCCACCTGGAGAGGGGCGTTCTGGTGCGGG
IRF8mid1R_A09_2012-09-13_Multiu(75>748) GCCAGCGGCAGGTCACCTGCAAGCTGCTGGGCCACCTGGAGAGGGGCGTTCTGGTGCGGG
IRF8mid2F_B03_2012-09-13_Multiu(58>731) GCCAGCGGCAGGTCACCTGCAAGCTGCTGGGCCACCTGGAGAGGGGCGTTCTGGTGCGGG
IRF8mid1F_A03_2012-09-13_Multiu(49>722) GCCAGCGGCAGGTCACCTGCAAGCTGCTGGGCCACCTGGAGAGGGGCGTTCTGGTGCGGG
910 920 930 940 950 960
CCAACCGCGAGGGCGTCTTCATCAAGAGGCTCTGCCAGAGCCGGGTGTTCTGGAGCGGCC
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937) CCAACCGCGAGGGCGTCTTCATCAAGAGGCTCTGCCAGAGCCGGGTGTTCTGGAGCGGCC
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879) CCAACCGCGAGGGCGTCTTCATCAAGAGGCTCTGCCAGAGCCGGGTGTTCTGGAGCGGCC
IRF8-ORF2_R_H02_2013-07-22_Multi(1>864) CCAACCGCGAGGGCGTCTTCATCAAGAGGCTCTGCCAGAGCCGGGTGTTCTGGAGCGGCC
IRF8mid2R_B09_2012-09-13_Multiu(77>750) CCAACCGCGAGGGCGTCTTCATCAAGAGGCTCTGCCAGAGCCGGGTGTTCTGGAGCGGCC
IRF8mid1R_A09_2012-09-13_Multiu(75>748) CCAACCGCGAGGGCGTCTTCATCAAGAGGCTCTGCCAGAGCCGGGTGTTCTGGAGCGGCC
IRF8mid2F_B03_2012-09-13_Multiu(58>731) CCAACCGCGAGGGCGTCTTCATCAAGAGGCTCTGCCAGAGCCGGGTGTTCTGGAGCGGCC
IRF8mid1F_A03_2012-09-13_Multiu(49>722) CCAACCGCGAGGGCGTCTTCATCAAGAGGCTCTGCCAGAGCCGGGTGTTCTGGAGCGGCC
970 980 990 1000 1010 1020
ACGGGGAGCACGGCCAACACCACGGCCCCGTGACCTGCAAGCTGGAGAGGGACGCCGTGG
IRF8-ORF1_F_G01_2013-07-22_Mult(48>937) ACGGGGAGCACGGCCAACACCACGGCCCC
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879) ACGGGGAGCACGGCCAACACCACGGCCCCGTGACCTGCAAGCTGGAGAGGGACGCCGTGG
IRF8-ORF2_R_H02_2013-07-22_Multi(1>864) ACGGGGAGCACGGCCAACACCACGGCCCCGTGACCTGCAAGCTGGAGAGGGACGCCGTGG
IRF8mid2R_B09_2012-09-13_Multiu(77>750) ACGGGGAGCACGGCCAACACCACGGCCCCGTGACCTGCAAGCTGGAGAGGGACGCCGTGG
IRF8mid1R_A09_2012-09-13_Multiu(75>748) ACGGGGAGCACGGCCAACACCACGGCCCCGTGACCTGCAAGCTGGAGAGGGACGCCGTGG
IRF8mid2F_B03_2012-09-13_Multiu(58>731) ACGGGGAGCACGGCCAACACCACGGCCCCGTGACCTGCAAGCTGGAGAGGGACGCCGTGG
IRF8mid1F_A03_2012-09-13_Multiu(49>722) ACGGGGAGCACGGCCAACACCACGGCCCCGTGACCTGCAAGCTGGAGAGGGACGCCGTGG

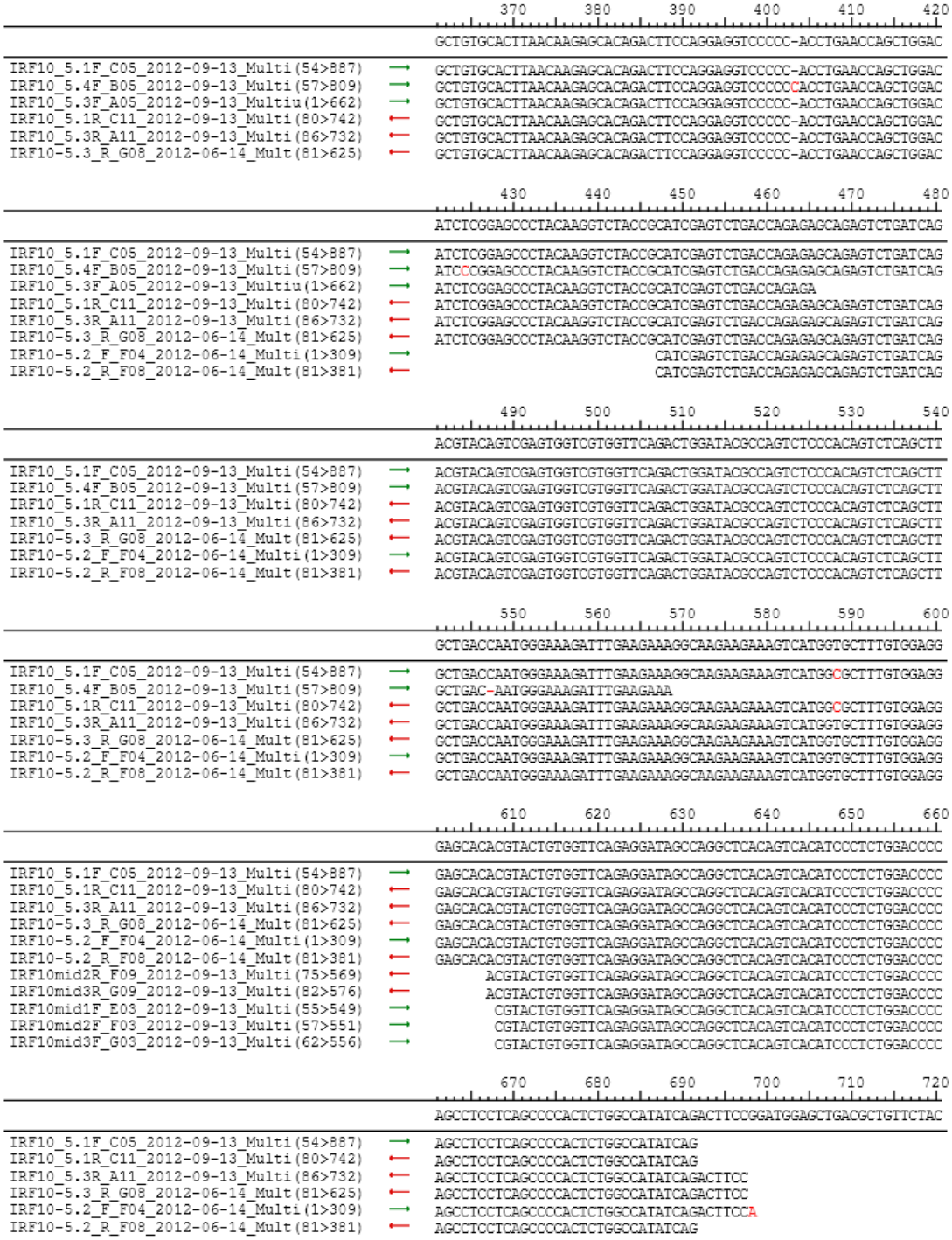

1030 1040 1050 1060 1070 1080
TGAAGATCTTCGACACGGGCCGCTTCCTGCACGCTCTTCAACTGCACCAAGAAGGCCAGA
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879)
IRF8-ORF2_R_H02_2013-07-22_Multi(1>864)
IRF8mid2R_B09_2012-09-13_Multiu(77>750)
IRF8mid1R_A09_2012-09-13_Multiu(75>748)
IRF8mid2F_B03_2012-09-13_Multiu(58>731)
IRF8mid1F_A03_2012-09-13_Multiu(49>722)
1090 1100 1110 1120 1130 1140
TCCCCGCACCTGACCCCACGGTGACGCTCTGTTTCGGGGAGGAACTACATGACCTCAGCA
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879)
IRF8-ORF2_R_H02_2013-07-22_Multi(1>864)
IRF8mid2R_B09_2012-09-13_Multiu(77>750)
IRF8mid1R_A09_2012-09-13_Multiu(75>748)
IRF8mid2F_B03_2012-09-13_Multiu(58>731)
IRF8mid1F_A03_2012-09-13_Multiu(49>722)
1150 1160 1170 1180 1190 1200
ACGCCAAGAACAAACTCATCCTGGTCCAGATCACCGCCATGAACTGTCAGCAGCTTCTTG
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879)
IRF8-ORF2_R_H02_2013-07-22_Multi(1>864)
IRF8mid2R_B09_2012-09-13_Multiu(77>750)
IRF8mid1R_A09_2012-09-13_Multiu(75>748)
IRF8mid2F_B03_2012-09-13_Multiu(58>731)
IRF8mid1F_A03_2012-09-13_Multiu(49>722)
IRF8_3.14F_D03_2012-07-05_Multiu(1>715)
IRF8_3.8R_B09_2012-07-05_Multius(1>585)
IRF8_3.11F_C03_2012-07-05_Multi(50>753)
IRF8_3.4R_A09_2012-07-05_Multiu(66>751)
IRF8_3.14R_D09_2012-07-05_Multi(75>759)
IRF8_3.11R_C09_2012-07-05_Multi(80>763)
TGCAGCTTCTTG
GCAGCTTCTTG
GCAGCTTCTTG
GCAGCTTCTTG
GCAGCTTCTTG
GCAGCTTCTTG
1210 1220 1230 1240 1250 1260
AGGCTGTGAACATGCGGGCTGTCCAGTCTTACAACCACAGCCCTTCTGTAGAGATGTCAG
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879)
IRF8-ORF2_R_H02_2013-07-22_Multi(1>864)
IRF8mid2R_B09_2012-09-13_Multiu(77>750)
IRF8mid1R_A09_2012-09-13_Multiu(75>748)
IRF8mid2F_B03_2012-09-13_Multiu(58>731)
IRF8mid1F_A03_2012-09-13_Multiu(49>722)
IRF8_3.14F_D03_2012-07-05_Multiu(1>715)
IRF8_3.8R_B09_2012-07-05_Multius(1>585)
IRF8_3.11F_C03_2012-07-05_Multi(50>753)
IRF8_3.4R_A09_2012-07-05_Multiu(66>751)
IRF8_3.14R_D09_2012-07-05_Multi(75>759)
IRF8_3.11R_C09_2012-07-05_Multi(80>763)
AGGCTGTGAACATGCGGGCTGTCCAGTCTTACAACCACAG
AGGCTGTGAACATGCGGGCTGTCCAGTCTTACAACCACAGA
1270 1280 1290 1300 1310 1320
ACGAGATGGCCAGTGACCAGATGGCACGCATCTACCAGGACCTGTGCAGCTACAGCGCCC
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879)
IRF8-ORF2_R_H02_2013-07-22_Multi(1>864)
IRF8_3.14F_D03_2012-07-05_Multiu(1>715)
IRF8_3.8R_B09_2012-07-05_Multius(1>585)
IRF8_3.11F_C03_2012-07-05_Multi(50>753)
IRF8_3.4R_A09_2012-07-05_Multiu(66>751)
IRF8_3.14R_D09_2012-07-05_Multi(75>759)
IRF8_3.11R_C09_2012-07-05_Multi(80>763)

