

**Full characterization and transcript expression profiling of the interferon regulatory factor (IRF) gene family in Atlantic cod (*Gadus morhua*)**

Authors: Sabrina M. Inkpen<sup>1</sup> (s.inkpen@mun.ca), Monica H. Solbakken<sup>2</sup> (m.h.solbakken@ibv.uio.no), Sissel Jentoft<sup>2</sup> (sissel.jentoft@ibv.uio.no), Khalil Eslamloo<sup>1</sup> (keslamloo@mun.ca), Matthew L. Rise<sup>1\*</sup> (mrise@mun.ca)

<sup>1</sup>Department of Ocean Sciences, Memorial University of Newfoundland, NL, Canada, A1C 5S7

<sup>2</sup>Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, Oslo, Norway

\*Corresponding author: Dr. Matthew L. Rise (email: mrise@mun.ca, phone: 709-864-7478)

## Abstract

1 Atlantic cod (*Gadus morhua*) represents a unique immune system among teleost fish,  
2 making it a species of interest for immunological studies, and especially for investigating the  
3 evolutionary history of immune gene families. The interferon regulatory factor (IRF) gene family  
4 encodes transcription factors which function in the interferon pathway, but also in areas  
5 including leukocyte differentiation, cell growth, autoimmunity, and development. We previously  
6 characterized several IRF family members in Atlantic cod (*Irf4a*, *Irf4b*, *Irf7*, *Irf8*, and two *Irf10*  
7 splice variants) at the cDNA and putative amino acid levels, and in the current study we took  
8 advantage of the new and improved Atlantic cod genome assembly in combination with rapid  
9 amplification of cDNA ends (RACE) to characterize the remaining family members (i.e. *Irf3*,  
10 *Irf5*, *Irf6*, *Irf9*, and two *Irf2* splice variants). Real-time quantitative PCR (QPCR) was used to  
11 investigate constitutive expression of all IRF transcripts during embryonic development,  
12 suggesting several putative maternal transcripts, and potential stage-specific roles. QPCR studies  
13 also showed 11 of 13 transcripts were responsive to stimulation with poly(I:C), while 6 of 13  
14 transcripts were responsive to lipopolysaccharide (LPS) in Atlantic cod head kidney  
15 macrophages, indicating roles for cod IRF family members in both antiviral and antibacterial  
16 responses. This study is the first to investigate expression of the complete IRF family in Atlantic  
17 cod, and suggests potential novel roles for several of these transcription factors within immunity  
18 as well as in early development of this species.

19

## 20 **Keywords:**

21 Atlantic cod, interferon regulatory factor, IRF, immune response, macrophage, development.

22

23

## 24        **1. Introduction**

25            The interferon (IFN) signalling pathway is a vital part of the vertebrate innate immune  
26 response to pathogens, and thus the study of genes involved in this complex process is important  
27 to our understanding of immune system function. Members of the interferon regulatory factor  
28 (IRF) family may either positively or negatively regulate the expression of IFNs and interferon  
29 stimulated genes (ISGs), playing an important role in the innate antiviral response (Honda and  
30 Taniguchi, 2006; Ozato *et al.*, 2007). The IRF gene family may be divided into four sub-groups:  
31 IRF1-G (*Irf1*, *Irf2*), IRF3-G (*Irf3*, *Irf7*), IRF4-G (*Irf4*, *Irf8*, *Irf9*, *Irf10*), and IRF5-G (*Irf5*, *Irf6*)  
32 based on molecular phylogenetic analyses, reflecting expansion over evolutionary history and  
33 diversification of function (Nehyba *et al.*, 2002; 2009). In most vertebrates studied, at least nine  
34 IRF genes (*Irf1*-*Irf9*) have been observed; a tenth family member (*Irf10*) is present in several  
35 avian and fish species, and another member of the IRF1-G sub-group (*Irf11* or *Irf1b*) has been  
36 identified in zebrafish (*Danio rerio*) and several other teleost fish (Stein *et al.*, 2007; Huang *et*  
37 *al.*, 2010; Huang *et al.*, 2017). The presence of additional paralogues of some IRF family  
38 members in several teleost species [e.g. *Irf4a* and *Irf4b* in zebrafish (Stein *et al.*, 2007)] is likely  
39 due to the whole genome duplication in the teleost lineage which occurred shortly after their  
40 divergence from lobe-finned fishes (Amores *et al.*, 1998). Additional paralogues of IRF family  
41 genes may be present in salmonid species (e.g. IRF7A, IRF7B in Atlantic salmon, *Salmo salar*)  
42 as a result of the more recent genome duplication in that lineage (Lien *et al.*, 2016 and references  
43 therein), although to our knowledge the entire IRF family has not yet been extensively studied in  
44 any salmonid species. Expansion of the vertebrate IRF gene family over evolutionary time is  
45 likely a factor in the observed diversity of function of these genes.

46           The amino (N) terminus of all vertebrate IRF proteins is well conserved, consisting of a  
47 DNA-binding domain (DBD) approximately 115-120 amino acids in length, including five  
48 conserved tryptophan (Trp) residues (Taniguchi *et al.*, 2001). The DBD forms a helix-loop-helix  
49 motif and recognizes the IFN stimulated response element (ISRE) in target genes, which has the  
50 consensus sequence A/GNGAAANNGAAACT (Darnell *et al.*, 1994) and is found in the  
51 promoters of Type I IFNs and many ISGs. The carboxyl (C) terminus of IRF family members is  
52 less conserved, except for an association module called the IRF associated domain (IAD) present  
53 in all except *Irf1* and *Irf2*, which is important for interaction with other transcription factors  
54 (Meraro *et al.*, 1999).

55           In addition to their role in the IFN signalling pathway, IRF family members are known to  
56 function in areas such as immune system regulation, growth, and development [see Ozato *et al.*,  
57 2007; Savitsky *et al.*, 2010; Matta *et al.*, 2017 for reviews]. For example, while IRF1 is known as  
58 a transcriptional activator of IFN $\alpha/\beta$ , mammalian studies have shown it is also required for DNA  
59 damage-induced apoptosis, and is thus known as a tumor suppressor (Tanaka *et al.*, 1996). IRF2  
60 can be said to act opposite to IRF1, negatively regulating type I IFN responses (Honda and  
61 Taniguchi 2006) and has been shown to have pro-oncogenic activity (reviewed by Yanai *et al.*,  
62 2012). Comparatively little is known about IRF functions outside of innate immunity in fish  
63 species, however, and further investigation could provide valuable insights into fish health and  
64 development.

65           In contrast to all other IRF family members, IRF6 has not been shown to have a role in  
66 IFN pathway regulation, though it is known to play a crucial role in the differentiation of  
67 epithelia: mutations in human IRF6 lead to Van der Woude syndrome, or cleft palate (Kondo *et*  
68 *al.*, 2002), and in zebrafish and the frog *Xenopus laevis*, *Irf6* has been shown to be a maternal

69 transcript necessary for epithelial differentiation (Ben *et al.*, 2005; Sabel *et al.*, 2009). While  
70 little is known about the roles of other IRF family members in early development of teleosts, *Irf7*  
71 in *Gadus morhua* (*GmIrf7*) has been identified as a maternal transcript with a wide range of  
72 expression in egg batches from different females (Rise *et al.*, 2014). Transcript expression of  
73 Atlantic cod *Irf1* (*GmIrf1*) and *GmIrf7* has also been observed to increase in early segmentation  
74 stage cod embryos, suggesting that these genes may have stage-specific functions during early  
75 development (Rise *et al.*, 2012). Thus, in the current study we investigate the expression of all  
76 IRF transcripts during early development, to identify any other maternal transcripts and predict  
77 additional stage-specific roles.

78 In earlier reports we characterized *GmIrf1* (Feng *et al.*, 2009), *GmIrf4a*, *GmIrf4b*,  
79 *GmIrf7*, *GmIrf8* and *GmIrf10* (Inkpen *et al.*, 2015) in Atlantic cod, and showed that several of  
80 these genes respond to immune stimulation (Rise *et al.*, 2008; Hori *et al.*, 2012). All previously  
81 characterized IRFs in cod showed an expression response to stimulation with the virus-like  
82 pathogen-associated molecular pattern (PAMP) polyriboinosinic polyribocytidylic acid  
83 [poly(I:C)], a synthetic double stranded RNA. *GmIrf4b*, *GmIrf7*, *GmIrf8* and *GmIrf10* were also  
84 responsive to stimulation with killed *Aeromonas salmonicida* (ASAL); furthermore, these  
85 responses were seen to be modulated by elevated temperature (Hori *et al.*, 2012; Inkpen *et al.*,  
86 2015). Others have shown that several cod *Irf* transcripts also respond to nervous necrosis virus  
87 infection in the brain, based on microarray analysis (Krasnov *et al.*, 2013), and *GmIrf1* was  
88 observed to respond to stimulation with heat-killed *Vibrio anguillarum* in the spleen (Caipang *et*  
89 *al.*, 2009).

90 IRF genes have been well characterized in several fish species in the past ~ decade [e.g.  
91 in mandarin fish, *Siniperca chuasti* (Sun *et al.*, 2007), rainbow trout, *Oncorhynchus mykiss*

92 (Holland *et al.*, 2008), Atlantic salmon, (Bergan *et al.*, 2010), rock bream, *Oplegnathus fasciatus*  
93 (Bathige *et al.*, 2012), orange spotted grouper, *Epinephelus coioides* (Huang *et al.*, 2017), blunt  
94 snout bream, *Megalobrama amblycephala* (Zhan *et al.*, 2017), and zebrafish (Ben *et al.*, 2005;  
95 Xiang *et al.*, 2010; Li *et al.*, 2011)]; but most studies have investigated only one or two  
96 transcripts at a time. However, the number of genome assemblies generated for non-model  
97 organisms including teleost fish has been steadily increasing, which has facilitated more in-depth  
98 characterizations of gene families of interest, contributing to our understanding of the evolution  
99 of the unique immune system of Atlantic cod and its relatives (Star *et al.*, 2011; Malmstrøm *et*  
100 *al.*, 2016). These studies show that the Gadiformes have lost important genes of the major  
101 histocompatibility complex (MHC) II pathway, and harbour expansions of several important  
102 immune gene families such as the MHC I and specific Toll-like receptors (TLRs) (Star *et al.*,  
103 2011; Malmstrøm *et al.*, 2016; Solbakken *et al.*, 2016; 2017). The characterization of all Atlantic  
104 cod IRFs will allow us to determine if the composition of this gene family is similar to or  
105 different from other species as they become more well studied within the teleost lineage.

106 In the current study, the remaining IRF family members (*GmIrf2*, *GmIrf3*, *GmIrf5*,  
107 *GmIrf6*, *GmIrf9*) were predicted using the most recent Atlantic cod genome assembly [i.e.,  
108 gadMor2, (Torresen *et al.*, 2017)], and verified using the same methods as the previously  
109 characterized Atlantic cod IRF transcripts [i.e., rapid amplification of cDNA ends (RACE), TA-  
110 cloning, and sequencing (Feng *et al.*, 2009; Inkpen *et al.*, 2015)]. Constitutive expression of all  
111 transcripts not previously studied was investigated in adult tissues using RT-PCR, and real-time  
112 quantitative PCR (QPCR) was used to observe expression of all cod IRFs during embryonic  
113 development, and to determine the expression response to viral and bacterial PAMP stimulation  
114 in isolated macrophages.

115

## 116 **2. Materials and Methods**

### 117 *2.1 cDNA characterization of cod IRF paralogues*

118 Parologue-specific RACE primers were designed for *GmIrf2*, *GmIrf3*, *GmIrf5*, *GmIrf6*,  
119 and *GmIrf9* (Table 1) based on predicted cDNA sequences from the Ensembl database (see web  
120 references), and on expressed sequence tag (EST) evidence as described previously (Inkpen *et*  
121 *al.*, 2015).

122 Column-purified RNA was pooled from two adult cod spleen samples as well as  
123 poly(I:C)-stimulated macrophages isolated from head kidney (sampled at 24 h post-stimulation)  
124 and used for RACE. The experimental design, sampling procedures, and method of RNA  
125 preparation are described in sections 2.3 and 2.4. Five micrograms of pooled RNA were used to  
126 prepare RACE-ready cDNA using the GeneRacer Kit (Invitrogen, Burlington, ON, Canada),  
127 according to the manufacturer's instructions. RACE was carried out in 50  $\mu$ L reactions  
128 containing 1  $\mu$ L (1 U/ $\mu$ L) DyNAzyme polymerase (Thermo Scientific, Ottawa, ON, Canada),  
129 DyNAzyme EXT buffer (1X final concentration), and either reverse gene specific primer (GSP)  
130 and GeneRacer 5' primer or forward GSP and GeneRacer 3' primer for 5' RACE or 3' RACE,  
131 respectively. Touchdown PCR was carried out using an initial denaturation at 94°C for 2 min  
132 followed by 40 cycles of [30 s at 94°C; 30 s at 70°C→60°C, decreasing 0.3°C per cycle; 2 min at  
133 72°C] and a final extension of 8 min at 72°C. Approximate size of PCR products was verified by  
134 agarose gel electrophoresis, and DNA bands were excised and purified using the QIAquick Gel  
135 Extraction Kit (QIAGEN, Mississauga, ON, Canada) as previously described (Inkpen *et al.*,  
136 2015).

137 RACE products were ligated into pGEM-T Easy vector (Promega, Madison, WI, USA)  
138 and replicated in Subcloning Efficiency DH5 $\alpha$  chemically competent cells (Invitrogen) as  
139 previously described (Inkpen *et al.*, 2015). Colonies containing recombinant plasmids with  
140 inserts were obtained by blue/white selection on LB agar/carbenicillin (50  $\mu$ g/mL) plates  
141 containing 40  $\mu$ L of 40 mg/mL X-gal (Sigma, Oakville, ON, Canada), and then grown overnight  
142 at 37°C with shaking at 225 rpm in liquid LB media containing 50  $\mu$ g/mL carbenicillin. Plasmid  
143 DNA was isolated using the QIAprep Spin Miniprep Kit (QIAGEN), following the  
144 manufacturer's instructions. Each RACE product was sequenced in both directions using M13F  
145 and M13R primers with help of the staff at the GaP (Genomics and Proteomics) facility,  
146 CREAT network, Memorial University of Newfoundland, as described previously (Inkpen *et*  
147 *al.*, 2015). Sequencing reactions were processed by capillary electrophoresis using the Applied  
148 Biosystems 3730 DNA Analyzer.