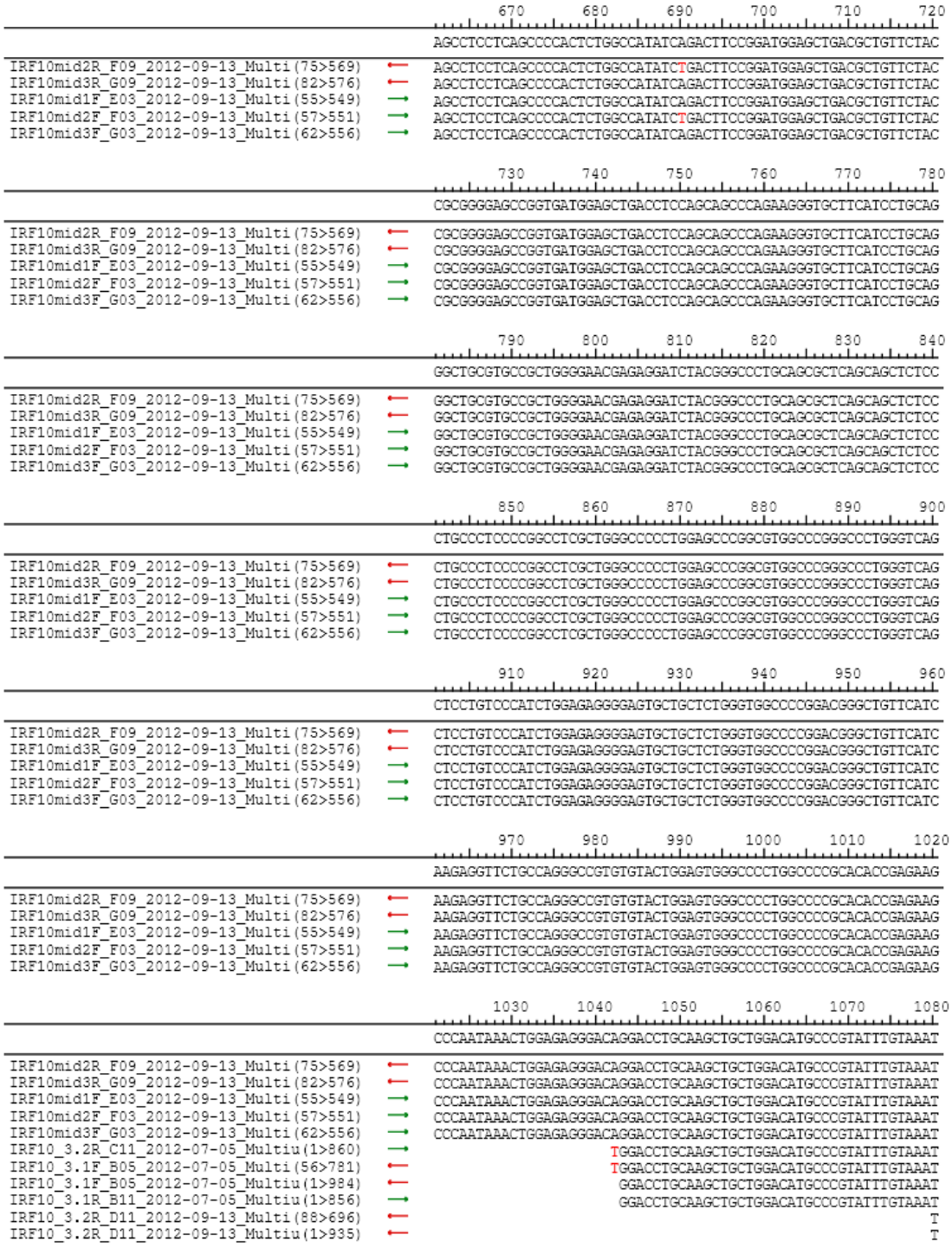

1330 1340 1350 1360 1370 1380
CCCAGAGGACAGACTGTTACAGGGACAACATGACCATTACCGCATGAGCTCCGGGCTTTA
IRF8-ORF1_R_G02_2013-07-22_Multi(1>879)
IRF8-ORF2_R_H02_2013-07-22_Multi(1>864)
IRF8_3.14F_D03_2012-07-05_Multiu(1>715)
IRF8_3.8R_B09_2012-07-05_Multius(1>585)
IRF8_3.11F_C03_2012-07-05_Multi(50>753)
IRF8_3.4R_A09_2012-07-05_Multiu(66>751)
IRF8_3.14R_D09_2012-07-05_Multi(75>759)
IRF8_3.11R_C09_2012-07-05_Multi(80>763)
1390 1400 1410 1420 1430 1440
AGCTCATGAACACACTGTCTCAACAGAGCCCGGTACATTATTGCTACTACGGTGGTTGTG
AGCTCA
AGCTCATGAACACACTGTCTCAACAGAGCCCGGTACATTATTGCTACTACGGTGGTTGTG
1450 1460 1470 1480 1490 1500
GATATACATTGTTCCCATCGTAGAAGGTACTGCTCTTTCCCCCAAACTCATTGTAATTAT
1510 1520 1530 1540 1550 1560
GATTTCATGAGGAATTTGTTCTCTAAGTCTGAATGCGTTTCTCTCATCTCATCTTTGTTT
1570 1580 1590 1600 1610 1620
TGTACTGTGTTCGTGAAACGCATGTCAAATTGACATTTTACTGTAAAGAGGGAGATAATT
1630 1640 1650 1660 1670 1680
GACTATGGTCAGAATCACATACACTTTATATTTTTATATGTTTGAGTGTAGTAAATGTTT

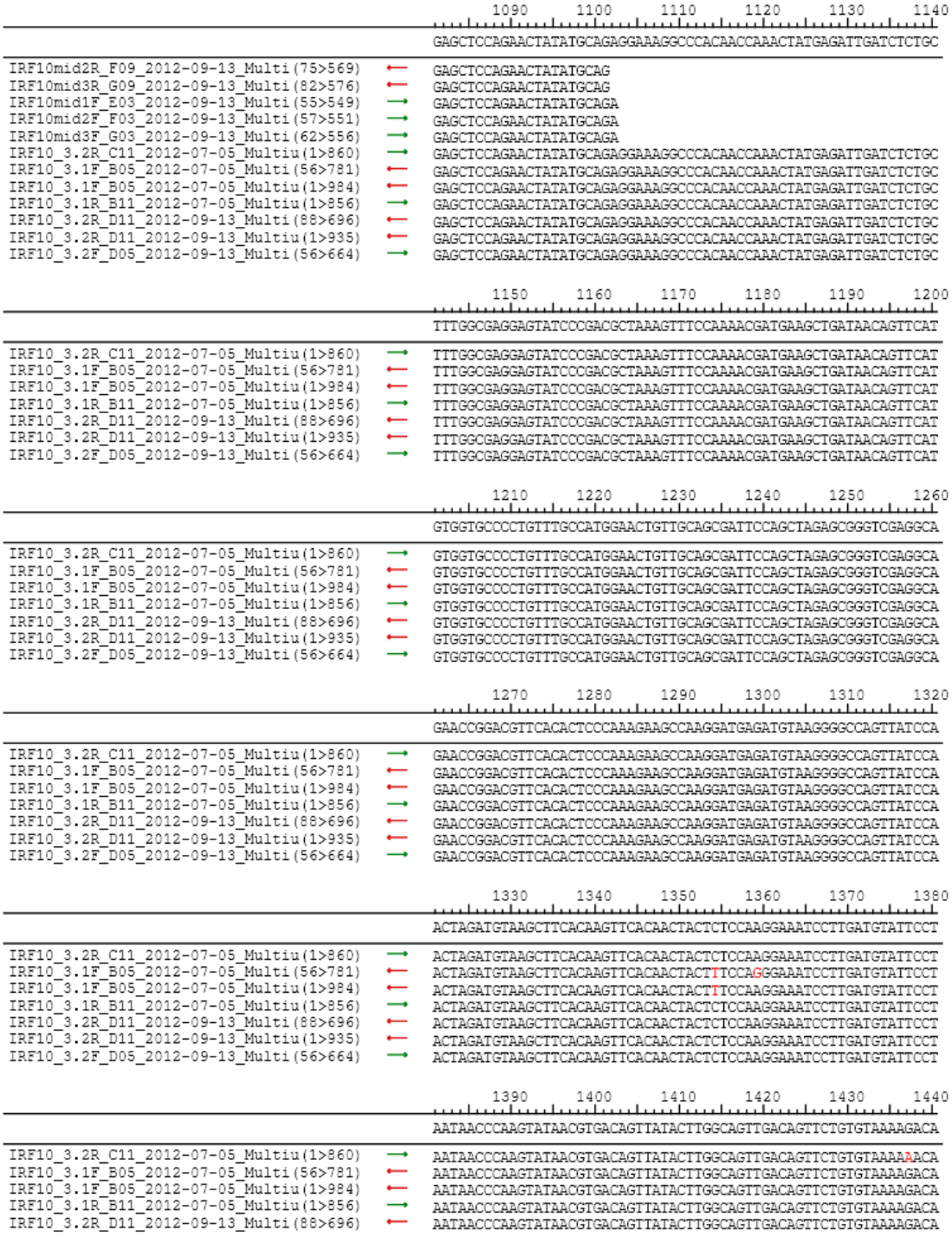

1690 1700 1710 1720 1730 1740
GTAAAAGTTGTTTATTAATCTGCAATGAAACCACTACAGATAGGTTTTACTATCTGTATT
IRF8_3.14F_D03_2012-07-05_Multiu(1>715) GTAAAAGTTGTTTATTAATCTGCAATGAAACCACTACAGATAGGTTTTACTATCTGTATT
IRF8_3.8R_B09_2012-07-05_Multius(1>585) GTAAAAGTTGTTTATTAATCTGCAATGAAACCACTACAGATAGGTTTTACTATCTGTATT
IRF8_3.11F_C03_2012-07-05_Multi(50>753) GTAAAAGTTGTTTATTAATCTGCAATGAAACCACTACAGATAGGTTTTACTATCTGTATT
IRF8_3.4R_A09_2012-07-05_Multiu(66>751) GTAAAAGTTGTTTATTAATCTGCAATGAAACCACTACAGATAGGTTTTACTATCTGTATT
IRF8_3.14R_D09_2012-07-05_Multi(75>759) GTAAAAGTTGTTTATTAATCTGCAATGAAACCACTACAGATAGGTTTTACTATCTGTATT
IRF8_3.11R_C09_2012-07-05_Multi(80>763) GTAAAAGTTGTTTATTAATCTGCAATGAAACCACTACAGATAGGTTTTACTATCTGTATT
1750 1760 1770 1780 1790 1800
GGCTACTGGCGATTACTTTCTCCTTATTCCTGTTATGTAGCTTTCATGAACTTCAGAACT
IRF8_3.14F_D03_2012-07-05_Multiu(1>715) GGCTACTGGCGATTACTTTCTCCTTATTCCTGTTATGTAGCTTTCATGAACTTCAGAACT
IRF8_3.8R_B09_2012-07-05_Multius(1>585) GGCTACTGGCGAT
IRF8_3.11F_C03_2012-07-05_Multi(50>753) GGCTACTGGCGATTACTTTCTCCTTATTCCTGTTATGTAGCTTTCATGAACTTCAGAACT
IRF8_3.4R_A09_2012-07-05_Multiu(66>751) GGCTACTGGCGATTACTTTCTCCTTATTCCTGTTATGTAGCTTTCATGAACTTCAGAACT
IRF8_3.14R_D09_2012-07-05_Multi(75>759) GGCTACTGGCGATTACTTTCTCCTTATTCCTGTTATGTAGCTTTCATGAACTTCAGAACT
IRF8_3.11R_C09_2012-07-05_Multi(80>763) GGCTACTGGCGATTACTTTCTCCTTATTCCTGTTATGTAGCTTTCATGAACTTCAGAACT
1810 1820 1830 1840 1850
TCTTAATAAATTCTTTACAAAACTTCTtAAAAAAAAAAAAAAAAAAAAAAAAAA
IRF8_3.14F_D03_2012-07-05_Multiu(1>715) TCTTAATAAATTCTTTACAAAACTT-TTAAAAAAAAAAAAAAAAAAAAAAAAA
IRF8_3.11F_C03_2012-07-05_Multi(50>753) TCTTAATAAATTCTTTACAAAACTTCTTAAAAAAAAAAAAAAAAAAAAAAAA
IRF8_3.4R_A09_2012-07-05_Multiu(66>751) TCTTAATAAATTCTTTACAAAACTTCT-AAAAAAAAAAAAAAAAAAAAAAAAA
IRF8_3.14R_D09_2012-07-05_Multi(75>759) TCTTAATAAATTCTTTACAAAACTTCT-AAAAAAAAAAAAAAAAAAAAAAAA
IRF8_3.11R_C09_2012-07-05_Multi(80>763) TCTTAATAAATTCTTTACAAAACTTCTTAAAAAAAAAAAAAAAAAAAAAA

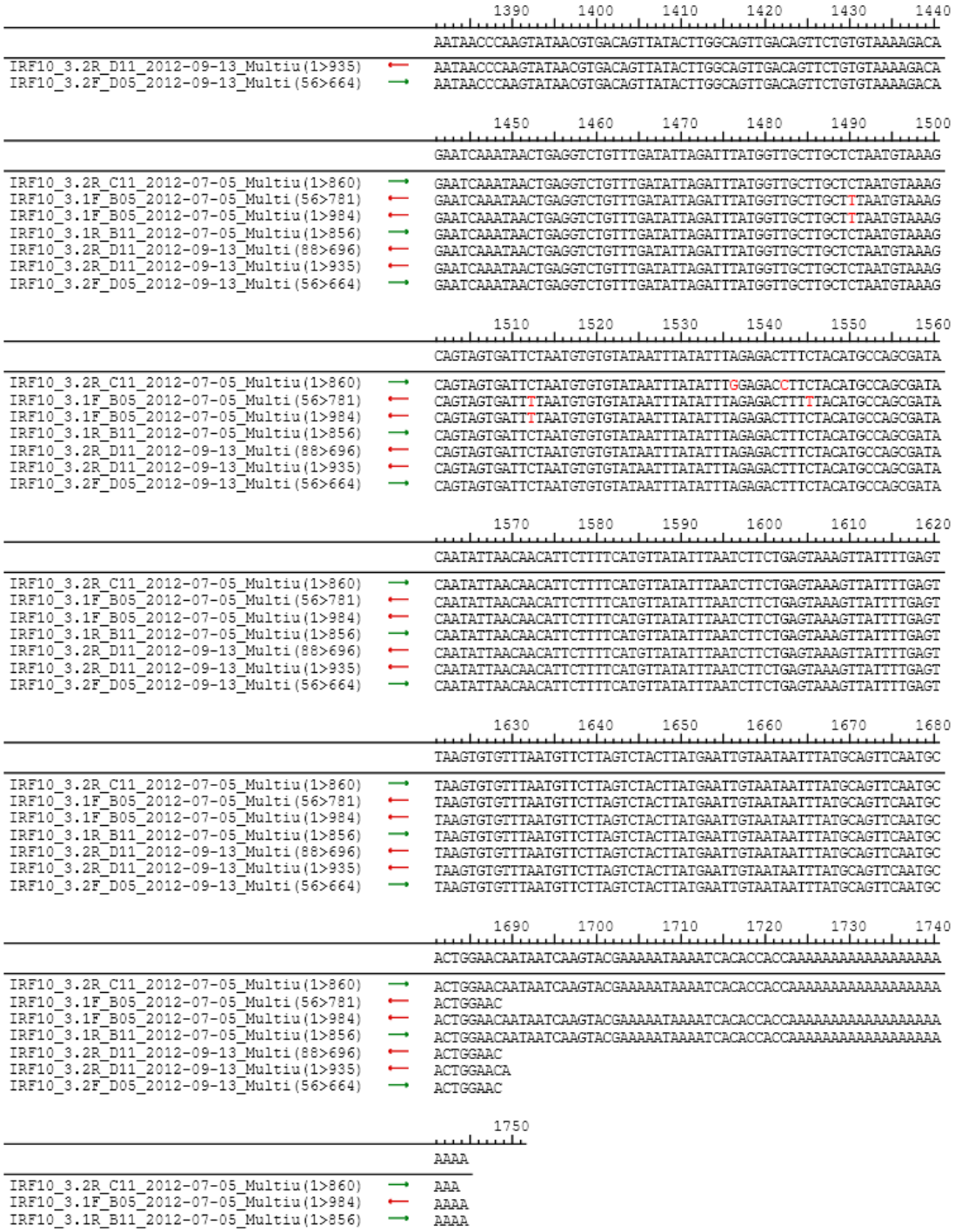
Appendix 6: Assembly of Atlantic cod *Irf10-v1* RACE and ORF PCR sequencing reads. Sequencing methods are described in section 2.1.2. Sequence data was assembled using Lasergene SeqMan Pro software (DNASTAR). Consensus sequence is indicated between horizontal lines. Note that sequences named "*Irf10*" or "*Irf10a*" were renamed as *Irf10-v1* after the second *Irf10* splice variant (*Irf10-v2*) was discovered.


10 20 30 40 50 60
TTGCATGAGGCGGCCTATTTGAAAGAAGGCTCGTTAAGTACGCTTCTAGGTGTTATTGTG
IRF10_5.1F_C05_2012-09-13_Multi(54>887)
IRF10_5.4F_B05_2012-09-13_Multi(57>809)
IRF10-5.3_F_G04_2012-06-14_Mult(43>505)
IRF10_5.3F_A05_2012-09-13_Multiu(1>662)
IRF10_5.1R_C11_2012-09-13_Multi(80>742)
IRF10_5.3R_A11_2012-09-13_Multi(86>732)
70 80 90 100 110 120
AATGAGCTTTACCAAGTCAGAGAACAGGCTACTATGATGTATTTAAAAGATGGAAGGCGA
130 140 150 160 170 180
TGGTAAAATGCACCTTAAAGAATGGCTCATAGCCCAAGTCGACAGTGAAAGGTTCGACGG
IRF10_5.3R_R_G08_2012-06-14_Mult(81>625)
190 200 210 220 230 240
GTTGCGGTGGGAGAACGAAGAGAAGACCATGTTCAGGATCCCC-TGGAAACATGCAGCTA
250 260 270 280 290 300
AG-AAGGACTACAGGCAGCAGG-ACGACGCGGCTCTCTTTAAGGCTTGGGCTGTGTACAA
310 320 330 340 350 360
AGGGAAATACAAGGTGGGCAGCGACAAGGACAACCCCACCATGTGGAAGACGCGCC-TGC