149

## 150 2.2 Sequence analysis and comparison to genome assembly

151 Sequence data based on RACE was compiled and analyzed using Lasergene SeqMan Pro  
152 software V.7.1.0 (DNASTAR, Inc., Madison, WI, USA) and Sequencher V5.4.6 (Gene Codes  
153 Corporation, Ann Arbor, MI, USA). Amino acid sequences for each paralogue were predicted  
154 based on cDNA sequence using the ExPASy Translate tool (see Web References). Separately,  
155 the Atlantic cod genome assembly gadMor2 (Torresen *et al.*, 2017) was scanned for all IRF-like  
156 sequences based on the previously characterized family members. For each gene, the putative  
157 transcript and genomic sequences were aligned using the BLAST 2 Sequences alignment tool  
158 (see Web References) to verify the transcript assembly and annotate intron/exon boundaries.



159 Following analysis, RACE PCR, cloning and sequencing of *GmIrf3*, *GmIrf5*, and *GmIrf9* were  
160 repeated as the 5' and 3' untranslated regions (UTRs) of these transcripts appeared to be  
161 incomplete.

162 Transcripts were further validated using the BLASTx alignment search tool with default  
163 parameters (see Web References) to compare with IRF protein sequences of other teleost fish in  
164 which this gene family has been fully or partially characterized. Homologous IRF protein  
165 sequences from other teleost species [zebrafish, Atlantic salmon, rainbow trout, Japanese  
166 flounder (*Paralichthys olivaceus*), grass carp (*Ctenopharyngodon idella*), rock bream] were  
167 collected from the NCBI non-redundant (nr) protein database using the BLASTx alignment  
168 search tool and Atlantic cod *Irf* transcripts as queries (see Supplemental Table 1). Predicted IRF  
169 amino acid sequences were aligned with the ClustalW function of MEGA7 software (Kumar *et*  
170 *al.*, 2016). Based on the multiple sequence alignment, a phylogenetic tree was constructed using  
171 the neighbour-joining method in MEGA7, where the bootstrap consensus tree was constructed  
172 from 2000 replicates, and evolutionary distances were computed using the Poisson correction  
173 method.

174

### 175 *2.3 RT-PCR expression analysis in Atlantic cod tissues*

176 All procedures involving sampling of early life stage or adult cod were conducted with  
177 approval of Memorial University of Newfoundland's Institutional Animal Care Committee,  
178 following the guidelines of the Canadian Council on Animal Care. In this experiment, two  
179 juvenile Atlantic cod were individually removed from a flow-through tank (5.0-6.5°C, 95-110%  
180 oxygen saturation) and quickly euthanized by a lethal dose of tricaine methanesulfonate (TMS;  
181 400 mg/L; Syndel Laboratories, Qualicum Beach, BC, Canada). Tissue samples (blood, brain,

182 eye, fin, gill, gonad, head kidney, heart, hindgut, liver, midgut, posterior kidney, pyloric caecum,  
183 dorsal skeletal muscle, dorsal and ventral skin, spleen, stomach) were collected and immediately  
184 flash-frozen in liquid nitrogen before storage at -80°C.

185 Total RNA was extracted from each tissue using TRIzol reagent (Invitrogen), and RNA  
186 cleanup and cDNA synthesis were carried out as previously described (Inkpen *et al.*, 2015).  
187 Briefly, RNA was treated with DNaseI (QIAGEN) and column-purified using the RNeasy Mini  
188 Kit (Invitrogen) following the manufacturer's instructions, and quality (A260/280 and  
189 A260/230) and concentration were assessed by Nanodrop (Thermo Scientific)  
190 spectrophotometry. One microgram of each clean RNA sample was used for cDNA synthesis  
191 with M-MLV (Moloney Murine Leukemia Virus) reverse transcriptase (Invitrogen), following  
192 the manufacturer's instructions.

193 Parologue-specific primers for *GmIrf2-v1*, *GmIrf2-v2*, *GmIrf3*, *GmIrf5*, *GmIrf6* and  
194 *GmIrf9* were designed using Primer3 software (see Web References) for use in all RT-PCR and  
195 QPCR experiments (Table 2). RT-PCR was carried out using TopTaq DNA polymerase  
196 (QIAGEN) as described in Inkpen *et al.* (2015). As in that study, *Ef1 $\alpha$*  (elongation factor 1  $\alpha$ )  
197 was used as a reference gene, and for each primer set, a no-template control containing all  
198 reaction components except cDNA was also run. PCR products were electrophoretically  
199 separated on 1.7% agarose/TAE gels (stained with ethidium bromide) alongside 1 Kb Plus DNA  
200 Ladder (Invitrogen).

201

#### 202 2.4 QPCR expression analysis in embryonic and larval development

203 Adult (broodstock) Atlantic cod involved in this study were handled by the staff of the  
204 Dr. Joe Brown Aquatic Research Building (JBARB) at the Ocean Sciences Centre of Memorial

205 University. After communal spawning, fertilized eggs were collected in 3 batches and ozonated  
206 at 1.5-2 ppm for 1.5 min and placed in three 250 L incubators with air stones. Temperature was  
207 recorded daily and maintained at 5-7 °C for the duration of sampling, and non-buoyant dead  
208 embryos and/or shells from hatched larvae were removed daily by draining from the bottom of  
209 each incubator before sampling. Each day from 0 to 17 days post-fertilization (dpf), 250 µL of  
210 embryos from each incubator were placed in a 1.5 mL RNase-free microcentrifuge tube, flash  
211 frozen in liquid nitrogen and stored at -80°C. Embryos were also observed under a light  
212 microscope to estimate developmental stage and confirm synchronous development, and pictures  
213 were taken of representative samples for each day (Supplemental Figure 1).

214 Total RNA was extracted, purified, and quality checked as above, and cDNA was  
215 prepared using M-MLV reverse transcriptase as above. Parologue-specific primers (Table 2) for  
216 *GmIrf4a*, *GmIrf4b*, *GmIrf7*, *GmIrf8*, *GmIrf10-v1*, and *GmIrf10-v2* were developed and used in  
217 previous studies (Inkpen *et al.*, 2015). All primer pairs were quality tested for the current study  
218 using pooled cDNA (i.e. 2 µg each of one sample from each time-point). For each assay, a 5-  
219 point, 4-fold dilution standard curve (starting with cDNA corresponding to 10 ng input RNA)  
220 was used to calculate amplification efficiency as described in Pfaffl (2001). Triplicate reactions  
221 were carried out for all standard curves, controls and experimental samples. The same pooled  
222 cDNA was also used as a linker in the QPCR study; this sample was included in all plates, and  
223 plates were only included for analysis if linker Ct values were within 1 cycle of each other.

224 All QPCR analyses were performed using SYBR Green chemistry and the 7500 Fast  
225 Real-Time PCR system (Applied Biosystems). Cycling conditions consisted of 1 cycle of 50°C  
226 for 2 min, 1 cycle of 95°C for 10 min, followed by 40 cycles of (95°C for 15 s and 60°C for 1  
227 min), including a final melt curve stage for primer quality testing assays. QPCR assays were

228 carried out in 13  $\mu\text{L}$  reactions containing 6.5  $\mu\text{L}$  Power SYBR Green master mix (Applied  
229 Biosystems), 0.52  $\mu\text{L}$  each of forward and reverse primers (50 nM final concentration), 3.46  $\mu\text{L}$   
230 nuclease-free water (Invitrogen) and 2.5  $\mu\text{L}$  cDNA (corresponding to 12.5 ng input RNA). All  
231 samples were run as triplicate technical replicates, and no-template controls were included for  
232 each primer set in each plate. To confirm the absence of any genomic DNA, a “no reverse  
233 transcription” (no-RT) control was also included in which a mock cDNA synthesis reaction  
234 using the linker RNA pool was carried out with all components except reverse transcriptase. The  
235 no-RT reaction product was run in triplicate (2.5  $\mu\text{L}$  of the no-RT reaction, as with cDNA  
236 samples), for each primer set, and no amplification was observed. Based on analysis of 4  
237 potential normalizers with the geNorm algorithm of qbase+ software (Biogazelle Zwijnaarde,  
238 Belgium), *Tubb2* (tubulin beta 2A) and *EIF3* (eukaryotic initiation factor 3) were chosen as  
239 normalizers (geNorm  $M < 0.5$ ). Gene of interest (GOI) expression was normalized to the  
240 geometric mean of *Tubb2* and *EIF3* expression, and relative quantity (RQ) of each QPCR target  
241 transcript for each individual was calculated using the 7500 Fast Software (Applied Biosystems)  
242 for Comparative  $C_T$  ( $\Delta\Delta C_T$ ) analysis (Livak and Schmittgen, 2001), incorporating calculated  
243 amplification efficiencies for each primer pair (Table 2). The sample with lowest normalized  
244 expression for each gene of interest was set as the calibrator (RQ set as 1.0) for analysis of that  
245 gene. RQ values were plotted using Prism v5.0 (GraphPad Software Inc., La Jolla, CA, USA).

246

#### 247 *2.5 QPCR expression analysis in immune-stimulated adherent head kidney macrophages*

248 Atlantic cod macrophages were isolated as described by Eslamloo *et al.* (2016). Briefly,  
249 six adult Atlantic cod were euthanized as above, and the head kidney was removed by dissection  
250 and transferred to Leibovitz’s 15+ [L-15 (Gibco, Carlsbad, CA) culture medium supplemented

251 with 2 mM L-glutamine, 4.2 mM NaHCO<sub>3</sub>, 25 mM HEPES, 1.8 mM glucose, 20 U ml<sup>-1</sup> heparin  
252 (Sigma-Aldrich, St. Louis, MO) and 100 U ml<sup>-1</sup> penicillin and 100 µg ml<sup>-1</sup> streptomycin  
253 (Gibco)]. The cells were passed through 100 µm nylon cell strainers (Fisherbrand™, Thermo  
254 Fisher Scientific, Waltham, MA), and the resulting cell suspension was centrifuged at 300 × g for  
255 40 min at 4°C on a 25%/51% Percoll gradient (GE Healthcare, Uppsala, Sweden) containing  
256 Hank's Balanced Salt Solution (HBSS; Thermo Fisher Scientific). The leukocytes were collected,  
257 and washed in L-15+, and placed in 35 mm culture dishes (Corning™, Corning, NY) at a density  
258 of 3 × 10<sup>7</sup> cells per dish (in 2 ml L-15+ containing 1% fetal bovine serum (FBS; Gibco) and no  
259 heparin). After ~24 h culture at 10°C, the cells were washed 3 times to remove non-adherent  
260 cells, and samples of the adherent cells were observed microscopically to confirm the majority of  
261 cells showed macrophage morphology. Media was then replaced with 2 ml L-15+ containing 20  
262 µg/mL poly(I:C) in phosphate buffered saline (PBS; Gibco), as in Eslamloo *et al.* (2016), 20  
263 µg/mL LPS (*E. coli* O26:B6; Sigma-Aldrich; as in Seppola *et al.*, 2015) in PBS, or 20 µl PBS.  
264 Sampling was carried out at 12, 24, 48, and 72 h post-stimulation (HPS), using aseptic  
265 techniques as described by Eslamloo *et al.* (2018), and samples were immediately stored at -  
266 80°C. Total RNA was extracted using *mirVana* miRNA isolation kit (Ambion, Life  
267 Technologies, Carlsbad, CA) according to the manufacturer's instruction. RNA quality was  
268 determined by agarose gel electrophoresis and Nanodrop spectrophotometry, and cDNA was  
269 prepared using M-MLV reverse transcriptase as previously described. QPCR reaction setup was  
270 performed as described above, except using 2 µL cDNA (corresponding to 10 ng input RNA).  
271 GOI expression was normalized to the geometric mean of *EF1α* and *rplp1* (acidic ribosomal  
272 protein P1) expression, (representing lowest geNorm M values (<0.4), following analysis of 4  
273 potential normalizers), and RQs were calculated, analyzed and plotted as above.

274

### 275 3. Results

#### 276 3.1 Characterization of *GmIrf2-v1*, *GmIrf2-v2*, *GmIrf3*, *GmIrf5*, *GmIrf6* and *GmIrf9* 277 *cDNA sequences*

278 Two *GmIrf2* splice variants were identified in this study. Assembly of *GmIrf2-v1*  
279 sequencing reads produced a 1354 bp cDNA sequence (excluding poly-A tail) (Figure 1A,  
280 Supplemental Figure 2). The sequence consists of a 1062 bp (353 AA) open reading frame  
281 (ORF), a 258 bp 5' UTR, and a 33 bp 3'-UTR. Assembly of sequencing reads for the second *Irf2*  
282 variant (*GmIrf2-v2*) produced a 1572 bp cDNA sequence (excluding poly-A tail) (Figure 1B,  
283 Supplemental Figure 3), consisting of a 723 bp (240 AA) ORF, a 258 bp 5'-UTR, and a 591 bp  
284 3'-UTR. The two transcripts are identical until exon 6, where *GmIrf2-v2* retains part of intron 6  
285 and reaches a stop codon.

286 *GmIrf3* sequencing reads were assembled to produce a 1323 bp transcript, consisting of a  
287 1083 bp (360 AA) ORF, a 74 bp 5'-UTR, and 166 bp of the 3'-UTR (Figure 1C, Supplemental  
288 Figure 4). *GmIrf5* sequencing reads were assembled to produce a 1487 bp cDNA sequence that  
289 contains a 1233 bp (410 AA) ORF, a 140 bp 5'-UTR, and 114 bp of the 3'-UTR (Figure 1D;  
290 Supplemental Figure 5). Assembly of *GmIrf6* sequencing reads produced a 2475 bp cDNA  
291 sequence that consisted of a 1431 bp (476 AA) ORF, a 207 bp 5'-UTR and a 837 bp 3'-UTR  
292 (Figure 1E; Supplemental Figure 6). *GmIrf9* sequencing reads were assembled into a 1989 bp  
293 transcript, consisting of a 1284 bp ORF (427 AA), a 397 bp 5'-UTR, and a 308 bp 3'-UTR  
294 (Figure 1F; Supplemental Figure 7). It should be noted that the 3'UTR sequences of *Irf3* and *Irf5*

295 obtained using the methods described above (section 2.1) appeared to be incomplete, not  
296 reaching the expected poly-A tail.

297 Alignment of each transcript against its genomic sequence using the gadMor2 genome  
298 assembly (Torresen *et al.*, 2017) allowed for the prediction of location and sizes of introns  
299 (Figure 1; Supplemental Figures 2-7). Phylogenetic analysis of Atlantic cod IRF proteins and  
300 those from selected other teleosts (Supplemental Table 1) indicated that all cod IRFs  
301 characterized in both the current and previous studies (Inkpen *et al.*, 2015) were putatively  
302 orthologous to IRFs from other fish species. Multiple sequence alignment showed that the DBDs  
303 (first 110-120 AA) of all teleost IRF sequences included were quite similar, including well-  
304 conserved Trp residues found in all IRFs (Supplemental Figure 8). In a phylogenetic tree based  
305 on the multiple sequence alignment (Figure 2), all teleost IRF sequences grouped into the four  
306 sub-groups described above (IRF1-G, IRF3-G, IRF4-G, IRF5-G), as expected.