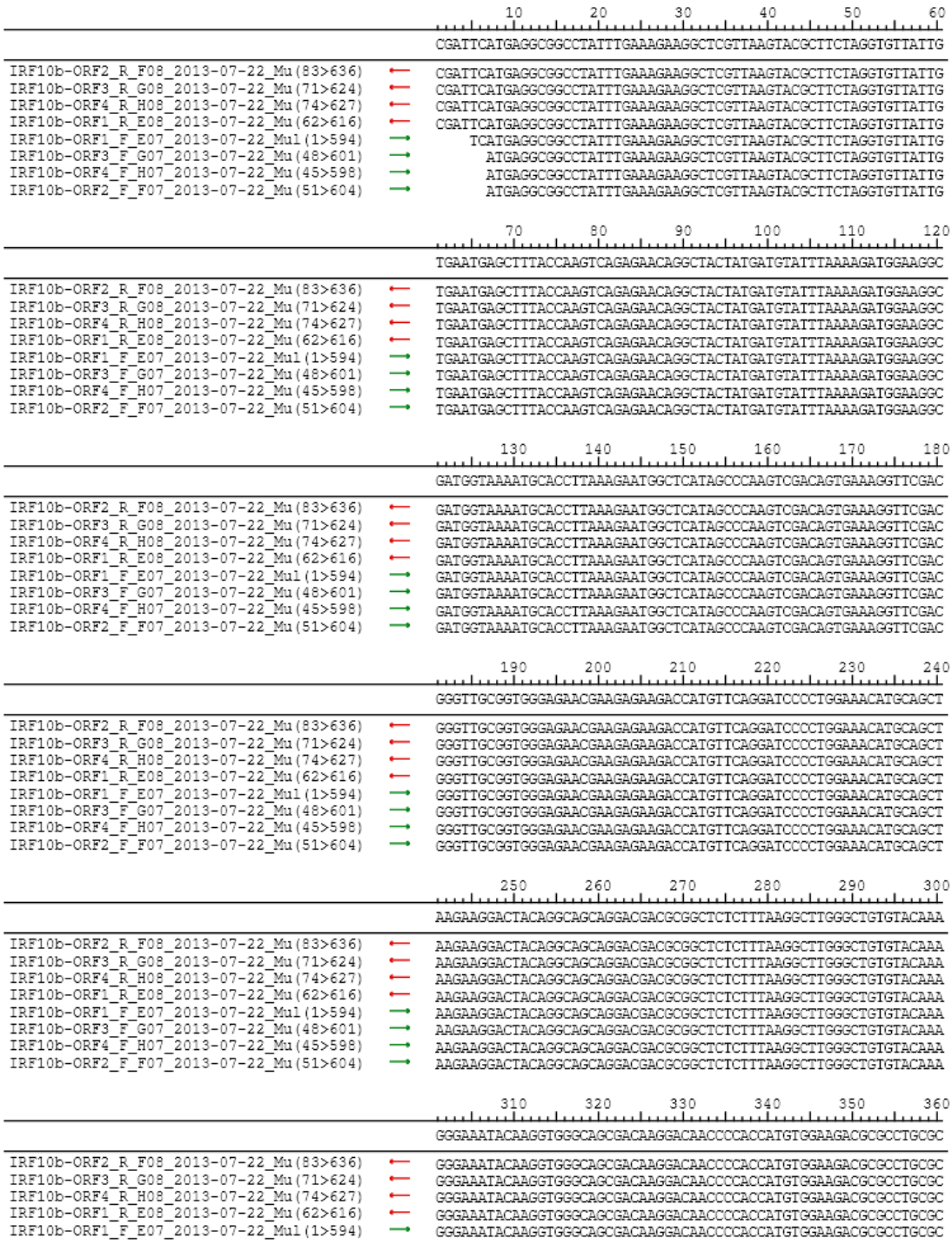

370 380 390 400 410 420
GCTGTGCACTTAACAAGAGCACAGACTTCCAGGAGGTCCCCC-ACCTGAACCAGCTGGAC
IRF10_5.1F_C05_2012-09-13_Multi(54>887) GCTGTGCACTTAACAAGAGCACAGACTTCCAGGAGGTCCCCC-ACCTGAACCAGCTGGAC
IRF10_5.4F_B05_2012-09-13_Multi(57>809) GCTGTGCACTTAACAAGAGCACAGACTTCCAGGAGGTCCCCCCACCTGAACCAGCTGGAC
IRF10_5.3F_A05_2012-09-13_Multiu(1>662) GCTGTGCACTTAACAAGAGCACAGACTTCCAGGAGGTCCCCC-ACCTGAACCAGCTGGAC
IRF10_5.1R_C11_2012-09-13_Multi(80>742) GCTGTGCACTTAACAAGAGCACAGACTTCCAGGAGGTCCCCC-ACCTGAACCAGCTGGAC
IRF10_5.3R_A11_2012-09-13_Multi(86>732) GCTGTGCACTTAACAAGAGCACAGACTTCCAGGAGGTCCCCC-ACCTGAACCAGCTGGAC
IRF10-5.3_R_G08_2012-06-14_Mult(81>625) GCTGTGCACTTAACAAGAGCACAGACTTCCAGGAGGTCCCCC-ACCTGAACCAGCTGGAC
430 440 450 460 470 480
ATCTCGGAGCCCTACAAGGTCTACCGCATCGAGTCTGACCAGAGAGCAGAGTCTGATCAG
IRF10_5.1F_C05_2012-09-13_Multi(54>887) ATCTCGGAGCCCTACAAGGTCTACCGCATCGAGTCTGACCAGAGAGCAGAGTCTGATCAG
IRF10_5.4F_B05_2012-09-13_Multi(57>809) ATCCCGGAGCCCTACAAGGTCTACCGCATCGAGTCTGACCAGAGAGCAGAGTCTGATCAG
IRF10_5.3F_A05_2012-09-13_Multiu(1>662) ATCTCGGAGCCCTACAAGGTCTACCGCATCGAGTCTGACCAGAGA
IRF10_5.1R_C11_2012-09-13_Multi(80>742) ATCTCGGAGCCCTACAAGGTCTACCGCATCGAGTCTGACCAGAGAGCAGAGTCTGATCAG
IRF10_5.3R_A11_2012-09-13_Multi(86>732) ATCTCGGAGCCCTACAAGGTCTACCGCATCGAGTCTGACCAGAGAGCAGAGTCTGATCAG
IRF10-5.3_R_G08_2012-06-14_Mult(81>625) ATCTCGGAGCCCTACAAGGTCTACCGCATCGAGTCTGACCAGAGAGCAGAGTCTGATCAG
IRF10-5.2_F_F04_2012-06-14_Multi(1>309) CATCGAGTCTGACCAGAGAGCAGAGTCTGATCAG
IRF10-5.2_R_F08_2012-06-14_Mult(81>381) CATCGAGTCTGACCAGAGAGCAGAGTCTGATCAG
490 500 510 520 530 540
ACGTACAGTCGAGTGGTCGTGGTTCAGACTGGATACGCCAGTCTCCCACAGTCTCAGCTT
IRF10_5.1F_C05_2012-09-13_Multi(54>887) ACGTACAGTCGAGTGGTCGTGGTTCAGACTGGATACGCCAGTCTCCCACAGTCTCAGCTT
IRF10_5.4F_B05_2012-09-13_Multi(57>809) ACGTACAGTCGAGTGGTCGTGGTTCAGACTGGATACGCCAGTCTCCCACAGTCTCAGCTT
IRF10_5.1R_C11_2012-09-13_Multi(80>742) ACGTACAGTCGAGTGGTCGTGGTTCAGACTGGATACGCCAGTCTCCCACAGTCTCAGCTT
IRF10_5.3R_A11_2012-09-13_Multi(86>732) ACGTACAGTCGAGTGGTCGTGGTTCAGACTGGATACGCCAGTCTCCCACAGTCTCAGCTT
IRF10-5.3_R_G08_2012-06-14_Mult(81>625) ACGTACAGTCGAGTGGTCGTGGTTCAGACTGGATACGCCAGTCTCCCACAGTCTCAGCTT
IRF10-5.2_F_F04_2012-06-14_Multi(1>309) ACGTACAGTCGAGTGGTCGTGGTTCAGACTGGATACGCCAGTCTCCCACAGTCTCAGCTT
IRF10-5.2_R_F08_2012-06-14_Mult(81>381) ACGTACAGTCGAGTGGTCGTGGTTCAGACTGGATACGCCAGTCTCCCACAGTCTCAGCTT
550 560 570 580 590 600
GCTGACCAATGGGAAAGATTTGAAGAAAGGCAAGAAGAAAGTCATGGTGCTTTGTGGAGG
IRF10_5.1F_C05_2012-09-13_Multi(54>887) GCTGACCAATGGGAAAGATTTGAAGAAAGGCAAGAAGAAAGTCATGGCGCTTTGTGGAGG
IRF10_5.4F_B05_2012-09-13_Multi(57>809) GCTGAC-AATGGGAAAGATTTGAAGAAA
IRF10_5.1R_C11_2012-09-13_Multi(80>742) GCTGACCAATGGGAAAGATTTGAAGAAAGGCAAGAAGAAAGTCATGGCGCTTTGTGGAGG
IRF10_5.3R_A11_2012-09-13_Multi(86>732) GCTGACCAATGGGAAAGATTTGAAGAAAGGCAAGAAGAAAGTCATGGTGCTTTGTGGAGG
IRF10-5.3_R_G08_2012-06-14_Mult(81>625) GCTGACCAATGGGAAAGATTTGAAGAAAGGCAAGAAGAAAGTCATGGTGCTTTGTGGAGG
IRF10-5.2_F_F04_2012-06-14_Multi(1>309) GCTGACCAATGGGAAAGATTTGAAGAAAGGCAAGAAGAAAGTCATGGTGCTTTGTGGAGG
IRF10-5.2_R_F08_2012-06-14_Mult(81>381) GCTGACCAATGGGAAAGATTTGAAGAAAGGCAAGAAGAAAGTCATGGTGCTTTGTGGAGG
610 620 630 640 650 660
GAGCACACGTACTGTGGTTCAGAGGATAGCCAGGCTCACAGTCACATCCCTCTGGACCCC
IRF10_5.1F_C05_2012-09-13_Multi(54>887) GAGCACACGTACTGTGGTTCAGAGGATAGCCAGGCTCACAGTCACATCCCTCTGGACCCC
IRF10_5.1R_C11_2012-09-13_Multi(80>742) GAGCACACGTACTGTGGTTCAGAGGATAGCCAGGCTCACAGTCACATCCCTCTGGACCCC
IRF10_5.3R_A11_2012-09-13_Multi(86>732) GAGCACACGTACTGTGGTTCAGAGGATAGCCAGGCTCACAGTCACATCCCTCTGGACCCC
IRF10-5.3_R_G08_2012-06-14_Mult(81>625) GAGCACACGTACTGTGGTTCAGAGGATAGCCAGGCTCACAGTCACATCCCTCTGGACCCC
IRF10-5.2_F_F04_2012-06-14_Multi(1>309) GAGCACACGTACTGTGGTTCAGAGGATAGCCAGGCTCACAGTCACATCCCTCTGGACCCC
IRF10-5.2_R_F08_2012-06-14_Mult(81>381) GAGCACACGTACTGTGGTTCAGAGGATAGCCAGGCTCACAGTCACATCCCTCTGGACCCC
IRF10mid2R_F09_2012-09-13_Multi(75>569) ACGTACTGTGGTTCAGAGGATAGCCAGGCTCACAGTCACATCCCTCTGGACCCC
IRF10mid3R_G09_2012-09-13_Multi(82>576) ACGTACTGTGGTTCAGAGGATAGCCAGGCTCACAGTCACATCCCTCTGGACCCC
IRF10mid1F_E03_2012-09-13_Multi(55>549) CGTACTGTGGTTCAGAGGATAGCCAGGCTCACAGTCACATCCCTCTGGACCCC
IRF10mid2F_F03_2012-09-13_Multi(57>551) CGTACTGTGGTTCAGAGGATAGCCAGGCTCACAGTCACATCCCTCTGGACCCC
IRF10mid3F_G03_2012-09-13_Multi(62>556) CGTACTGTGGTTCAGAGGATAGCCAGGCTCACAGTCACATCCCTCTGGACCCC
670 680 690 700 710 720
AGCCTCCTCAGCCCCACTCTGGCCATATCAGACTTCCGGATGGAGCTGACGCTGTTCTAC
IRF10_5.1F_C05_2012-09-13_Multi(54>887) AGCCTCCTCAGCCCCACTCTGGCCATATCAG
IRF10_5.1R_C11_2012-09-13_Multi(80>742) AGCCTCCTCAGCCCCACTCTGGCCATATCAG
IRF10_5.3R_A11_2012-09-13_Multi(86>732) AGCCTCCTCAGCCCCACTCTGGCCATATCAGACTTCC
IRF10-5.3_R_G08_2012-06-14_Mult(81>625) AGCCTCCTCAGCCCCACTCTGGCCATATCAGACTTCC
IRF10-5.2_F_F04_2012-06-14_Multi(1>309) AGCCTCCTCAGCCCCACTCTGGCCATATCAGACTTCCA
IRF10-5.2_R_F08_2012-06-14_Mult(81>381) AGCCTCCTCAGCCCCACTCTGGCCATATCAG

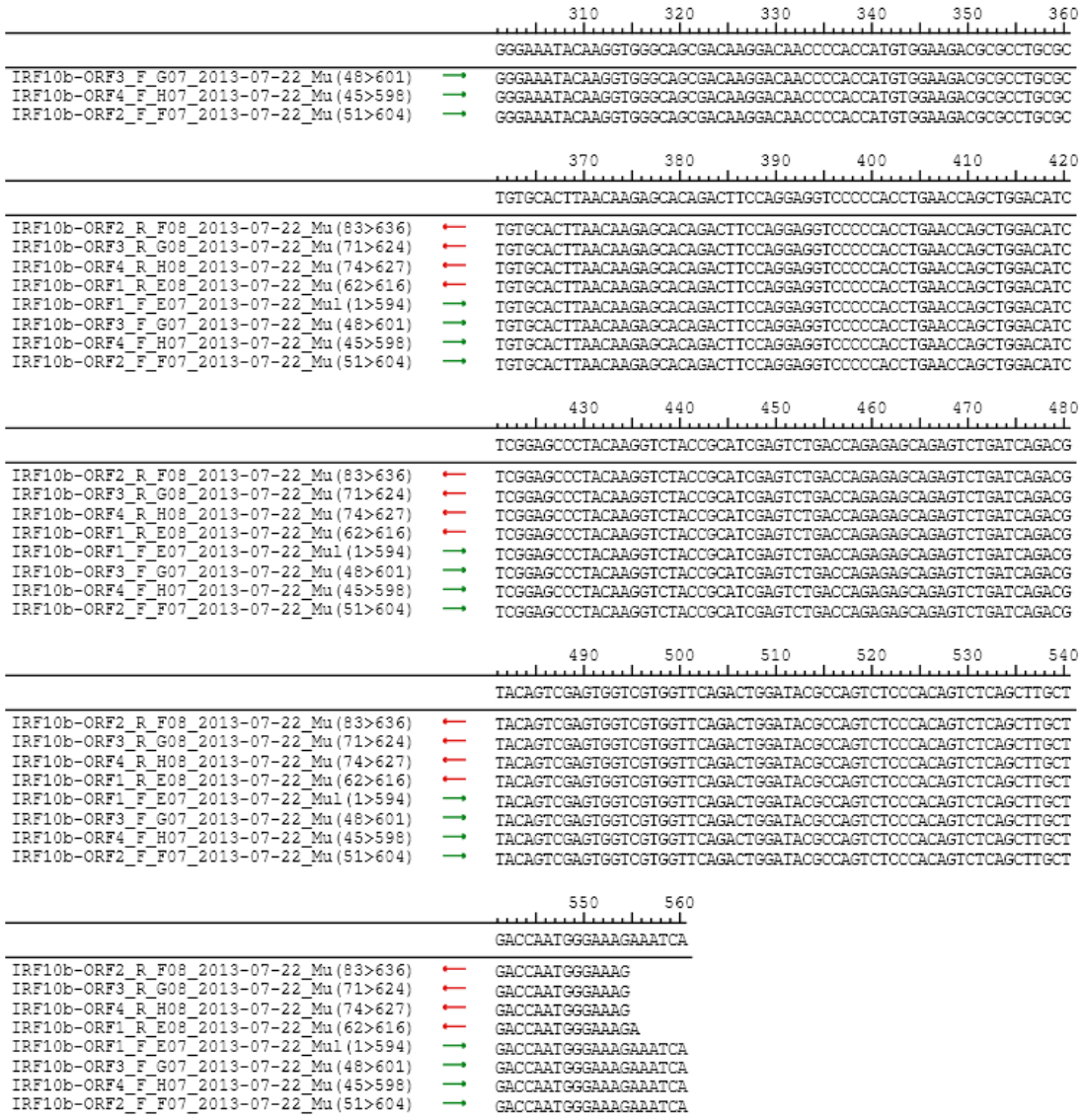

670 680 690 700 710 720
AGCCTCCTCAGCCCCACTCTGGCCATATCAGACTTCCGGATGGAGCTGACGCTGTTCTAC
IRF10mid2R_F09_2012-09-13_Multi(75>569)
IRF10mid3R_G09_2012-09-13_Multi(82>576)
IRF10mid1F_E03_2012-09-13_Multi(55>549)
IRF10mid2F_F03_2012-09-13_Multi(57>551)
IRF10mid3F_G03_2012-09-13_Multi(62>556)
730 740 750 760 770 780
CGCGGGGAGCCGGTGATGGAGCTGACCTCCAGCAGCCCAGAAGGGTGCTTCATCCTGCAG
IRF10mid2R_F09_2012-09-13_Multi(75>569)
IRF10mid3R_G09_2012-09-13_Multi(82>576)
IRF10mid1F_E03_2012-09-13_Multi(55>549)
IRF10mid2F_F03_2012-09-13_Multi(57>551)
IRF10mid3F_G03_2012-09-13_Multi(62>556)
790 800 810 820 830 840
GGCTGCGTGCCGCTGGGGAACGAGAGGATCTACGGGCCCTGCAGCGCTCAGCAGCTCTCC
IRF10mid2R_F09_2012-09-13_Multi(75>569)
IRF10mid3R_G09_2012-09-13_Multi(82>576)
IRF10mid1F_E03_2012-09-13_Multi(55>549)
IRF10mid2F_F03_2012-09-13_Multi(57>551)
IRF10mid3F_G03_2012-09-13_Multi(62>556)
850 860 870 880 890 900
CTGCCCTCCCCGGCCTCGCTGGGCCCCCTGGAGCCCGGCGTGGCCCGGGCCCTGGGTCAG
IRF10mid2R_F09_2012-09-13_Multi(75>569)
IRF10mid3R_G09_2012-09-13_Multi(82>576)
IRF10mid1F_E03_2012-09-13_Multi(55>549)
IRF10mid2F_F03_2012-09-13_Multi(57>551)
IRF10mid3F_G03_2012-09-13_Multi(62>556)
910 920 930 940 950 960
CTCCTGTCCCATCTGGAGAGGGGAGTGCTGCTCTGGGTGGCCCCGGACGGGCTGTTCATC
IRF10mid2R_F09_2012-09-13_Multi(75>569)
IRF10mid3R_G09_2012-09-13_Multi(82>576)
IRF10mid1F_E03_2012-09-13_Multi(55>549)
IRF10mid2F_F03_2012-09-13_Multi(57>551)
IRF10mid3F_G03_2012-09-13_Multi(62>556)
970 980 990 1000 1010 1020
AAGAGGTTCTGCCAGGGCCGTGTGTACTGGAGTGGGCCCCTGGCCCCGCACACCGAGAAG
IRF10mid2R_F09_2012-09-13_Multi(75>569)
IRF10mid3R_G09_2012-09-13_Multi(82>576)
IRF10mid1F_E03_2012-09-13_Multi(55>549)
IRF10mid2F_F03_2012-09-13_Multi(57>551)
IRF10mid3F_G03_2012-09-13_Multi(62>556)
1030 1040 1050 1060 1070 1080
CCCAATAAACTGGAGAGGGACAGGACCTGCAAGCTGCTGGACATGCCCGTATTTGTAAAT
IRF10mid2R_F09_2012-09-13_Multi(75>569)
IRF10mid3R_G09_2012-09-13_Multi(82>576)
IRF10mid1F_E03_2012-09-13_Multi(55>549)
IRF10mid2F_F03_2012-09-13_Multi(57>551)
IRF10mid3F_G03_2012-09-13_Multi(62>556)
IRF10_3.2R_C11_2012-07-05_Multiu(1>860)
IRF10_3.1F_B05_2012-07-05_Multiu(56>781)
IRF10_3.1F_B05_2012-07-05_Multiu(1>984)
IRF10_3.1R_B11_2012-07-05_Multiu(1>856)
IRF10_3.2R_D11_2012-09-13_Multi(88>696)
IRF10_3.2R_D11_2012-09-13_Multiu(1>935)