307

### 308 *3.2 RT-PCR expression analysis in Atlantic cod tissues*

309 *GmIrf2*, *GmIrf3*, *GmIrf5*, *GmIrf6*, and *GmIrf9* transcript expression was observed in 18  
310 cod tissues using RT-PCR and agarose gel electrophoresis (Supplemental Figure 9). While most  
311 of these transcripts showed some expression in all tissues, *GmIrf6* appeared to be absent or show  
312 very low expression in blood, eye, brain, heart, spleen, and muscle (Supplemental Figure 9F).  
313 Notably, differences in expression were observed between the two *GmIrf2* splice variants  
314 identified in this study (Supplemental Figure 9B, C). The longer splice variant (*GmIrf2-v1*)  
315 showed more uniform expression in all tissues, while the shorter splice variant (*GmIrf2-v2*)  
316 appeared to have very low expression in the gonad, muscle and digestive system (i.e. stomach,

317 pyloric caecum, midgut, hindgut). *GmIrf5* also appeared to have lower expression in muscle and  
318 tissues of the digestive system relative to other tissues (Supplemental Figure 9E).

319

### 320 3.3 QPCR expression analysis in embryonic and early larval development

321 Atlantic cod IRF transcripts showed distinct expression profiles during early development  
322 (0-17 dpf). It should be noted that, while *Tubb2* and *EIF3* were chosen as acceptable normalizers  
323 by geNorm analysis (section 2.4), expression of both transcripts in 0-1 dpf samples was lower  
324 than at other time-points (normalizer Ct values are presented in Supplemental Table 2). The  
325 overall expression profiles of *GmIrf1* and *GmIrf7* show some similarities (Figure 3A, I), both  
326 peaking during segmentation (8-10 dpf for *GmIrf1* and 7-8 dpf for *GmIrf7*), though *GmIrf1*  
327 expression was low overall and was in fact undetectable by QPCR in some stages. Splice  
328 variants of *GmIrf2* again showed differences in constitutive expression (e.g. *GmIrf2-v1*  
329 expression was highest at 2 dpf, while *GmIrf2-v2* expression was highest at 10 dpf; Figure 3B,  
330 C), though it should be noted that variation in expression among biological replicates within  
331 time-points was high. The two *GmIrf10* splice variants identified previously (Inkpen *et al.*, 2015)  
332 also had different expression profiles: *GmIrf10-v1* appeared to increase from 4-8 dpf, peaking  
333 and then decreasing gradually, while *GmIrf10-v2* expression appeared to increase drastically at 2  
334 dpf, remaining high through most of segmentation and dropping again at 11 dpf (Figure 3L, M).  
335 Interestingly, the expression profile of *GmIrf4b* (Figure 3F) was very similar to that of *GmIrf10-*  
336 *v2*, with relatively higher expression from 2-10 dpf, lower expression during hatching and an  
337 apparent increase in expression post-hatch. *GmIrf5*, *GmIrf6*, and *GmIrf8* were all relatively  
338 highly expressed in early cleavage stages (0-1 dpf), then dropped drastically and remained  
339 relatively low for the rest of the study (Figure 3G, H, J). As noted above, the relatively low



340 expression of both normalizer transcripts at 0-1 dpf (Supplemental Table 2) may have impacted  
341 the presented RQ values at these time-points. Expression of *GmIrf3* showed a general decrease  
342 with time (Figure 3D), with lowest relative expression occurring during hatching (13-14 dpf),  
343 while *GmIrf4a* and *GmIrf9* transcripts both showed increases in expression over time, with  
344 highest relative expression in hatched larvae (15-17 dpf; Figure 3E, K).

345

#### 346 *3.4 QPCR expression analysis in immune-stimulated head kidney macrophages*

347 The expression of all *Irf* transcripts in Atlantic cod macrophages stimulated with  
348 poly(I:C) or LPS was analyzed using QPCR at four time-points (12, 24, 48 and 72 HPS). To  
349 confirm that both treatments induced an immune response, additional antiviral (viperin, ISG15)  
350 and antibacterial (IL8, IL-1 $\beta$ ) transcripts were also analyzed (Supplemental Figure 10), showing  
351 increased expression in response to poly(I:C) and LPS, respectively. Increased transcript  
352 expression of IFN $\gamma$  in response to both treatments also indicates stimulation of interferon  
353 signalling pathways.

354 A summary of the observed *Irf* transcript responses in comparison with previous cod  
355 studies is presented in Table 3. Most transcripts were significantly up-regulated in response to  
356 poly(I:C) stimulation in at least one time-point, and none were down-regulated by poly(I:C)  
357 stimulation (Figure 4). Neither *GmIrf6* nor *GmIrf8* showed any significant difference in  
358 expression between the control (PBS-treated) and poly(I:C) treated cells, or within treatments at  
359 different time points. *GmIrf1* showed a higher expression in poly(I:C)-treated cells than PBS  
360 controls at 12 and 48 HPS; upregulation was also apparent at 24 h, but this difference was not  
361 statistically significant (Figure 4A). Once again, differences in expression were observed  
362 between *GmIrf2-v1* and *GmIrf2-v2* splice variants (Figure 4B, C). *GmIrf2-v1* transcript

363 expression was up-regulated in response to poly(I:C) treatment at all four time-points compared  
364 with time-matched PBS controls, while *GmIrf2-v2* up-regulation was only observed at 24 HPS.  
365 Expression of *GmIrf2-v2* was also observed to change in the control (PBS) cells, decreasing from  
366 12 to 24 HPS. *GmIrf3* and *GmIrf7* transcript expression was up-regulated by poly(I:C) at all  
367 time-points; the greatest increases (over 7-fold) in *GmIrf3* were observed at 48 and 72 HPS  
368 (Figure 4D), while *GmIrf7* expression appeared to be most responsive to poly(I:C) at 12 and 24  
369 HPS (8.4 and 10.3-fold, respectively; Figure 4I). Both *GmIrf4* paralogues showed similar up-  
370 regulation in poly(I:C) stimulated cells compared to PBS controls at all time-points included in  
371 the study (Figure 4E, F), while expression of *GmIrf4b* also differed significantly between 12 and  
372 24 HPS PBS (control) samples. *GmIrf5* transcript expression was only slightly up-regulated (< 2-  
373 fold) in response to poly(I:C) stimulation at 12 and 24 HPS, while no significant response was  
374 observed at the later time-points (Figure 4G). Expression at 24 HPS was, however, significantly  
375 lower than at any other time-point for *GmIrf5*, in both PBS and poly(I:C) treated cells. Very  
376 similar expression profiles were observed for *GmIrf9* and *GmIrf10-v1* transcripts in response to  
377 poly(I:C) stimulation (Figure 4K, L); as they were up-regulated at all time-points, with the  
378 highest expression observed in poly(I:C) treated cells at 48 HPS (5.7-fold and 23.5-fold  
379 upregulation in *GmIrf9* and *GmIrf10-v1*, respectively) and 72 HPS (4.7-fold and 20.9-fold  
380 upregulation in *GmIrf9* and *GmIrf10-v1*, respectively). These transcripts also showed lower  
381 expression in control samples at 24 HPS than at all other time-points. *GmIrf10-v1* showed the  
382 greatest increase in expression overall, with over 20-fold upregulation in poly(I:C) treated cells  
383 compared to PBS controls at 48 and 72 HPS. The shorter *GmIrf10* splice variant (*GmIrf10-v2*),  
384 however, was only 1.7- and 2-fold up-regulated at 48 and 72 HPS respectively, and was non-  
385 responsive to poly(I:C) at 12 and 24 HPS (Figure 4M).

386           Only 6 of 13 Atlantic cod *Irf* transcripts were LPS-responsive, each showing up-  
387 regulation of 2-fold or less compared to time-matched PBS controls (Figure 5). *GmIrf1*, *GmIrf2-*  
388 *v1*, *GmIrf3*, *GmIrf4a* and *GmIrf10-v2* were all up-regulated at only 72 HPS, while *GmIrf5* was  
389 up-regulated only at 12 HPS. While *GmIrf2-v2* expression appeared to be up-regulated in  
390 response to LPS at several time-points (Figure 5C), expression of this transcript varied widely  
391 among individuals within treatment groups, and these changes were not statistically significant.  
392 Both *GmIrf6* and *GmIrf8* were non-responsive to LPS (Figure 5H, J) as well as poly(I:C) as  
393 noted above, and neither of these transcripts showed significant changes in expression among  
394 time points. *GmIrf7* expression, while responsive to poly(I:C), showed no significant response to  
395 LPS stimulation (Figure 5I). Notably, several transcripts (i.e. *GmIrf2-v1*, *GmIrf3*, *GmIrf4b*,  
396 *GmIrf5*, *GmIrf9*, *GmIrf10-v1*) showed similar patterns of expression over time, wherein  
397 expression decreased from 12 to 24 HPS and then increased from 24 HPS to 48 HPS in PBS  
398 and/or LPS-treated cells (Figure 5B, D, F, G, K, L).

399

## 400 **4. Discussion**

### 401 *4.1 IRF transcript characterization and phylogenetic analysis*

402           Following the characterization of *GmIrf1*, *GmIrf4a*, *GmIrf4b*, *GmIrf7*, *GmIrf8* and  
403 *GmIrf10* (Feng *et al.*, 2009; Inkpen *et al.*, 2015) by our research group, the objective of the  
404 current study was to complete the characterization of all Atlantic cod IRF family member  
405 transcripts. Six additional transcripts (*GmIrf2-v1*, *GmIrf2-v2*, *GmIrf3*, *GmIrf5*, *GmIrf6* and  
406 *GmIrf9*) were identified and characterized at the cDNA and putative amino acid levels, and all  
407 cod IRFs were compared with those of other teleost fish species by molecular phylogenetic

408 analysis. Long and short splice variants of Atlantic cod *Irf2* (*GmIrf2-v1* and *GmIrf2-v2*,  
409 respectively) were identified, and while analysis of the gadMor2 genome assembly also indicated  
410 an additional putative *Irf2* paralogue located on a scaffold region, this coding region was 99%  
411 identical to the RACE-identified *GmIrf2* located in linkage group (LG) 10 of the genome,  
412 suggesting a possible error in assembly rather than an additional gene copy. No other additional  
413 paralogues of the previously characterized cod IRFs were identified using the updated genome  
414 assembly.

415         Interestingly, similar to *GmIrf10-v2* (Inkpen *et al.*, 2015), *GmIrf2-v2* appears to be  
416 produced by intron retention, leading to a truncated transcript. IRF family member splice  
417 variants, particularly in human research, have been shown to perform distinct roles mainly  
418 associated with disease. For example, multiple variants, including splice variants, of human *Irf5*  
419 have been associated with risk of developing systemic lupus erythematosus (Graham *et al.*,  
420 2007), and human *Irf1* splice variants were observed to negatively regulate wild type *Irf1* in  
421 cervical cancer tissue (Lee *et al.*, 2006). However, to our knowledge, differential splicing of  
422 either *Irf2* or *Irf10* has not been observed in any other fish species; therefore, it is difficult to  
423 predict the function of these shorter transcripts in cod. Furthermore, we have not yet investigated  
424 the expression of these variants at the protein level. Thus, further expression and functional  
425 studies are required to determine if *GmIrf2* and *GmIrf10* variants have different regulatory roles.

426         Multiple sequence alignment and phylogenetic analysis indicated that all Atlantic cod  
427 IRFs fall into the expected four sub-groups (IRF1-G, IRF3-G, IRF4-G, IRF5-G; as in Nehyba *et*  
428 *al.*, 2002), and all contain the amino terminal DBD and associated conserved tryptophan residues  
429 common to all IRFs. As noted previously (Inkpen *et al.*, 2015), teleost IRF7 orthologues appear  
430 to lack one Trp compared to all other orthologues, while teleost IRF1-G sequences (i.e. IRF1 and

431 IRF2) have an additional Trp residue in the DBD. As these conserved amino acids play an  
432 important role in binding the ISRE (Escalante *et al.*, 1998), it is possible that variations may  
433 affect target gene specificity of IRF family members. When all Atlantic cod IRFs were  
434 compared, significant sequence variation was observed outside of the DBD, as expected  
435 (Supplemental Figure 8B). Variation in the IAD, at the carboxyl region of all IRF proteins except  
436 IRF1-G, is indicative of the wide range of functions of IRFs outside of IFN regulation, as this  
437 region is important for protein-protein interactions (Meraro *et al.*, 1999). Thus, *GmIrf10-v2* for  
438 example, which lacks the IAD, is expected to function quite differently from *GmIrf10-v1*, as  
439 suggested by the expression profiles discussed below.

#### 440 4.2 RT-PCR expression analysis in Atlantic cod tissues

441 All transcripts characterized in the current study appeared to be ubiquitously expressed in  
442 juvenile cod tissues except *GmIrf6*, which appeared to have little or no expression in some  
443 important immune related tissues such as the spleen and blood (Supplemental Figure 9F). This is  
444 consistent with our previous understanding of IRF6 function, as it has long been thought to be  
445 the only IRF family member without a known role in the innate immune response (Savitsky *et*  
446 *al.*, 2010). IRF6 has been shown to be necessary for epithelial development in other species such  
447 as zebrafish and the frog *Xenopus laevis* (Ben *et al.*, 2005; Sabel *et al.*, 2009); expression of  
448 *GmIrf6* was observed in Atlantic cod tissues such as skin and gill (as well as fin, kidney, gonad,  
449 and gut), suggesting a role for IRF6 in those tissues in juvenile fish.

450 As with the previously identified *Irf10* splice variants (Inkpen *et al.*, 2015), different  
451 expression patterns were observed among the *GmIrf2* variants based on RT-PCR analysis, with  
452 *GmIrf2-v1* appearing to be more ubiquitously expressed and more highly expressed overall

453 (Supplemental Figure 9B, C). IRF2 often acts as a transcriptional repressor, in opposition to  
454 IRF1 and IRF9, and has been shown to be pro-oncogenic (Savitsky *et al.*, 2010; Yanai *et al.*,  
455 2012). Its potentially conserved role in the IFN pathway in cod, as with most other IRFs, is  
456 supported by relatively high constitutive transcript expression in the spleen, head kidney, and  
457 blood. The very low expression of *GmIrf2-v2* in the gonad and several areas of the gut suggests it  
458 has a less significant role in those tissues than *GmIrf2-v1*; however, little is known of IRF2  
459 function in these areas. Several cod IRF transcripts show ubiquitous expression in the tissues  
460 studied (i.e. *GmIrf2-v1*, *GmIrf3*, *GmIrf9* in the current study and *GmIrf7*, *GmIrf8* and *GmIrf10-*  
461 *v1* in Inkpen *et al.* 2015), suggesting more ubiquitous function of these transcription factors in  
462 many cell types.