1090 1100 1110 1120 1130 1140
GAGCTCCAGAACTATATGCAGAGGAAAGGCCCACAACCAAACTATGAGATTGATCTCTGC
IRF10mid2R_F09_2012-09-13_Multi(75>569) GAGCTCCAGAACTATATGCAG
IRF10mid3R_G09_2012-09-13_Multi(82>576) GAGCTCCAGAACTATATGCAG
IRF10mid1F_E03_2012-09-13_Multi(55>549) GAGCTCCAGAACTATATGCAGA
IRF10mid2F_F03_2012-09-13_Multi(57>551) GAGCTCCAGAACTATATGCAGA
IRF10mid3F_G03_2012-09-13_Multi(62>556) GAGCTCCAGAACTATATGCAGA
IRF10_3.2R_C11_2012-07-05_Multiu(1>860) GAGCTCCAGAACTATATGCAGAGGAAAGGCCCACAACCAAACTATGAGATTGATCTCTGC
IRF10_3.1F_B05_2012-07-05_Multi(56>781) GAGCTCCAGAACTATATGCAGAGGAAAGGCCCACAACCAAACTATGAGATTGATCTCTGC
IRF10_3.1F_B05_2012-07-05_Multiu(1>984) GAGCTCCAGAACTATATGCAGAGGAAAGGCCCACAACCAAACTATGAGATTGATCTCTGC
IRF10_3.1R_B11_2012-07-05_Multiu(1>856) GAGCTCCAGAACTATATGCAGAGGAAAGGCCCACAACCAAACTATGAGATTGATCTCTGC
IRF10_3.2R_D11_2012-09-13_Multi(88>696) GAGCTCCAGAACTATATGCAGAGGAAAGGCCCACAACCAAACTATGAGATTGATCTCTGC
IRF10_3.2R_D11_2012-09-13_Multiu(1>935) GAGCTCCAGAACTATATGCAGAGGAAAGGCCCACAACCAAACTATGAGATTGATCTCTGC
IRF10_3.2F_D05_2012-09-13_Multi(56>664) GAGCTCCAGAACTATATGCAGAGGAAAGGCCCACAACCAAACTATGAGATTGATCTCTGC
1150 1160 1170 1180 1190 1200
TTTGGCGAGGAGTATCCCGACGCTAAAGTTTCCAAAACGATGAAGCTGATAACAGTTCAT
IRF10_3.2R_C11_2012-07-05_Multiu(1>860) TTTGGCGAGGAGTATCCCGACGCTAAAGTTTCCAAAACGATGAAGCTGATAACAGTTCAT
IRF10_3.1F_B05_2012-07-05_Multi(56>781) TTTGGCGAGGAGTATCCCGACGCTAAAGTTTCCAAAACGATGAAGCTGATAACAGTTCAT
IRF10_3.1F_B05_2012-07-05_Multiu(1>984) TTTGGCGAGGAGTATCCCGACGCTAAAGTTTCCAAAACGATGAAGCTGATAACAGTTCAT
IRF10_3.1R_B11_2012-07-05_Multiu(1>856) TTTGGCGAGGAGTATCCCGACGCTAAAGTTTCCAAAACGATGAAGCTGATAACAGTTCAT
IRF10_3.2R_D11_2012-09-13_Multi(88>696) TTTGGCGAGGAGTATCCCGACGCTAAAGTTTCCAAAACGATGAAGCTGATAACAGTTCAT
IRF10_3.2R_D11_2012-09-13_Multiu(1>935) TTTGGCGAGGAGTATCCCGACGCTAAAGTTTCCAAAACGATGAAGCTGATAACAGTTCAT
IRF10_3.2F_D05_2012-09-13_Multi(56>664) TTTGGCGAGGAGTATCCCGACGCTAAAGTTTCCAAAACGATGAAGCTGATAACAGTTCAT
1210 1220 1230 1240 1250 1260
GTGGTGCCCCTGTTTGCCATGGAACTGTTGCAGCGATTCCAGCTAGAGCGGGTCGAGGCA
IRF10_3.2R_C11_2012-07-05_Multiu(1>860) GTGGTGCCCCTGTTTGCCATGGAACTGTTGCAGCGATTCCAGCTAGAGCGGGTCGAGGCA
IRF10_3.1F_B05_2012-07-05_Multi(56>781) GTGGTGCCCCTGTTTGCCATGGAACTGTTGCAGCGATTCCAGCTAGAGCGGGTCGAGGCA
IRF10_3.1F_B05_2012-07-05_Multiu(1>984) GTGGTGCCCCTGTTTGCCATGGAACTGTTGCAGCGATTCCAGCTAGAGCGGGTCGAGGCA
IRF10_3.1R_B11_2012-07-05_Multiu(1>856) GTGGTGCCCCTGTTTGCCATGGAACTGTTGCAGCGATTCCAGCTAGAGCGGGTCGAGGCA
IRF10_3.2R_D11_2012-09-13_Multi(88>696) GTGGTGCCCCTGTTTGCCATGGAACTGTTGCAGCGATTCCAGCTAGAGCGGGTCGAGGCA
IRF10_3.2R_D11_2012-09-13_Multiu(1>935) GTGGTGCCCCTGTTTGCCATGGAACTGTTGCAGCGATTCCAGCTAGAGCGGGTCGAGGCA
IRF10_3.2F_D05_2012-09-13_Multi(56>664) GTGGTGCCCCTGTTTGCCATGGAACTGTTGCAGCGATTCCAGCTAGAGCGGGTCGAGGCA
1270 1280 1290 1300 1310 1320
GAACCGGACGTTCACACTCCCAAAGAAGCCAAGGATGAGATGTAAGGGGCCAGTTATCCA
IRF10_3.2R_C11_2012-07-05_Multiu(1>860) GAACCGGACGTTCACACTCCCAAAGAAGCCAAGGATGAGATGTAAGGGGCCAGTTATCCA
IRF10_3.1F_B05_2012-07-05_Multi(56>781) GAACCGGACGTTCACACTCCCAAAGAAGCCAAGGATGAGATGTAAGGGGCCAGTTATCCA
IRF10_3.1F_B05_2012-07-05_Multiu(1>984) GAACCGGACGTTCACACTCCCAAAGAAGCCAAGGATGAGATGTAAGGGGCCAGTTATCCA
IRF10_3.1R_B11_2012-07-05_Multiu(1>856) GAACCGGACGTTCACACTCCCAAAGAAGCCAAGGATGAGATGTAAGGGGCCAGTTATCCA
IRF10_3.2R_D11_2012-09-13_Multi(88>696) GAACCGGACGTTCACACTCCCAAAGAAGCCAAGGATGAGATGTAAGGGGCCAGTTATCCA
IRF10_3.2R_D11_2012-09-13_Multiu(1>935) GAACCGGACGTTCACACTCCCAAAGAAGCCAAGGATGAGATGTAAGGGGCCAGTTATCCA
IRF10_3.2F_D05_2012-09-13_Multi(56>664) GAACCGGACGTTCACACTCCCAAAGAAGCCAAGGATGAGATGTAAGGGGCCAGTTATCCA
1330 1340 1350 1360 1370 1380
ACTAGATGTAAGCTTCACAAGTTCACAACTACTCTCCAAGGAAATCCTTGATGTATTCCT
IRF10_3.2R_C11_2012-07-05_Multiu(1>860) ACTAGATGTAAGCTTCACAAGTTCACAACTACTCTCCAAGGAAATCCTTGATGTATTCCT
IRF10_3.1F_B05_2012-07-05_Multi(56>781) ACTAGATGTAAGCTTCACAAGTTCACAACTACTTTCCAGGGAAATCCTTGATGTATTCCT
IRF10_3.1F_B05_2012-07-05_Multiu(1>984) ACTAGATGTAAGCTTCACAAGTTCACAACTACTTTCCAAGGAAATCCTTGATGTATTCCT
IRF10_3.1R_B11_2012-07-05_Multiu(1>856) ACTAGATGTAAGCTTCACAAGTTCACAACTACTCTCCAAGGAAATCCTTGATGTATTCCT
IRF10_3.2R_D11_2012-09-13_Multi(88>696) ACTAGATGTAAGCTTCACAAGTTCACAACTACTCTCCAAGGAAATCCTTGATGTATTCCT
IRF10_3.2R_D11_2012-09-13_Multiu(1>935) ACTAGATGTAAGCTTCACAAGTTCACAACTACTCTCCAAGGAAATCCTTGATGTATTCCT
IRF10_3.2F_D05_2012-09-13_Multi(56>664) ACTAGATGTAAGCTTCACAAGTTCACAACTACTCTCCAAGGAAATCCTTGATGTATTCCT
1390 1400 1410 1420 1430 1440
AATAACCCAAGTATAACGTGACAGTTATACTTGGCAGTTGACAGTTCTGTGTAAAAGACA
IRF10_3.2R_C11_2012-07-05_Multiu(1>860) AATAACCCAAGTATAACGTGACAGTTATACTTGGCAGTTGACAGTTCTGTGTAAAAACA
IRF10_3.1F_B05_2012-07-05_Multi(56>781) AATAACCCAAGTATAACGTGACAGTTATACTTGGCAGTTGACAGTTCTGTGTAAAAGACA
IRF10_3.1F_B05_2012-07-05_Multiu(1>984) AATAACCCAAGTATAACGTGACAGTTATACTTGGCAGTTGACAGTTCTGTGTAAAAGACA
IRF10_3.1R_B11_2012-07-05_Multiu(1>856) AATAACCCAAGTATAACGTGACAGTTATACTTGGCAGTTGACAGTTCTGTGTAAAAGACA
IRF10_3.2R_D11_2012-09-13_Multi(88>696) AATAACCCAAGTATAACGTGACAGTTATACTTGGCAGTTGACAGTTCTGTGTAAAAGACA