#### 463 *4.3 QPCR expression analysis in embryonic and larval development*

464 All Atlantic cod IRF transcripts, characterized in our previous (Inkpen *et al.*, 2015) and  
465 current studies, were included in the QPCR study of early developmental expression herein.  
466 Notably, most transcripts included in the current study showed high variation in expression  
467 among biological replicate pools during early development. Our research group has previously  
468 described high variation in *GmIrf7* transcript expression among egg batches from different  
469 females in Atlantic cod (Rise *et al.*, 2014), and thus while the replicates in the current study  
470 represented pooled embryos from multiple parents (communally spawned), for future studies a  
471 larger number of replicate batches from multiple parents may be useful. However, clear changes  
472 in transcript expression over time were observed despite the limited sample size, suggesting  
473 stage-specific functions of several transcripts.

474 *GmIrf1* and *GmIrf7* were included in a previous study of embryonic transcript expression,  
475 using different egg batches from those used in the current study (Rise *et al.*, 2012). The  
476 expression profiles observed for both transcripts, particularly *GmIrf7*, agreed well with the  
477 previous study, in both cases showing a peak in expression during segmentation (Figure 3A, I).  
478 Although there has been little study of IRF function during embryonic development in other  
479 species, a recent study of blunt snout bream (*Megalobrama amblycephala*) showed *Irf7*  
480 expression peaking during hatching (Zhan *et al.*, 2017), and *Irf7* has been observed as a potential  
481 indicator of egg quality in both Atlantic halibut (*Hippoglossus hippoglossus* L) and sea bass  
482 (*Dicentrarchus labrax*), showing significantly higher expression in “high quality” egg or embryo  
483 groups than “low quality” groups (Mommens *et al.*, 2014; Zarski *et al.*, 2017). Contrary to those  
484 studies, previous analysis of Atlantic cod *Irf7* expression in fertilized and unfertilized eggs  
485 showed no correlation with egg quality (Rise *et al.*, 2014). Thus, while our data suggest  
486 important roles for IRF7, IRF1, and potentially their target genes in the IFN pathway during  
487 early development, these roles may vary among teleost species. This study also supports  
488 previous research showing *GmIrf7* to be a maternal transcript (Rise *et al.*, 2012; 2014), as  
489 expression was relatively high at 0-1 dpf. The lower expression observed at 2-5 dpf may indicate  
490 degradation of a maternal transcript during the maternal-to-zygotic transition (MZT; for review  
491 see Lee *et al.*, 2014).

492 Several other Atlantic cod IRF transcripts show high relative expression at the earliest  
493 time-points studied (0-1 dpf) compared to later stage embryos, indicating potential maternal  
494 transcripts. As noted by Hall *et al.* (2004), Atlantic cod embryos at approximately 36 hours post  
495 fertilization (hpf) have entered the mid-blastula transition (MBT), during which the embryo  
496 begins to transcribe its own mRNA. Therefore, we may assume that transcripts with high

497 expression at 0-1 dpf are highly expressed in the unfertilized egg. Particularly, *GmIrf5*, *GmIrf6*,  
498 and *GmIrf8* all showed dramatically higher expression at 0-1 dpf than at any time-point  
499 thereafter (Figure 3G, H, J). Zebrafish IRF8 has been shown to function in embryonic  
500 macrophage development (Shiau *et al.*, 2015), though to our knowledge it has not been described  
501 as a maternal transcript. Functional studies of *GmIrf8* (e.g. using morpholino knockdown in  
502 Atlantic cod embryos or cultured macrophages) will be of interest in the future, to determine the  
503 potential role of this transcript in early development and particularly in embryonic myelopoiesis.  
504 *Irf6* has previously been described as a maternal transcript in zebrafish and *X. laevis*, and is  
505 known to have important roles in embryonic development (Sabel *et al.*, 2009). The authors of  
506 that study showed that maternally-derived *Irf6* seems to be vital for epithelial differentiation,  
507 whereas blocking expression of embryonic IRF6 had little effect. Similar knockdown studies of  
508 *GmIrf6* should be performed in the future to determine if this role is conserved in Atlantic cod.

509

#### 510 *4.4 QPCR expression analysis in immune-stimulated head kidney macrophages*

511 In previous reports, we have analyzed the transcript expression responses of *GmIrf1*,  
512 *GmIrf4a*, *GmIrf4b*, *GmIrf7*, *GmIrf8*, *GmIrf10-v1* (identified as *Irf10* in most studies), and  
513 *GmIrf10-v2* stimulated with poly(I:C) and/or killed *Aeromonas salmonicida* (ASAL) in immune  
514 tissues (Rise *et al.*, 2008; Feng *et al.*, 2009; Hori *et al.*, 2012; Inkpen *et al.*, 2015; Eslamloo *et*  
515 *al.*, 2016). In the current study, all 13 identified Atlantic cod *Irf* transcripts were analyzed in  
516 response to stimulation with either poly(I:C) or LPS (i.e. inducing an antiviral or antibacterial  
517 cellular response, respectively) in isolated head kidney macrophages. It should be noted that the  
518 minimal or absent response to LPS stimulation observed for many transcripts may indicate that



519 they have less importance to the antibacterial response than the antiviral response, but could also  
520 be partially due to variation in the response to LPS compared to a live or killed bacterium. We  
521 have observed, for example, different responses to LPS than to killed ASAL (see Table 3 for  
522 comparison). As our group has discussed previously (Smith *et al.*, 2018 and references therein),  
523 most fish lack TLR4 as an LPS receptor, but do respond to immune stimulation with LPS  
524 (Seppola *et al.*, 2015 and references therein), perhaps through an as-yet uncharacterized  
525 mechanism. Expression of Il-8 and Il-1 $\beta$ , commonly used as biomarkers of the antibacterial  
526 response, were both increased in response to LPS (Supplemental Figure 10), indicating the LPS  
527 stimulation in this study was effective. Protein contaminants in the LPS used may also have  
528 contributed to the transcript expression response observed, as noted by Smith *et al.* (2018).

529         IRF1 has been well studied in vertebrates as a transcriptional activator of IFN $\alpha/\beta$ , whose  
530 expression can be induced by IFNs and other cytokines, or by viral infection (reviewed in  
531 Taniguchi *et al.* 2001), and we have previously indicated its role in the antiviral response of  
532 Atlantic cod (see Table 3). In isolated macrophages, *GmIrf1* was significantly upregulated (> 5-  
533 fold) in response to poly(I:C) at 12 and 48 HPS, though the response appeared to decrease or end  
534 by 72 HPS (Figure 4A). The potential role of IRF1 in antibacterial responses is not well  
535 understood, though we have observed up-regulation in response to ASAL stimulation in Atlantic  
536 cod previously (Feng *et al.*, 2009). In response to LPS, *GmIrf1* was up-regulated slightly (<2-  
537 fold compared to PBS controls) and only at 72 HPS (Figure 5A).

538         Expression of both Atlantic cod *Irf2* splice variants increased in response to poly(I:C),  
539 though while *GmIrf2-v1* was significantly up-regulated (2.5 to 4.5-fold) at all four time-points,  
540 *GmIrf2-v2* up-regulation (2.3-fold) was only observed at 24 HPS (Figure 4B, C). *Irf2* has  
541 previously been shown to be poly(I:C)-responsive in Atlantic salmon cell lines (Bergen *et al.*,

2010), in head kidney of orange-spotted grouper (Shi *et al.*, 2010), in rainbow trout gonad cells (Collet *et al.*, 2003), and in several grass carp tissues (Gu *et al.*, 2015). However, it is notable that *Irf2* expression in salmon was not changed in response to infectious salmon anemia virus, unlike other *Irf* transcripts (Bergen *et al.*, 2010). In LPS-stimulated Atlantic cod macrophages, *GmIrf2-v1* was significantly up-regulated (1.4-fold) only at 72 HPS, and while *GmIrf2-v2* expression appeared to be up-regulated at several time-points, these differences were not statistically significant (Figure 5B, C). *Irf2* was previously uncharacterized in Atlantic cod, and to our knowledge no studies have investigated *Irf2* splice variants *Irf2* in other fish species, while only one study has described IRF2 variants in mammalian cells (Koenig Merediz *et al.*, 2000). Therefore, while the differences observed in response to poly(I:C) and LPS stimulation suggest different roles for the two variants in the cellular immune response, further functional studies will be required to elucidate these roles.

*GmIrf3* was not fully characterized prior to this study, though we previously characterized the closely related *GmIrf7* and observed it to be slightly up-regulated in response to ASAL in the spleen (Inkpen *et al.*, 2015), and more highly up-regulated in response to poly(I:C) in the spleen (Rise *et al.*, 2008; Hori *et al.*, 2012), and in head kidney macrophages (Eslamloo *et al.*, 2016). In the current study, both *GmIrf3* and *GmIrf7* were significantly up-regulated in response to poly(I:C), though while *GmIrf3* was most responsive at later time-points (over 7-fold up-regulated at 48 and 72 HPS), *GmIrf7* showed the greatest increase (10.3 fold) at 24 HPS (Figure 4D, I). *Irf3* has previously been shown to be poly(I:C) responsive in Atlantic cod larvae cells (Krasnov *et al.*, 2013), and in other teleost species [e.g. in rainbow trout (Holland *et al.*, 2008), Atlantic salmon (Bergen *et al.*, 2010), turbot, *Scophthalmus maximus* (Hu *et al.*, 2011), and tilapia, *Oreochromis niloticus* (Gu *et al.*, 2016)], though only the rainbow trout study

565 investigated macrophage expression. The roles of both IRF3 and IRF7 as important regulators of  
566 the antiviral response are well understood in mammalian species; both IRF3 and IRF7 are  
567 expressed in the cytosol, activated downstream of the TLR3-dependent pathway and then  
568 translocate to the nucleus following activation in response to viral infection (Honda and  
569 Taniguchi, 2006). Fish IRF3 and IRF7 appear to function similarly to mammalian orthologues in  
570 the antiviral response. In all vertebrates, the mechanisms by which these genes may regulate  
571 antibacterial responses are less understood. In the current study *GmIrf7* expression in cod  
572 macrophages showed no response to LPS stimulation, while *GmIrf3* was slightly up-regulated  
573 (1.4-fold) at only 72 HPS (Figure 5D, I).

574         The IRF5 sub-group (*Irf5*, *Irf6*) had, to our knowledge, never been studied in cod prior to  
575 this study; and the function of IRF6 in most fish species is not well understood. The role of IRF5  
576 in the teleost antiviral immune response, however, has been indicated in several species,  
577 including Japanese flounder (Hu *et al.*, 2012), tongue sole (*Cynoglossus semilaevis*; Zhang *et al.*,  
578 2015), common carp (*Cyprinus carpio* L.; Zhu *et al.*, 2016), and zebrafish (Ai *et al.*, 2018). In  
579 the current study, *GmIrf5* expression increased in response to poly(I:C) stimulation, though, only  
580 less than 2-fold increases over time-matched PBS controls were observed (Figure 4G). *GmIrf5*  
581 was also up-regulated in response to LPS, though only at the earliest time-point in the study (12  
582 HPS; Figure 5G). Interestingly, this is similar to the expression profile observed for *Il-8*, a  
583 commonly used biomarker for antibacterial responses (Supplemental Figure 10). Others have  
584 shown *Irf5* expression to be responsive to bacterial infection in common carp (Zhu *et al.*, 2016),  
585 and tongue sole (Zhang *et al.*, 2015), which along with the current study suggest that teleost  
586 IRF5 may be important to the innate antibacterial response in addition to the antiviral immune  
587 response.

588 *GmIrf6* expression showed no significant change in response to poly(I:C) and LPS  
589 stimulation in Atlantic cod head kidney macrophages (Figure 4H; Figure 5H). While these  
590 results are consistent with the current understanding that the primary function of IRF6 is in  
591 epithelial development (Ben *et al.*, 2005; Sabel *et al.*, 2009), a few studies have shown evidence  
592 of *Irf6* up-regulation in response to poly(I:C) in zebrafish (Li *et al.*, 2016) and bacterial infection  
593 in tongue sole (Zhang *et al.*, 2015). Thus, further study of *GmIrf6* in additional tissues and in  
594 response to additional viral and bacterial antigens and live pathogens is needed to rule out a role  
595 in the cod innate immune response.

596 Within the IRF4 sub-group, we previously characterized two *Irf4* paralogues (*GmIrf4a*;  
597 *GmIrf4b*), and *GmIrf8*, and saw that both *GmIrf4b* and *GmIrf8* expression increased slightly in  
598 the spleen in response to poly(I:C) and ASAL, though *GmIrf8* expression initially decreased  
599 slightly with ASAL stimulation (Inkpen *et al.*, 2015). *GmIrf4a* in that study had no response to  
600 ASAL, and a 2-fold decrease in response to poly(I:C). In the current study, *GmIrf8* had no  
601 significant response to either poly(I:C) or LPS (Figure 4J; Figure 5J). When taken with the  
602 previous study, this suggests that if this transcript does play a role in the cellular immune  
603 response, it may be a more subtle role, and its functions as a constitutively expressed gene may  
604 be more important. However, *Irf8* expression was responsive to poly(I:C) in several other teleost  
605 species [rainbow trout (Holland *et al.*, 2010), rock bream (Bathidge *et al.*, 2012), turbot (Chen *et*  
606 *al.*, 2012, and Japanese flounder (Hu *et al.*, 2013)], suggesting our observations of *GmIrf8*  
607 expression may be unique to Atlantic cod, though notably those studies did not investigate  
608 expression in isolated macrophages. Both *GmIrf4a* and *GmIrf4b* increased in expression with  
609 poly(I:C) stimulation at all time-points in the current study (Figure 4E, F), differing from our  
610 observations in the spleen as noted above. Interestingly, the response of these paralogues to LPS

611 was somewhat opposite to the response observed to ASAL in the spleen (Inkpen *et al.*, 2015), as  
612 *GmIrf4b* had no response, and *GmIrf4a* showed only a 1.6-fold increase at 72 HPS (Figure 5E,  
613 F). Further studies may help determine if these differences are indicative of tissue- or cell-  
614 specific functions. IRF4 and IRF8 have been shown in mammalian species to have important  
615 roles in myelopoiesis and the differentiation of macrophages (Tamura *et al.*, 2015 and references  
616 therein; Nam and Lim, 2016 and references therein), and similar function of IRF8 has also been  
617 observed in zebrafish (Li *et al.*, 2011). Although we have not yet determined if this function is  
618 conserved in Atlantic cod, it would support the hypothesis that *GmIrf8* is less involved in the  
619 macrophage antiviral and antibacterial responses and more important to other processes.