1390 1400 1410 1420 1430 1440
AATAACCCAAGTATAACGTGACAGTTATACTTGGCAGTTGACAGTTCTGTGTAAAAGACA
IRF10_3.2R_D11_2012-09-13_Multiu(1>935) AATAACCCAAGTATAACGTGACAGTTATACTTGGCAGTTGACAGTTCTGTGTAAAAGACA
IRF10_3.2F_D05_2012-09-13_Multi(56>664) AATAACCCAAGTATAACGTGACAGTTATACTTGGCAGTTGACAGTTCTGTGTAAAAGACA
1450 1460 1470 1480 1490 1500
GAATCAAATAACTGAGGTCTGTTTGATATTAGATTTATGGTTGCTTGCTCTAATGTAAAG
IRF10_3.2R_C11_2012-07-05_Multiu(1>860) GAATCAAATAACTGAGGTCTGTTTGATATTAGATTTATGGTTGCTTGCTCTAATGTAAAG
IRF10_3.1F_B05_2012-07-05_Multi(56>781) GAATCAAATAACTGAGGTCTGTTTGATATTAGATTTATGGTTGCTTGCTTTAATGTAAAG
IRF10_3.1F_B05_2012-07-05_Multiu(1>984) GAATCAAATAACTGAGGTCTGTTTGATATTAGATTTATGGTTGCTTGCTTTAATGTAAAG
IRF10_3.1R_B11_2012-07-05_Multiu(1>856) GAATCAAATAACTGAGGTCTGTTTGATATTAGATTTATGGTTGCTTGCTCTAATGTAAAG
IRF10_3.2R_D11_2012-09-13_Multi(88>696) GAATCAAATAACTGAGGTCTGTTTGATATTAGATTTATGGTTGCTTGCTCTAATGTAAAG
IRF10_3.2R_D11_2012-09-13_Multiu(1>935) GAATCAAATAACTGAGGTCTGTTTGATATTAGATTTATGGTTGCTTGCTCTAATGTAAAG
IRF10_3.2F_D05_2012-09-13_Multi(56>664) GAATCAAATAACTGAGGTCTGTTTGATATTAGATTTATGGTTGCTTGCTCTAATGTAAAG
1510 1520 1530 1540 1550 1560
CAGTAGTGATTCTAATGTGTGTATAATTTATATTTAGAGACTTTCTACATGCCAGCGATA
IRF10_3.2R_C11_2012-07-05_Multiu(1>860) CAGTAGTGATTCTAATGTGTGTATAATTTATATTTGGAGACCTTCTACATGCCAGCGATA
IRF10_3.1F_B05_2012-07-05_Multi(56>781) CAGTAGTGATTTAATGTGTGTATAATTTATATTTAGAGACTTTTTACATGCCAGCGATA
IRF10_3.1F_B05_2012-07-05_Multiu(1>984) CAGTAGTGATTTAATGTGTGTATAATTTATATTTAGAGACTTTCTACATGCCAGCGATA
IRF10_3.1R_B11_2012-07-05_Multiu(1>856) CAGTAGTGATTCTAATGTGTGTATAATTTATATTTAGAGACTTTCTACATGCCAGCGATA
IRF10_3.2R_D11_2012-09-13_Multi(88>696) CAGTAGTGATTCTAATGTGTGTATAATTTATATTTAGAGACTTTCTACATGCCAGCGATA
IRF10_3.2R_D11_2012-09-13_Multiu(1>935) CAGTAGTGATTCTAATGTGTGTATAATTTATATTTAGAGACTTTCTACATGCCAGCGATA
IRF10_3.2F_D05_2012-09-13_Multi(56>664) CAGTAGTGATTCTAATGTGTGTATAATTTATATTTAGAGACTTTCTACATGCCAGCGATA
1570 1580 1590 1600 1610 1620
CAATATTAACAACATTCTTTTCATGTTATATTTAATCTTCTGAGTAAAGTTATTTTGAGT
IRF10_3.2R_C11_2012-07-05_Multiu(1>860) CAATATTAACAACATTCTTTTCATGTTATATTTAATCTTCTGAGTAAAGTTATTTTGAGT
IRF10_3.1F_B05_2012-07-05_Multi(56>781) CAATATTAACAACATTCTTTTCATGTTATATTTAATCTTCTGAGTAAAGTTATTTTGAGT
IRF10_3.1F_B05_2012-07-05_Multiu(1>984) CAATATTAACAACATTCTTTTCATGTTATATTTAATCTTCTGAGTAAAGTTATTTTGAGT
IRF10_3.1R_B11_2012-07-05_Multiu(1>856) CAATATTAACAACATTCTTTTCATGTTATATTTAATCTTCTGAGTAAAGTTATTTTGAGT
IRF10_3.2R_D11_2012-09-13_Multi(88>696) CAATATTAACAACATTCTTTTCATGTTATATTTAATCTTCTGAGTAAAGTTATTTTGAGT
IRF10_3.2R_D11_2012-09-13_Multiu(1>935) CAATATTAACAACATTCTTTTCATGTTATATTTAATCTTCTGAGTAAAGTTATTTTGAGT
IRF10_3.2F_D05_2012-09-13_Multi(56>664) CAATATTAACAACATTCTTTTCATGTTATATTTAATCTTCTGAGTAAAGTTATTTTGAGT
1630 1640 1650 1660 1670 1680
TAAGTGTGTTTAATGTTCTTAGTCTACTTATGAATTGTAATAATTTATGCAGTTCAATGC
IRF10_3.2R_C11_2012-07-05_Multiu(1>860) TAAGTGTGTTTAATGTTCTTAGTCTACTTATGAATTGTAATAATTTATGCAGTTCAATGC
IRF10_3.1F_B05_2012-07-05_Multi(56>781) TAAGTGTGTTTAATGTTCTTAGTCTACTTATGAATTGTAATAATTTATGCAGTTCAATGC
IRF10_3.1F_B05_2012-07-05_Multiu(1>984) TAAGTGTGTTTAATGTTCTTAGTCTACTTATGAATTGTAATAATTTATGCAGTTCAATGC
IRF10_3.1R_B11_2012-07-05_Multiu(1>856) TAAGTGTGTTTAATGTTCTTAGTCTACTTATGAATTGTAATAATTTATGCAGTTCAATGC
IRF10_3.2R_D11_2012-09-13_Multi(88>696) TAAGTGTGTTTAATGTTCTTAGTCTACTTATGAATTGTAATAATTTATGCAGTTCAATGC
IRF10_3.2R_D11_2012-09-13_Multiu(1>935) TAAGTGTGTTTAATGTTCTTAGTCTACTTATGAATTGTAATAATTTATGCAGTTCAATGC
IRF10_3.2F_D05_2012-09-13_Multi(56>664) TAAGTGTGTTTAATGTTCTTAGTCTACTTATGAATTGTAATAATTTATGCAGTTCAATGC
1690 1700 1710 1720 1730 1740
ACTGGAACAATAATCAAGTACGAAAAATAAAATCACACCACCAAAAAAAAAAAAAAAAAA
IRF10_3.2R_C11_2012-07-05_Multiu(1>860) ACTGGAACAATAATCAAGTACGAAAAATAAAATCACACCACCAAAAAAAAAAAAAAAAAA
IRF10_3.1F_B05_2012-07-05_Multi(56>781) ACTGGAAC
IRF10_3.1F_B05_2012-07-05_Multiu(1>984) ACTGGAACAATAATCAAGTACGAAAAATAAAATCACACCACCAAAAAAAAAAAAAAAAAA
IRF10_3.1R_B11_2012-07-05_Multiu(1>856) ACTGGAACAATAATCAAGTACGAAAAATAAAATCACACCACCAAAAAAAAAAAAAAAAAA
IRF10_3.2R_D11_2012-09-13_Multi(88>696) ACTGGAAC
IRF10_3.2R_D11_2012-09-13_Multiu(1>935) ACTGGAACA
IRF10_3.2F_D05_2012-09-13_Multi(56>664) ACTGGAAC
1750
AAAA
IRF10_3.2R_C11_2012-07-05_Multiu(1>860) AAA
IRF10_3.1F_B05_2012-07-05_Multiu(1>984) AAAA
IRF10_3.1R_B11_2012-07-05_Multiu(1>856) AAAA

Appendix 7: Assembly of Atlantic cod *Irf10-2* RACE and ORF PCR sequencing reads. Sequencing methods are described in section 2.1.2. Sequence data was assembled using Lasergene SeqMan Pro software (DNASTAR). Consensus sequence is indicated between horizontal lines. Note that sequences named "*Irf10b*" were renamed as *Irf10-v2* when it was determined the sequence was an *Irf10* splice variant.


10 20 30 40 50 60
CGATTCATGAGGCGGCCTATTTGAAAGAAGGCTCGTTAAGTACGCTTCTAGGTGTTATTG
IRF10b-ORF2_R_F08_2013-07-22_Mu(83>636)
IRF10b-ORF3_R_G08_2013-07-22_Mu(71>624)
IRF10b-ORF4_R_H08_2013-07-22_Mu(74>627)
IRF10b-ORF1_R_E08_2013-07-22_Mu(62>616)
IRF10b-ORF1_F_E07_2013-07-22_Mul(1>594)
IRF10b-ORF3_F_G07_2013-07-22_Mu(48>601)
IRF10b-ORF4_F_H07_2013-07-22_Mu(45>598)
IRF10b-ORF2_F_F07_2013-07-22_Mu(51>604)
70 80 90 100 110 120
TGAATGAGCTTTACCAAGTCAGAGAACAGGCTACTATGATGTATTTAAAAGATGGAAGGC
IRF10b-ORF2_R_F08_2013-07-22_Mu(83>636)
IRF10b-ORF3_R_G08_2013-07-22_Mu(71>624)
IRF10b-ORF4_R_H08_2013-07-22_Mu(74>627)
IRF10b-ORF1_R_E08_2013-07-22_Mu(62>616)
IRF10b-ORF1_F_E07_2013-07-22_Mul(1>594)
IRF10b-ORF3_F_G07_2013-07-22_Mu(48>601)
IRF10b-ORF4_F_H07_2013-07-22_Mu(45>598)
IRF10b-ORF2_F_F07_2013-07-22_Mu(51>604)
130 140 150 160 170 180
GATGGTAAAATGCACCTTAAAGAATGGCTCATAGCCCAAGTCGACAGTGAAAGGTTCGAC
IRF10b-ORF2_R_F08_2013-07-22_Mu(83>636)
IRF10b-ORF3_R_G08_2013-07-22_Mu(71>624)
IRF10b-ORF4_R_H08_2013-07-22_Mu(74>627)
IRF10b-ORF1_R_E08_2013-07-22_Mu(62>616)
IRF10b-ORF1_F_E07_2013-07-22_Mul(1>594)
IRF10b-ORF3_F_G07_2013-07-22_Mu(48>601)
IRF10b-ORF4_F_H07_2013-07-22_Mu(45>598)
IRF10b-ORF2_F_F07_2013-07-22_Mu(51>604)
190 200 210 220 230 240
GGGTTGCGGTGGGAGAACGAAGAGAAGACCATGTTCAGGATCCCCTGGAAACATGCAGCT
IRF10b-ORF2_R_F08_2013-07-22_Mu(83>636)
IRF10b-ORF3_R_G08_2013-07-22_Mu(71>624)
IRF10b-ORF4_R_H08_2013-07-22_Mu(74>627)
IRF10b-ORF1_R_E08_2013-07-22_Mu(62>616)
IRF10b-ORF1_F_E07_2013-07-22_Mul(1>594)
IRF10b-ORF3_F_G07_2013-07-22_Mu(48>601)
IRF10b-ORF4_F_H07_2013-07-22_Mu(45>598)
IRF10b-ORF2_F_F07_2013-07-22_Mu(51>604)
250 260 270 280 290 300
AAGAAGGACTACAGGCAGCAGGACGACGCGGCTCTCTTTAAGGCTTGGGCTGTGTACAAA
IRF10b-ORF2_R_F08_2013-07-22_Mu(83>636)
IRF10b-ORF3_R_G08_2013-07-22_Mu(71>624)
IRF10b-ORF4_R_H08_2013-07-22_Mu(74>627)
IRF10b-ORF1_R_E08_2013-07-22_Mu(62>616)
IRF10b-ORF1_F_E07_2013-07-22_Mul(1>594)
IRF10b-ORF3_F_G07_2013-07-22_Mu(48>601)
IRF10b-ORF4_F_H07_2013-07-22_Mu(45>598)
IRF10b-ORF2_F_F07_2013-07-22_Mu(51>604)
310 320 330 340 350 360
GGGAAATACAAGGTGGGCAGCGACAAGGACAACCCCACCATGTGGAAGACGCGCCTGCGC
IRF10b-ORF2_R_F08_2013-07-22_Mu(83>636)
IRF10b-ORF3_R_G08_2013-07-22_Mu(71>624)
IRF10b-ORF4_R_H08_2013-07-22_Mu(74>627)
IRF10b-ORF1_R_E08_2013-07-22_Mu(62>616)
IRF10b-ORF1_F_E07_2013-07-22_Mul(1>594)