620 Atlantic cod *Irf10* was also shown to be immune responsive in our previous experiments  
621 (Hori *et al.*, 2012; Inkpen *et al.*, 2015), with increased expression in response to both poly(I:C)  
622 and ASAL in the spleen. Xu *et al.* (2016) also showed IRF10 in three different teleost species  
623 [rainbow trout (two paralogues), grass carp, and swamp eel, *Monopterus albus*] to be responsive  
624 to viral infection and/or poly(I:C). In the current study, both *GmIrf10* splice variants were up-  
625 regulated in response to poly(I:C) stimulation, though while *GmIrf10-v1* expression was  
626 increased at all time-points and showed the highest fold changes in the study (up to 23.5-fold;  
627 Figure 4L), *GmIrf10-v2* was only up-regulated at 48 HPS and 72 HPS, and only showed a ~2-  
628 fold increase (Figure 4M). With LPS stimulation, *GmIrf10-v1* showed no significant response,  
629 while *GmIrf10-v2* increased in expression slightly (< 2-fold) at 72 HPS (Figure 5L, M).  
630 Similarly, *GmIrf10-v2* was slightly more responsive than *GmIrf10-v1* to ASAL in the spleen  
631 (Inkpen *et al.*, 2015), though both variants showed some increase in expression in that study.  
632 Collectively, these studies may suggest that the two variants have different roles in the innate

633 immune response, with *GmIrf10-v1* potentially acting more in the antiviral response and  
634 *GmIrf10-v2* in the antibacterial response.

635 IRF9 has not previously been studied in Atlantic cod, though its role in the IFN pathway  
636 has been well studied in mammalian species – it forms the ISGF3 transcription factor complex  
637 along with STAT1 and STAT2, which activates several IFN pathway genes (Taniguchi *et al.*,  
638 2001; Yanai *et al.*, 2012, and references therein). To our knowledge IRF9 has only been studied  
639 in a few teleost species, but within those studies it has been shown to be up-regulated in response  
640 to IFN stimulation in zebrafish (Shi *et al.*, 2013), in response to poly(I:C) and/or viral infection  
641 in Japanese flounder (Hu *et al.*, 2014), tongue sole (Zhang *et al.*, 2015), miiuy croaker, *Miichthys*  
642 *miuy* (Yang *et al.*, 2017), and mandarin fish (Laghari *et al.*, 2018), and in response to bacterial  
643 infection in tongue sole (Zhang *et al.*, 2015) and blunt snout bream (Zhan *et al.*, 2017). In the  
644 current study, *GmIrf9* showed increased expression in response to poly(I:C) at all time-points,  
645 but had no significant response to LPS (Figure 4K; Figure 5K). As noted above, further study of  
646 the fish response to LPS, along with the use of live bacterial infection and other bacterial  
647 antigens, would be valuable to better understand any potential roles of *Irf9* and other IRFs in the  
648 teleost antibacterial response.

649

## 650 **Conclusions:**

651 In total, we have now characterized 13 IRF transcripts in Atlantic cod, completing the  
652 characterization of this gene family at the transcript and hypothetical amino acid levels. All cod  
653 IRF transcripts appear to be constitutively expressed in multiple tissues, and all were observed to  
654 be expressed during embryonic and early larval development. These findings suggest potential  
655 roles for IRF family members in Atlantic cod outside of their function in the innate antiviral

656 response, which may be similar to those observed in other vertebrate species. Immune  
657 stimulation of isolated Atlantic cod macrophages with poly(I:C) produced increases in  
658 expression of all but two IRF transcripts, including several transcripts which had been previously  
659 uncharacterized and therefore not shown to be immune responsive prior to this study. Several  
660 IRF transcripts were also responsive to stimulation with bacterial LPS, suggesting roles in the  
661 innate antibacterial response.

662

663 **Acknowledgments:**

664 This study was supported by a Natural Sciences and Engineering Research Council of Canada  
665 (NSERC) Discovery Grant to MLR. The genomic analyses were supported by The Research  
666 Council of Norway (Grant number 222378/F20) and performed on the Abel Supercomputing  
667 Cluster (Norwegian metacenter for High Performance Computing (NOTUR) and the University  
668 of Oslo) operated by the Research Computing Services group at USIT, the University of Oslo IT-  
669 Department (<http://www.hpc.uio.no/>). The authors would like to thank the Dr. Joe Brown  
670 Aquatic Research Building (JBARB) staff (Ocean Sciences Centre, Memorial University, NL),  
671 for their assistance with fish husbandry, Atefeh Ghorbani for help with tissue sampling, and  
672 Heather Fifield for assistance with sequencing.

673

675 **References:**

- 676 Ai K, Luo K, Xia L, Gao W, Hu W, Qi Z, Xu Q. 2018. Functional characterization of interferon  
677 regulatory factor 5 and its role in the innate antiviral immune response. *Fish Shellfish Immunol*  
678 72:31-36.
- 679 Amores A, Force A, Yan Y, Joly L, Amemiya C, Fritz A, Ho RK, Langeland J, Prince V, Wang  
680 Y, Westerfield M, Ekker M, Postlethwait JH. 1998. Zebrafish *hox* clusters and vertebrate  
681 genome evolution. *Science* 282:1711-14.
- 682 Bathige SDNK, Whang I, Umasuthan N, Lim BS, Park MA, Kim E, Park HC, Lee J. 2012.  
683 Interferon regulatory factors 4 and 8 in rock bream, *Oplegnathus fasciatus*: Structural and  
684 expressional evidence for their antimicrobial role in teleosts. *Fish Shellfish Immunol* 33:857-71.
- 685 Ben J, Wang Jabs E, and Chong SS. 2005. Genomic, cDNA and embryonic expression analysis  
686 of zebrafish IRF6, the gene mutated in the human oral clefting disorders Van der Woude and  
687 popliteal pterygium syndromes. *Gene Expression Patterns* 5:629-38.
- 688 Bergan V, Kileng Ø, Sun B, Robertsen B. 2010. Regulation and function of interferon regulatory  
689 factors of Atlantic salmon. *Mol Immunol* 47:2005-14.
- 690 Caipang CM, Brinchmann MF, Kiron V. 2009. Profiling gene expression in the spleen of  
691 Atlantic cod, *Gadus morhua* upon vaccination with *Vibrio anguillarum* antigen. *Comp Biochem*  
692 *Physiol B Biochem Mol Biol* 153(3):261-67.
- 693 Chen X, Hu G, Dong X, Liu Q, Zhang S. 2012. Molecular cloning and expression analysis of  
694 interferon regulatory factor 8 (IRF8) in turbot, *Scophthalmus maximus*. *Vet Immunol*  
695 *Immunopathol* 149(1-2):143-50.
- 696 Collet B, Hovens GCJ, Mazzoni D, Hirono I, Aoki T, Secombes CJ. 2003. Cloning and  
697 expression analysis of rainbow trout *Oncorhynchus mykiss* interferon regulatory factor 1 and 2  
698 (IRF-1 and IRF-2). *Dev Comp Immunol* 27(2):111-26.
- 699 Darnell JE, Kerr IM, Stark GR. 1994. Jak-STAT pathways and transcriptional activation in  
700 response to IFNs and other extracellular signaling proteins. *Science* 264:1415-21.
- 701 Escalante CR, Yie J, Thanos D, Aggarwal AK. 1998. Structure of IRF-1 with bound DNA  
702 reveals determinants of interferon regulation. *Nature* 319:103-06.
- 703 Eslamloo K, Xue X, Booman M, Smith NC, Rise ML. 2016. Transcriptome profiling of the  
704 antiviral immune response in Atlantic cod macrophages. *Dev Comp Immunol* 63:187-205.
- 705 Eslamloo K, Inkpen SM, Rise ML, Andreassen R. 2018. Discovery of microRNAs associated  
706 with the antiviral immune response of Atlantic cod macrophages. *Mol Immunol* 93:152-61.



707 Feng CY, Johnson SC, Hori TS, Rise M, Hall JR, Gamperl AK, Hubert S, Kimball J, Bowman S,  
708 Rise ML. 2009. Identification and analysis of differentially expressed genes in immune tissues of  
709 Atlantic cod stimulated with formalin-killed, atypical *Aeromonas salmonicida*. *Physiol*  
710 *Genomics* 37(3):149-63.

711 Graham RR, Kyogoku C, Sigurdsson S, Vlasova IA, Davies LRL, Baechler EC, Plenge RM,  
712 Koeuth T, Ortmann WA, Hom G, et al., 2007. Three functional variants of IFN regulatory factor  
713 5 (IRF5) define risk and protective haplotypes for human lupus. *Proc Natl Acad Sci USA*  
714 104:6758-63.

715 Gu M, Lin G, Lai Q, Zhong B, Liu Y, Mi Y, Chen H, Wang B, Fan L, Hu C. 2015.  
716 *Ctenopharyngodon idella* IRF2 plays an antagonistic role to IRF1 in transcriptional regulation of  
717 IFN and ISG genes. *Dev Comp Immunol* 49(1):103-12.

718 Gu YF, Wei Q, Tang SJ, Chen XW, Zhao JL. 2016. Molecular characterization and functional  
719 analysis of IRF3 in tilapia (*Oreochromis niloticus*). *Dev Comp Immunol* 55:130-7.

720 Hall TE, Smith P, Johnson IA. 2004. Stages of embryonic development in the Atlantic cod  
721 *Gadus morhua*. *J Morphol* 259(3):255-0.

722 Holland JW, Bird S, Williamson B, Woudstra C, Mustafa A, Wang T, Zou J, Blaney SC, Collet  
723 B, Secombes CJ. 2008. Molecular characterization of IRF3 and IRF7 in rainbow trout,  
724 *Oncorhynchus mykiss*: Functional analysis and transcriptional modulation. *Mol Immunol* 46:269-  
725 85.

726 Holland JW, Karim A, Wang T, Alnabulsi A, Scott J, Collet B, Mughal MS, Secombes CJ, Bird  
727 S. 2010. Molecular cloning and characterization of interferon regulatory factors 4 and 8 (IRF-4  
728 and IRF-8) in rainbow trout, *Oncorhynchus mykiss*. *Fish Shellfish Immunol* 29(1):157-66.

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

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

734 Hu GB, Xia J, Lou HM, Chen XL, Li J, Liu QM. 2011. An IRF-3 homolog that is up-regulated  
735 by DNA virus and poly I:C in turbot, *Scophthalmus maximus*. *Fish Shellfish Immunol*  
736 31(6):1224-31.

737 Hu GB, Lou HM, Dong XZ, Liu QM, Zhang SC. 2012. Characteristics of the interferon  
738 regulatory factor 5 (IRF5) and its expression in response to LCDV and poly I:C challenges in  
739 Japanese flounder, *Paralichthys olivaceus*. *Dev Comp Immunol* 38(2):377-82.

- 740 Hu G, Chen X, Gong Q, Liu Q, Zhang S, Dong X. 2013. Structural and expression studies of  
741 interferon regulatory factor 8 in Japanese flounder, *Paralichthys olivaceus*. Fish Shellfish  
742 Immunol 35(3):1016-24.
- 743 Hu GB, Zhao MY, Lin JY, Liu QM, Zhang SC. 2014. Molecular cloning and characterization of  
744 interferon regulatory factor 9 (IRF9) in Japanese flounder, *Paralichthys olivaceus*. Fish Shellfish  
745 Immunol 39(2):138-44.
- 746 Huang B, Qi Z, Xu Z, Nie P. 2010. Global characterization of interferon regulatory factor (IRF)  
747 genes in vertebrates: Glimpse of the diversification in evolution. BMC Immunol 11:22.
- 748 Huang WS, Zhu MH, Chen S, Wang ZW, Liang Y, Huang B, Nei P. 2017. Molecular cloning  
749 and expression analysis of a fish specific interferon regulatory factor, IRF11, in orange spotted  
750 grouper, *Epinephelus coioides*. Fish Shellfish Immunol 60:368-379.
- 751 Inkpen SM, Hori TS, Gamperl AK, Nash GW, Rise ML. 2015. Characterization and expression  
752 analyses of five interferon regulatory factor transcripts (Irf4a, Irf4b, Irf7, Irf8, Irf10) in Atlantic  
753 cod (*Gadus morhua*). Fish Shellfish Immunol 44(1):365-81.
- 754 Koenig Merediz SA, Schmidt M, Hoppe GJ, Alfken J, Meraro D, Levi BZ, Neubauer A, Wittig  
755 B. 2000. Cloning of an interferon regulatory factor 2 isoform with different regulatory ability.  
756 Nucleic Acids Res 28(21):4219-24.
- 757 Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, Howard E, Ferreira  
758 de Lima RL, Daack-Hirsch S, *et al.* 2002. Mutations in IRF6 cause Van der Woude and popliteal  
759 pterygium syndromes. Nature Genetics 32:285-89.
- 760 Krasnov A, Kileng O, Skugor S, Jorgensen SM, Afanasyev S, Timmerhaus G, Sommer AI,  
761 Jensen I. 2013. Genomic analysis of the host response to nervous necrosis virus in Atlantic cod  
762 (*Gadus morhua*) brain. Mol Immunol 54:443-52.
- 763 Kumar S, Stecher G, Tamura K (2016). MEGA7: Molecular Evolutionary Genetics Analysis  
764 version 7.0 for bigger datasets. Mol Biol Evol 33:1870-1874.
- 765 Laghari ZA, Li L, Chen SN, Huo HJ, Huang B, Zhou Y, Nie P. 2018. Composition and  
766 transcription of all interferon regulatory factors (IRFs), IRF1–11 in a perciform fish, the  
767 mandarin fish, *Siniperca chuatsi*. Dev Comp Immunol 81:127-40.
- 768 Lee E, Jo M, Park J, Zhang W, Lee J. 2006. Alternative splicing variants of IRF-1 lacking exons  
769 7, 8, and 9 in cervical cancer. Biochem Biophys Res Comm 347:882-88.
- 770 Lee MT, Bonneau AR, Giraldez AJ. 2014. Zygotic genome activation during the maternal-to-  
771 zygotic transition. Annu Rev Cell Dev Biol 30:581-613.
- 772 Li L, Jin H, Xu J, Shi Y, Wen Z. 2011. Irf8 regulates macrophage versus neutrophil fate during  
773 zebrafish primitive myelopoiesis. Blood 117:1359-69.

774 Li S, Lu LF, Wang ZX, Chen DD, Zhang YA. 2016. Fish IRF6 is a positive regulator of IFN  
775 expression and involved in both of the MyD88 and TBK1 pathways. *Fish Shellfish Immunol*  
776 57:262-68.

777 Lien S, Koop BF, Sandve SR, Miller JR, Kent MP, Nome T, et al. 2016. The Atlantic salmon  
778 genome provides insights into rediploidization. *Nature* 533(7602):200-5.

779 Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time  
780 quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25:402-8.

781 Malmstrøm M, Matschiner M, Tørresen OK, Star B, Snipen LG, Hansen TF, Baalsrud HT,  
782 Nederbragt AJ, Hanel R, Salzburger W, Stenseth NC, Jakobsen KS, Jentoft S. 2016. Evolution of  
783 the immune system influences speciation rates in teleost fishes. *Nat Genet* 48:1204-10.

784 Matta B, Song S, Li D, Barnes BJ. 2017. Interferon regulatory factor signaling in autoimmune  
785 disease. *Cytokine* 98:15-26.

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

789 Mommens M, Fernandes JMO, Tollefsen KE, Johnston IA, Babiak I. 2014. Profiling of the  
790 embryonic Atlantic halibut (*Hippoglossus hippoglossus* L.) transcriptome reveals maternal  
791 transcripts as potential markers of embryo quality. *BMC Genomics* 15:829.

792 Nam S, Lim JS. 2016. Essential role of interferon regulatory factor 4 (IRF4) in immune cell  
793 development. *Arch Pharm Res* 39(11):1548-55.

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

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

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

801 Pfaffl M.W. 2001. A new mathematical model for relative quantification in real-time RT-PCR.  
802 *Nucleic Acids Res* 29(9):e45.

803 Rise ML, Hall J, Rise M, Hori T, Gamperl AK, Kimball J, Hubert S, Bowman S, Johnson SC.  
804 2008. Functional genomic analysis of the response of Atlantic cod (*Gadus morhua*) spleen to the  
805 viral mimic polyriboinosinic polyribocytidylic acid (pIC). *Dev Comp Immunol* 32:916-31.

- 806 Rise ML, Hall JR, Alcock BP, Hori TS. 2012. Dynamic expression profiles of virus-responsive  
807 and putative antimicrobial peptide-encoding transcripts during Atlantic cod (*Gadus morhua*)  
808 embryonic and early larval development. *Gene* 509:232-46.
- 809 Rise ML, Nash GW, Hall JR, Booman M, Hori TS, Trippel EA, Gamperl AK. 2014. Variation in  
810 embryonic mortality and maternal transcript expression among Atlantic cod (*Gadus morhua*)  
811 broodstock: a functional genomics study. *Marine Genomics Pt A* 3-20.
- 812 Sabel JL, d'Alençon C, O'Brien EK, Otterloo EV, Lutz K, Cuykendall TN, Schutte BC, Houston  
813 DW, Cornell RA. 2009. Maternal interferon regulatory factor 6 is required for the differentiation  
814 of primary superficial epithelia in *Danio* and *Xenopus* embryos. *Dev Biol* 325:249-62.
- 815 Savitsky D, Yanai H, Taniguchi T. 2010. Regulation of immunity and oncogenesis by the IRF  
816 transcription factor family. *Cancer Immunol Immunother* 59:489-510.
- 817 Seppola M, Mikkelsen H, Johansen A, Steiro K, Myrnes B, Nilsen IW. 2015. Ultrapure LPS  
818 induces inflammatory and antibacterial responses attenuated in vitro by exogenous sera in  
819 Atlantic cod and Atlantic salmon. *Fish Shellfish Immunol* 44(1):66-78.
- 820 Shi Y, Zhao Z, Yin JK, Zhu XP, Chen KC, Pan DB, Gui JF. 2010. Interferon regulatory factor-2  
821 in orange-spotted grouper (*Epinephelus coioides*): Gene, inductive expression pattern and  
822 subcellular localization. *Comp Biochem Physiol B Biochem Mol Biol* 155(2):110-17.
- 823 Shi J, Zhang YB, Zhang JS, Gui JF. 2013. Expression regulation of zebrafish interferon  
824 regulatory factor 9 by promoter analysis. *Dev Comp Immunol* 41(4):534-43.
- 825 Shiao CE, Kaufman Z, Meireles AM, Talbot WS. 2015. Differential requirement for irf8 in  
826 formation of embryonic and adult macrophages in zebrafish. *PLoS One* 10(1):e0117513.
- 827 Smith NC, Christian SL, Taylor RG, Santander J, Rise ML. 2018. Immune modulatory  
828 properties of 6-gingerol and resveratrol in Atlantic salmon macrophages. *Mol Immunol* 95:10-  
829 19. Corrigendum: Smith NC *et al.* 2018. *Mol Immunol* 104:139.
- 830 Solbakken MH, Rise ML, Jakobsen KS, Jentoft S. 2016. Successive losses of central immune  
831 genes characterize the Gadiformes' alternate immunity. *Genome Biol Evol* 8(11):3508-15.
- 832 Solbakken MH, Voje KL, Jakobsen KS, Jentoft S. 2017. Linking species habitat and past  
833 palaeoclimatic events to evolution of the teleost innate immune system. *Proc Biol Sci*  
834 284(1853):20162810.
- 835 Star B, Nederbragt AJ, Jentoft S, Grimholt U, Malmstrøm M, Gregers TF, *et al.* 2011. The  
836 genome sequence of Atlantic cod reveals a unique immune system. *Nature* 477(7363):207-10.
- 837 Stein C, Caccamo M, Laird G, Leptin M. 2007. Conservation and divergence of gene families  
838 encoding components of innate immune response systems in zebrafish. *Genome Biol* 8:R251.

- 839 Sun BJ, Chang MX, Song Y, Yao WJ, Nie P. 2007. Gene structure and transcription of IRF-1  
840 and IRF-7 in the mandarin fish *Siniperca chuatsi*. *Vet Immunol Immunopathol* 116:26-36.
- 841 Tamura T, Kurotaki D, Koizumi S. 2015. Regulation of myelopoiesis by the transcription factor  
842 IRF8. *Int J Hematol* 101(4):342-51.
- 843 Tanaka N, Ishihara M, Lamphier MS, Nozawa H, Matsuyama T, Mak TW, Aizawa S, Tokino T,  
844 Oren M, Taniguchi T. 1996. Cooperation of the tumor suppressors IRF-1 and p53 in response to  
845 DNA damage. *Nature* 382(6594):816-18.
- 846 Taniguchi T, Ogasawara K, Takaoka A, Tanaka N. 2001. IRF family of transcription factors as  
847 regulators of host defense. *Annu Rev Immunol* 19:623-55.
- 848 Torresen OK, Star B, Jentoft S, Reinart WB, Grove H, Miller JR, Walenz BP, Knight J, Ekholm  
849 JM, Peluso P, Edvardsen RB, Tooming-Klunderud A, Skage M, Lien S, Jakobsen KS,  
850 Nederbragt AJ. 2017. An improved genome assembly uncovers prolific tandem repeats in  
851 Atlantic cod. *BMC Genomics* 18:95.
- 852 Xiang Z, Dong C, Qi L, Chen W, Huang L, Li Z, Xia Q, Liu D, Huang M, Weng S, *et al.*, 2010.  
853 Characteristics of the interferon regulatory factor pairs zfIRF5/7 and their stimulation expression  
854 by ISKNV infection in zebrafish (*Danio rerio*). *Dev Comp Immunol* 34:1263-73.
- 855 Xu Q, Jiang Y, Wangkahart E, Zou J, Chang M, Yang D, Secombes CJ, Nie P, Wang T. 2016.  
856 Sequence and Expression Analysis of Interferon Regulatory Factor 10 (IRF10) in Three Diverse  
857 Teleost Fish Reveals Its Role in Antiviral Defense. *PLoS One* 11(1):e0147181.
- 858 Yanai H, Negishi H, Taniguchi T. 2012. The IRF family of transcription factors: inception,  
859 impact and implications in oncogenesis. *OncoImmunology* 1(8):1376-86.
- 860 Yang Q, Cui J, Song W, Zhao X, Xu T. 2017. The evolution and functional characterization of  
861 miiuy croaker interferon regulatory factor 9 involved in immune response. *Fish Shellfish*  
862 *Immunol* 66:524-30.
- 863 Źarski D, Nguyen T, Le Cam A, Montfort J, Dutto G, Vidal MO, Fauvel C, Bobe J. 2017.  
864 Transcriptomic Profiling of Egg Quality in Sea Bass (*Dicentrarchus labrax*) Sheds Light on  
865 Genes Involved in Ubiquitination and Translation. *Mar Biotechnol* 19:102-15.
- 866 Zhan FB, Liu H, Lai RF, Jakovlic I, Wang WM. 2017. Expression and functional  
867 characterization of interferon regulatory factors (*irf2*, *irf7* and *irf9*) in the blunt snout bream  
868 (*Megalobrama amblycephala*). *Dev Comp Immunol* 67:239-248.
- 869 Zhang J, Li YX, Hu YH. 2015. Molecular characterization and expression analysis of eleven  
870 interferon regulatory factors in half-smooth tongue sole, *Cynoglossus semilaevis*. *Fish Shellfish*  
871 *Immunol* 44(1):272-82.

872 Zhu Y, Qi C, Shan S, Zhang F, Li H, An L, Yang G. 2016. Characterization of common carp  
873 (*Cyprinus carpio L.*) interferon regulatory factor 5 (IRF5) and its expression in response to viral  
874 and bacterial challenges. BMC Vet Res 12(1):127.

875

876

### Web References

877 <http://bioinfo.ut.ee/primer3-0.4.0/>; Web interface for Primer3.

878 <https://blast.ncbi.nlm.nih.gov/Blast.cgi>; National Center for Biotechnology Information BLAST  
879 web interface (BLASTx; BLASTn; BLAST 2 sequences).

880

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

882 <http://www.ensembl.org/>; Ensembl genome browser web interface.

884

885 **Figure Legends:**

886 **Figure 1: Schematic structure of Atlantic cod *Irf2-v1*, *Irf2-v2*, *Irf3*, *Irf5*, *Irf6*, and *Irf9*.**

887 Exons (E) are depicted as shaded boxes, where 1 cm represents 100 bp. 5' and 3' untranslated  
888 regions (UTR) and introns (I) are depicted as lines. Introns longer than 150 bp and UTRs are  
889 represented by bent lines (not to scale). A portion of I6 in *Irf2-v1* is expressed in E6 of *Irf2-v2*,  
890 and is shaded black. Intron lengths are inferred from the gadMor2 genome assembly.

891

892 **Figure 2: Phylogenetic analysis of teleost IRF sequences.** Putative Atlantic cod amino acid  
893 sequences were aligned with IRF proteins from other teleost fish species (see Supplemental  
894 Table 1 for GenBank accession numbers) using MEGA7 software (Kumar *et al.*, 2016). Based  
895 on the multiple sequence alignment, the evolutionary history was inferred using the neighbour-  
896 joining method. The bootstrap consensus tree was constructed from 2000 replicates, where  
897 numbers at the branch points represent percentage of replicates in which the associated taxa  
898 grouped together. Branch lengths are proportional to calculated evolutionary distances. Atlantic  
899 cod sequences determined by our group in the current and previous studies are indicated by an  
900 asterisk.

901

902 **Figure 3: Atlantic cod transcript expression of IRF family members throughout embryonic**  
903 **and early larval development, measured by QPCR.** Data is presented as mean +/- SEM, of 3  
904 pools normalized to the geometric mean of *Tubb2* and *EIF3* expression, with the lowest  
905 expressing sample set to RQ=1. Normalizer Ct values are presented in Supplemental Table 2.

906 Developmental stages at each sampling point (x-axis) are based on observations and photographs  
907 taken daily using light microscopy (Supplemental Figure 1), with reference to the work of Hall *et*  
908 *al.* (2004).

909 **Figure 4: Atlantic cod macrophage transcript expression response of IRF family members**  
910 **to poly(I:C), measured by QPCR.** Data is presented as mean +/- SEM, normalized to the  
911 geometric mean of *EF1α* and *rplp1* expression, with the lowest expressing sample set to RQ=1.  
912 Normalizer Ct values are presented in Supplemental Table 2. Note that *Irf1* expression is  
913 represented as log<sub>2</sub> of RQ, due to the wide range of RQ values observed for that transcript. An  
914 asterisk (\*) represents a significant difference between a poly(I:C) stimulated group and the  
915 time-matched PBS group (p < 0.05). [\*\* = p<0.01; \*\*\* = p<0.001; \*\*\*\* = p<0.0001]. Capital  
916 letters and lower case letters represent similarity among PBS-treated groups and poly(I:C)-  
917 treated groups across time points, respectively. Fold change is calculated as [mean poly(I:C)  
918 RQ/mean PBS RQ].

919

920 **Figure 5: Atlantic cod macrophage transcript expression response of IRF family members**  
921 **to LPS, measured by QPCR.** Data is presented as mean +/- SEM, normalized to the geometric  
922 mean of *EF1α* and *rplp1* expression, with the lowest expressing sample set to RQ=1. Normalizer  
923 Ct values are presented in Supplemental Table 2. Note that *Irf1* expression is represented as log<sub>2</sub>  
924 of RQ, due to the wide range of RQ values observed for that transcript. An asterisk (\*) represents  
925 a significant difference between a poly(I:C) stimulated group and the time-matched PBS group  
926 (p < 0.05). Capital letters and lower case letters represent similarity among PBS-treated groups

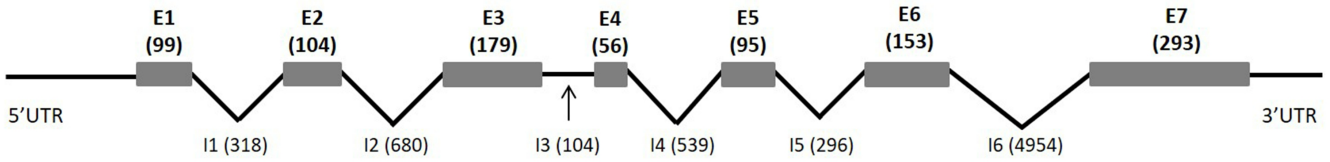


927 and LPS-treated groups across time points, respectively. Fold change is calculated as (mean LPS  
928 RQ)/(mean PBS RQ).