310 320 330 340 350 360
GGGAAATACAAGGTGGGCAGCGACAAGGACAACCCCACCATGTGGAAGACGCGCCTGCGC
IRF10b-ORF3_F_G07_2013-07-22_Mu(48>601)
IRF10b-ORF4_F_H07_2013-07-22_Mu(45>598)
IRF10b-ORF2_F_F07_2013-07-22_Mu(51>604)
370 380 390 400 410 420
TGTGCACTTAACAAGAGCACAGACTTCCAGGAGGTCCCCCACCTGAACCAGCTGGACATC
IRF10b-ORF2_R_F08_2013-07-22_Mu(83>636)
IRF10b-ORF3_R_G08_2013-07-22_Mu(71>624)
IRF10b-ORF4_R_H08_2013-07-22_Mu(74>627)
IRF10b-ORF1_R_E08_2013-07-22_Mu(62>616)
IRF10b-ORF1_F_E07_2013-07-22_Mul(1>594)
IRF10b-ORF3_F_G07_2013-07-22_Mu(48>601)
IRF10b-ORF4_F_H07_2013-07-22_Mu(45>598)
IRF10b-ORF2_F_F07_2013-07-22_Mu(51>604)
430 440 450 460 470 480
TCGGAGCCCTACAAGGTCTACCGCATCGAGTCTGACCAGAGAGCAGAGTCTGATCAGACG
IRF10b-ORF2_R_F08_2013-07-22_Mu(83>636)
IRF10b-ORF3_R_G08_2013-07-22_Mu(71>624)
IRF10b-ORF4_R_H08_2013-07-22_Mu(74>627)
IRF10b-ORF1_R_E08_2013-07-22_Mu(62>616)
IRF10b-ORF1_F_E07_2013-07-22_Mul(1>594)
IRF10b-ORF3_F_G07_2013-07-22_Mu(48>601)
IRF10b-ORF4_F_H07_2013-07-22_Mu(45>598)
IRF10b-ORF2_F_F07_2013-07-22_Mu(51>604)
490 500 510 520 530 540
TACAGTCGAGTGGTCGTGGTTCAGACTGGATACGCCAGTCTCCCACAGTCTCAGCTTGCT
IRF10b-ORF2_R_F08_2013-07-22_Mu(83>636)
IRF10b-ORF3_R_G08_2013-07-22_Mu(71>624)
IRF10b-ORF4_R_H08_2013-07-22_Mu(74>627)
IRF10b-ORF1_R_E08_2013-07-22_Mu(62>616)
IRF10b-ORF1_F_E07_2013-07-22_Mul(1>594)
IRF10b-ORF3_F_G07_2013-07-22_Mu(48>601)
IRF10b-ORF4_F_H07_2013-07-22_Mu(45>598)
IRF10b-ORF2_F_F07_2013-07-22_Mu(51>604)
550 560
GACCAATGGGAAAGAAATCA
IRF10b-ORF2_R_F08_2013-07-22_Mu(83>636)
IRF10b-ORF3_R_G08_2013-07-22_Mu(71>624)
IRF10b-ORF4_R_H08_2013-07-22_Mu(74>627)
IRF10b-ORF1_R_E08_2013-07-22_Mu(62>616)
IRF10b-ORF1_F_E07_2013-07-22_Mul(1>594)
IRF10b-ORF3_F_G07_2013-07-22_Mu(48>601)
IRF10b-ORF4_F_H07_2013-07-22_Mu(45>598)
IRF10b-ORF2_F_F07_2013-07-22_Mu(51>604)

Appendix 8: Percent identity tables indicating similarity between Atlantic cod putative amino acid sequences. Percentages are based on alignment of sequences using Clustal Omega software (see Web References). A) Based on alignment of complete amino acid sequences. B) Based on alignment of sequences trimmed to the length of the shortest sequence (IRF10-v2; 126 AA).

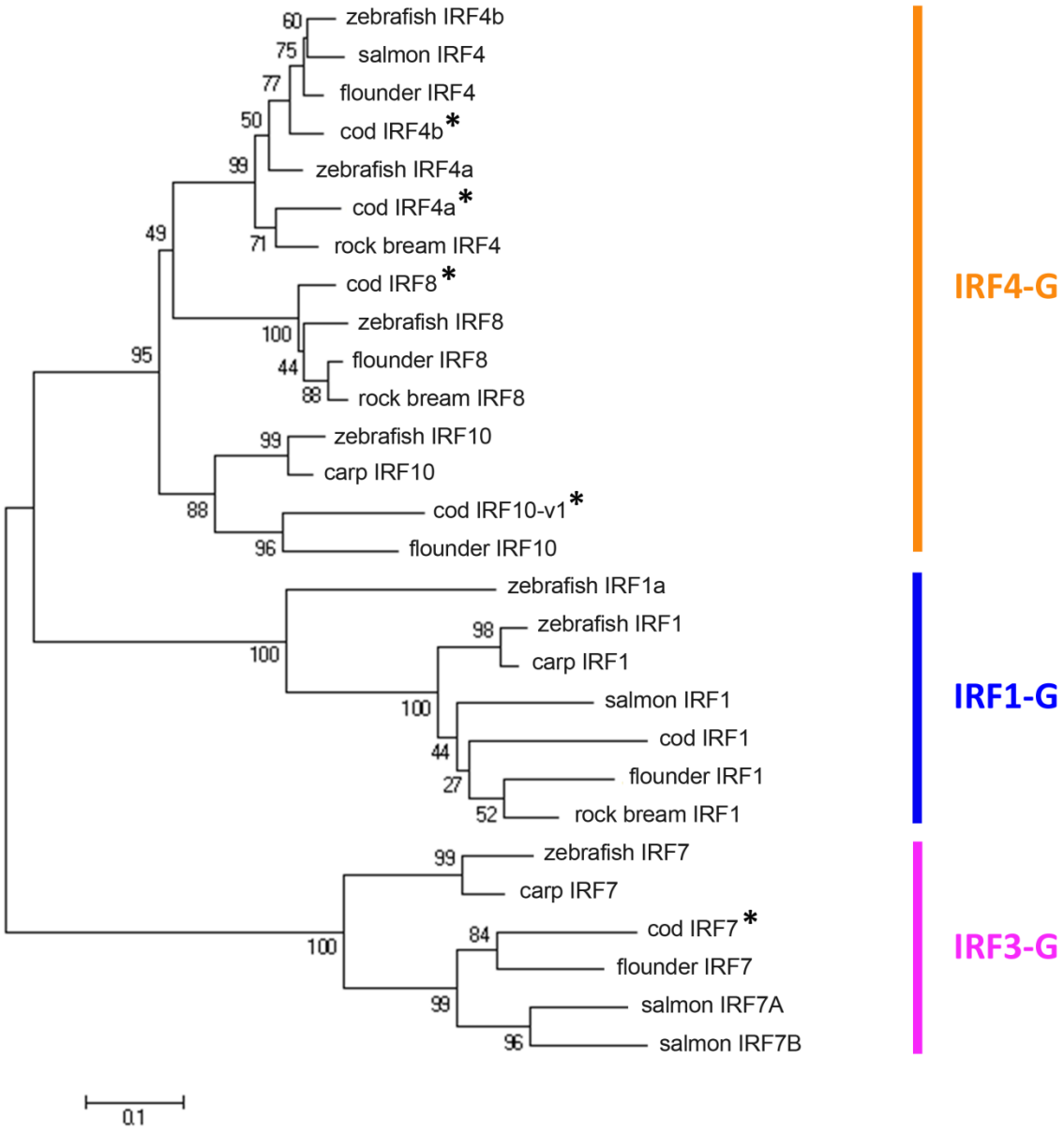
A

Irf4a	*Irf4b*	*Irf7*	*Irf8*	*Irf10-v1*	*Irf10-v2*	
-	73.6	32.3	58.7	54.1	56.0	*Irf4a*
	-	24.8	33.3	38.3	56.8	*Irf4b*
		-	26.4	28.2	31.4	*Irf7*
			-	39.7	56.8	*Irf8*
				-	93.7	*Irf10-v1*
					-	*Irf10-v2*

B

Irf4a	*Irf4b*	*Irf7*	*Irf8*	*Irf10-v1*	*Irf10-v2*	
-	81.0	34.8	65.5	59.8	59.8	*Irf4a*
	-	40.1	69.0	59.8	59.8	*Irf4b*
		-	36.1	31.4	31.4	*Irf7*
			-	57.6	56.8	*Irf8*
				-	93.7	*Irf10-v1*
					-	*Irf10-v2*

Appendix 9: Alternative phylogenetic analysis of teleost IRF family members. Putative cod amino acid sequences were aligned with IRF proteins from selected other teleost fish species using MEGA5 software (Tamura *et al.,* 2011) as in Figure 14, with sequences trimmed to the length of cod IRF4a (144 AA). Based on the multiple sequence alignment, the evolutionary history was inferred using the neighbour-joining method. The bootstrap consensus tree was constructed from 5000 replicates, where numbers at the branch points represent percentage of replicates in which the associated taxa grouped together. Branch lengths are proportional to calculated evolutionary distances.


zebrafish IRF4b
salmon IRF4
flounder IRF4
cod IRF4b*
zebrafish IRF4a
cod IRF4a*
rock bream IRF4
cod IRF8*
zebrafish IRF8
flounder IRF8
rock bream IRF8
zebrafish IRF10
carp IRF10
cod IRF10-v1*
flounder IRF10
zebrafish IRF1a
zebrafish IRF1
carp IRF1
salmon IRF1
cod IRF1
flounder IRF1
rock bream IRF1
zebrafish IRF7
carp IRF7
cod IRF7*
flounder IRF7
salmon IRF7A
salmon IRF7B
0.1
IRF4-G
IRF1-G
IRF3-G