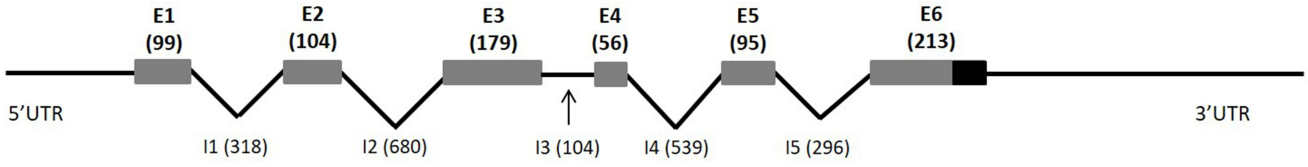
929

930

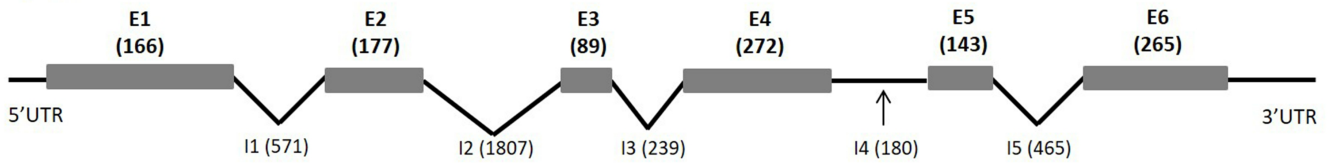
### A) IRF2-v1



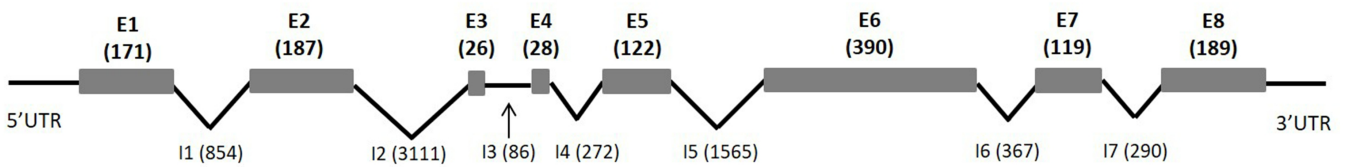
### B) IRF2-v2



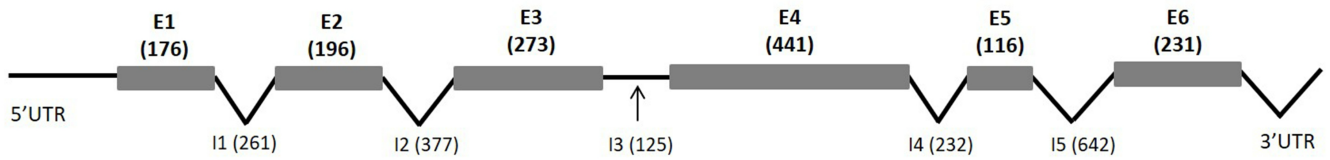
### C) IRF3



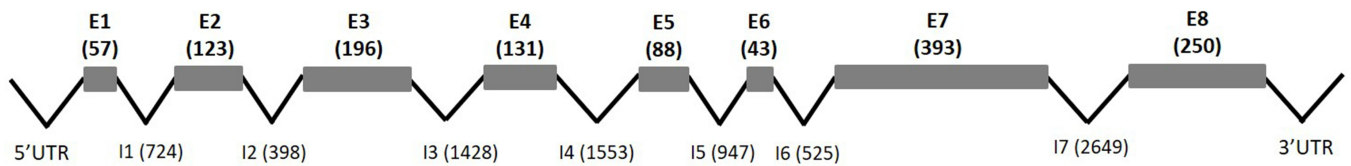
### D) IRF5



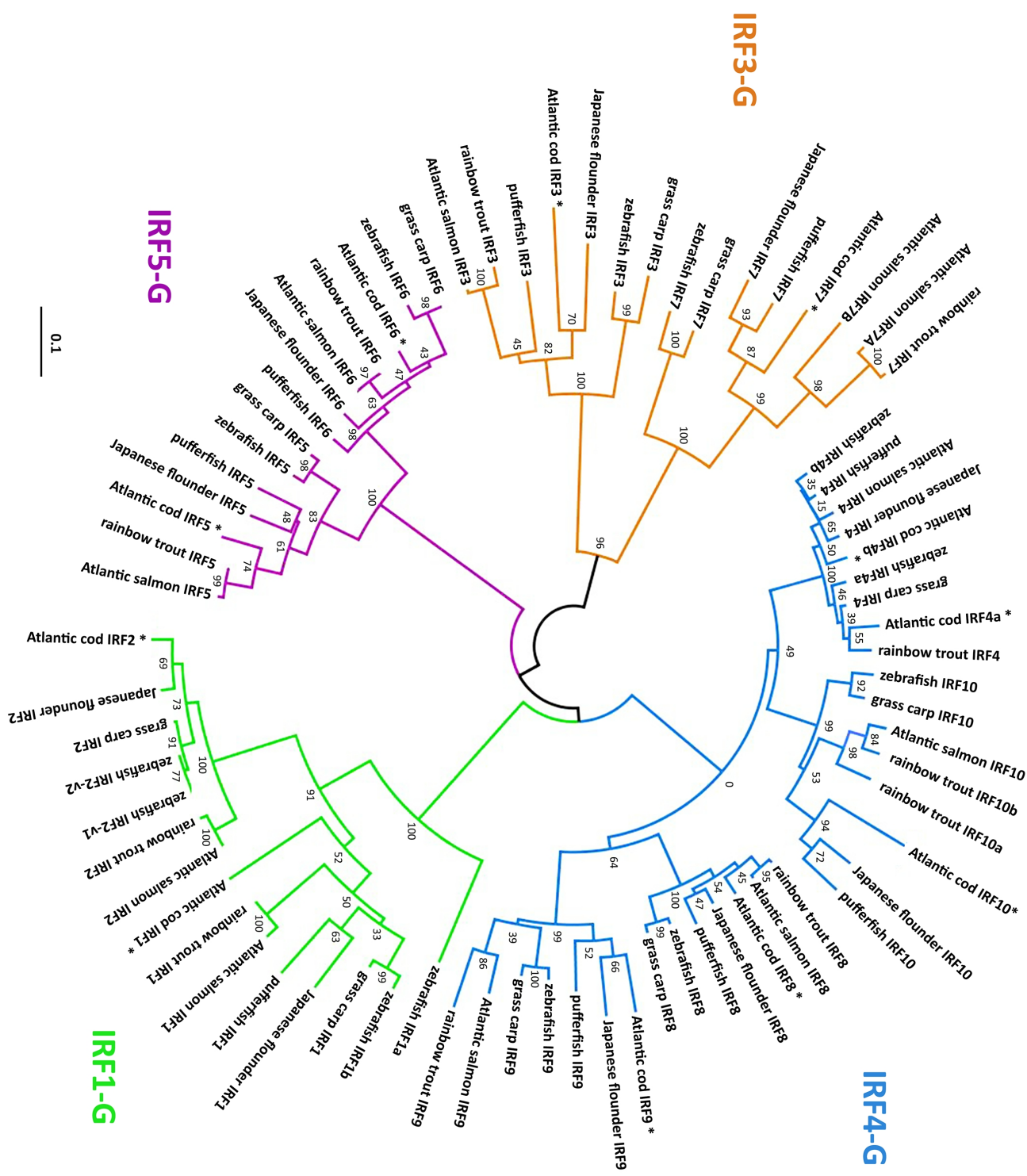
### E) IRF6

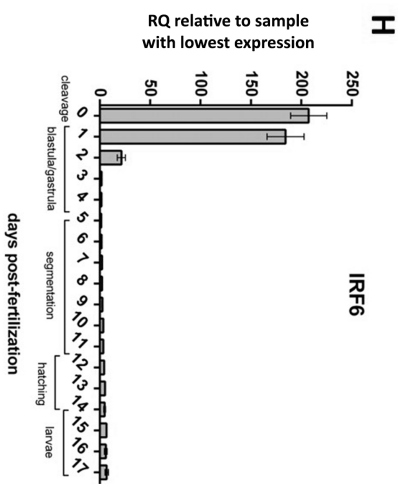
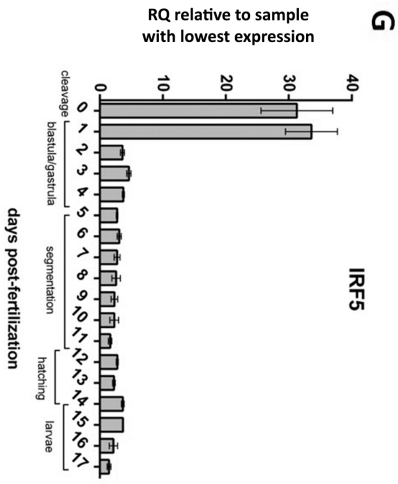
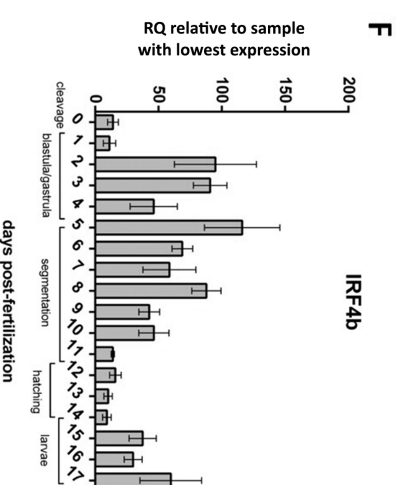
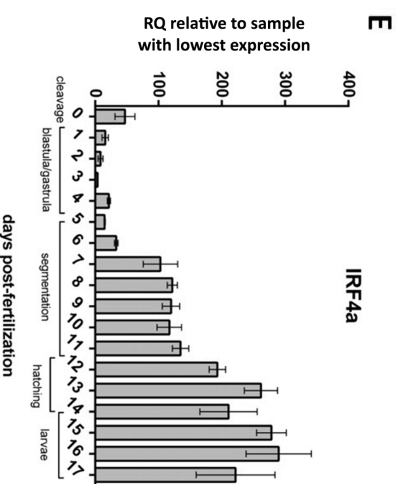
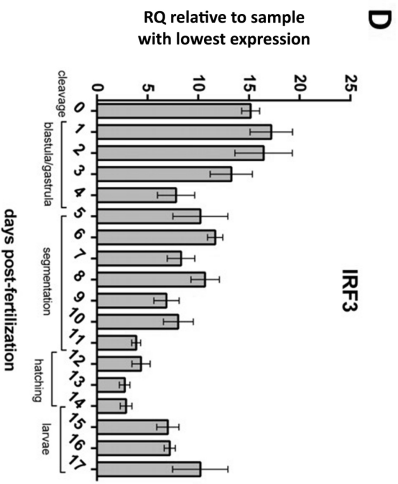
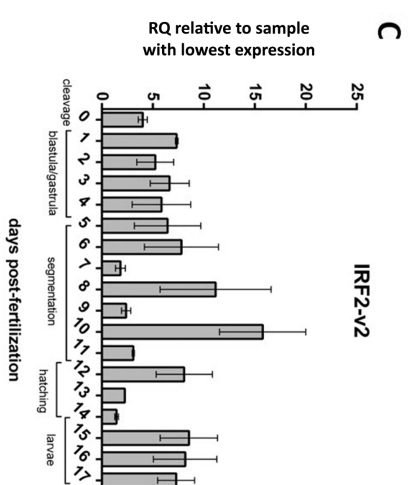
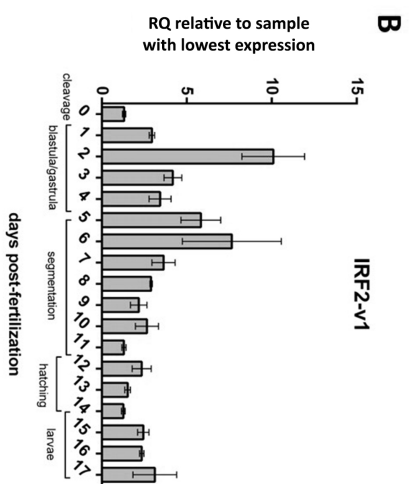
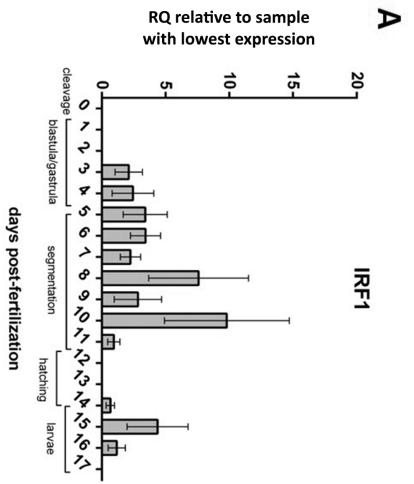


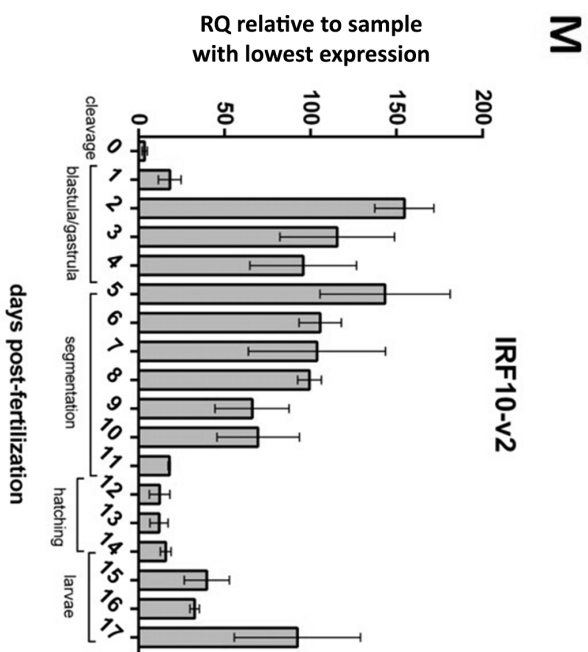
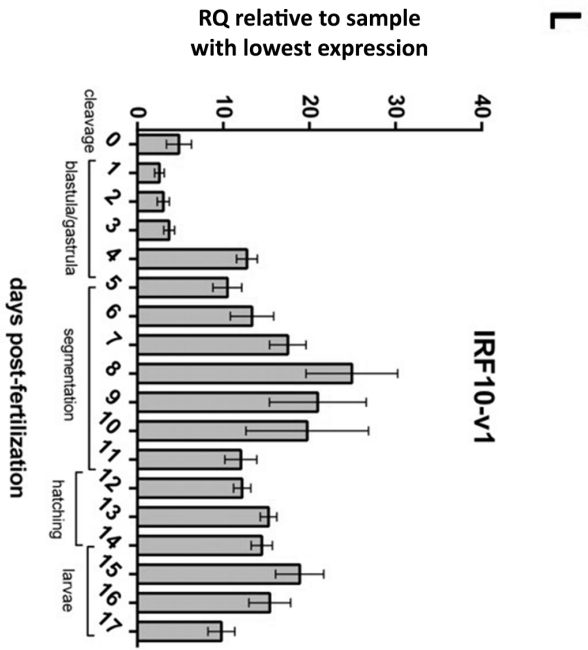
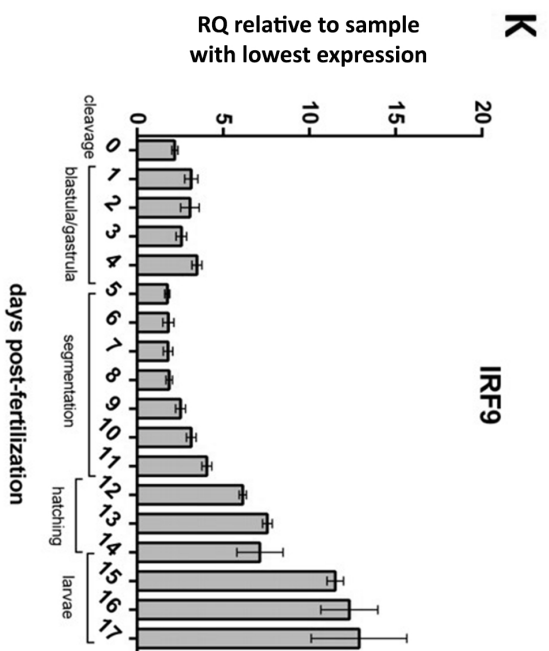
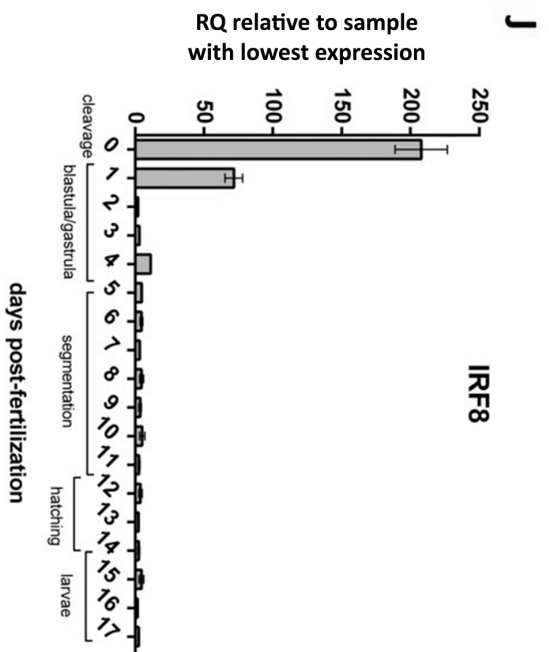
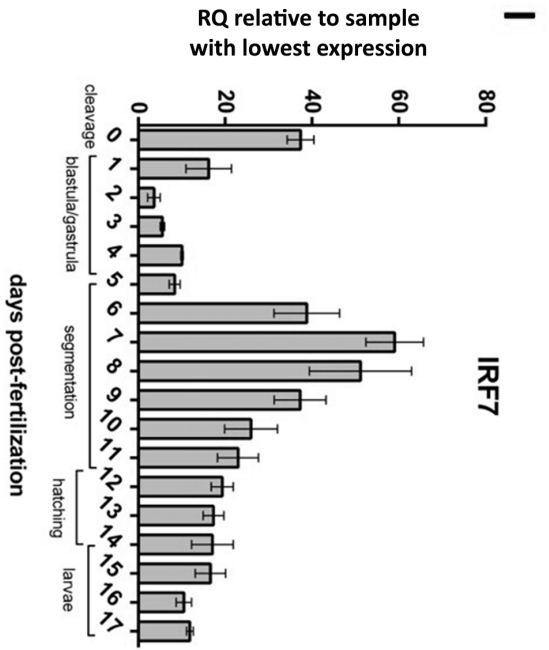
### F) IRF9

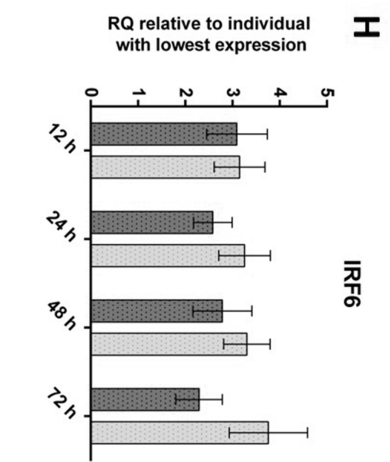
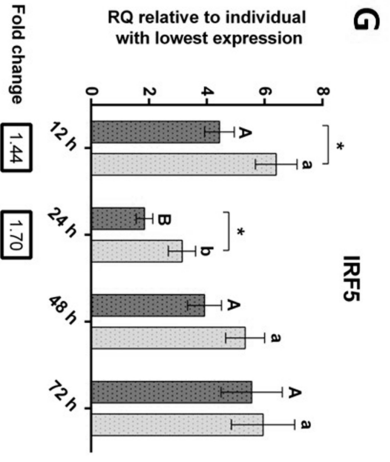
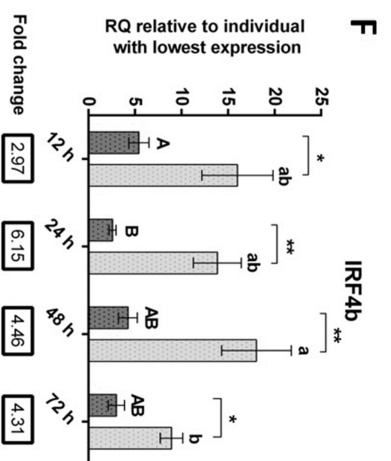
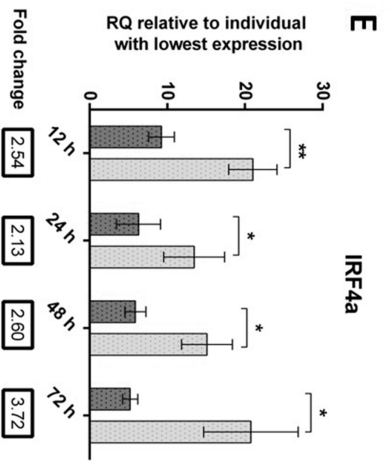
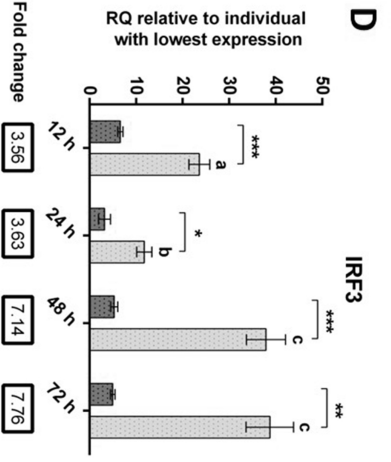
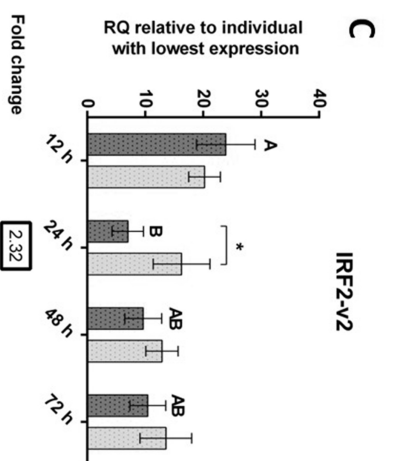
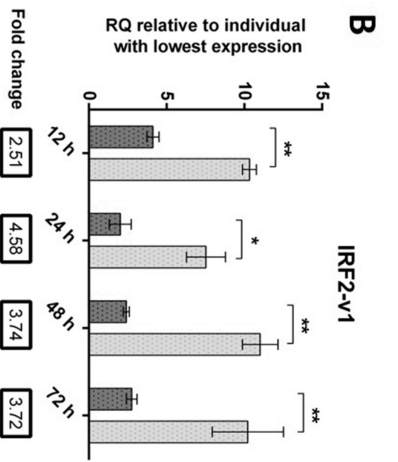
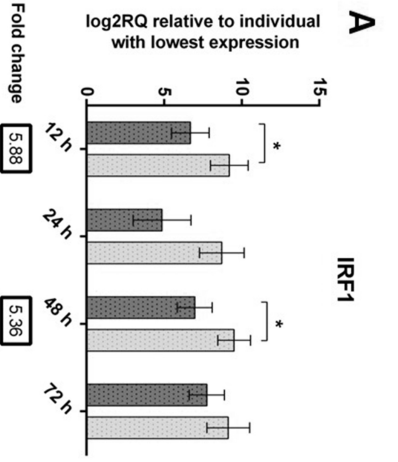


■ Exon — Intron; UTR — 200 nt

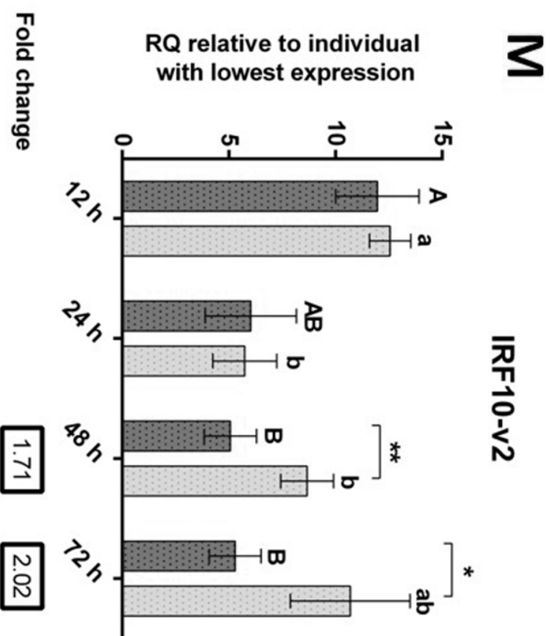
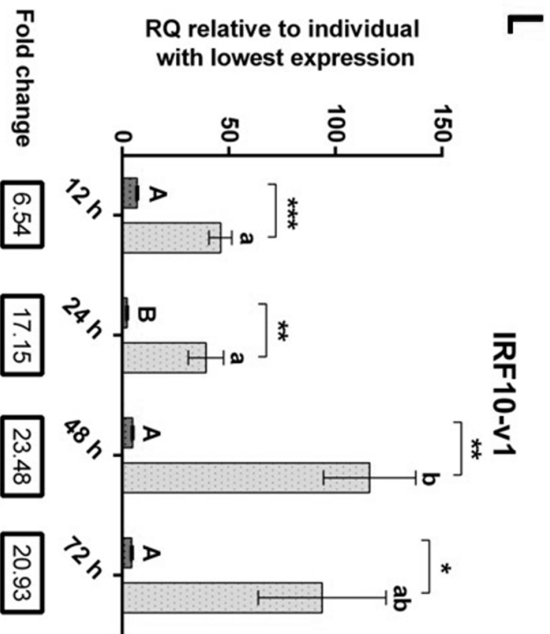
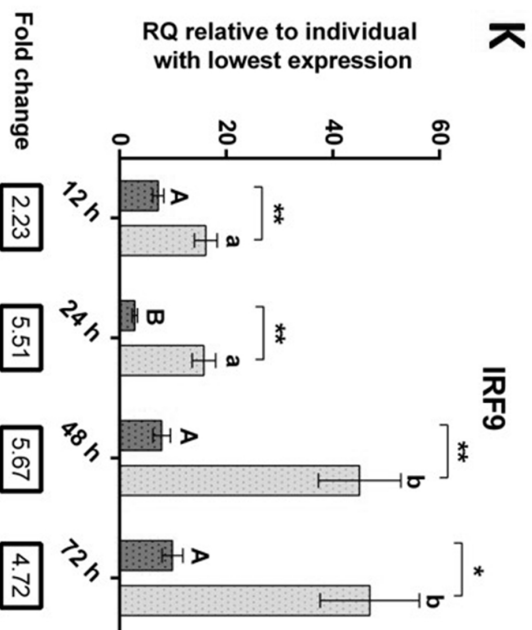
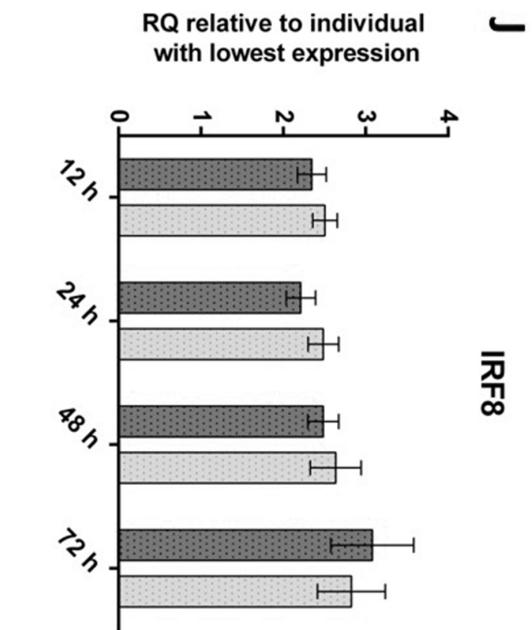
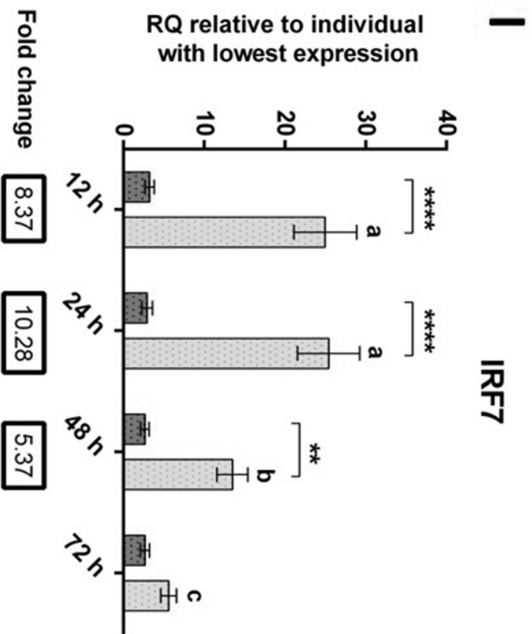




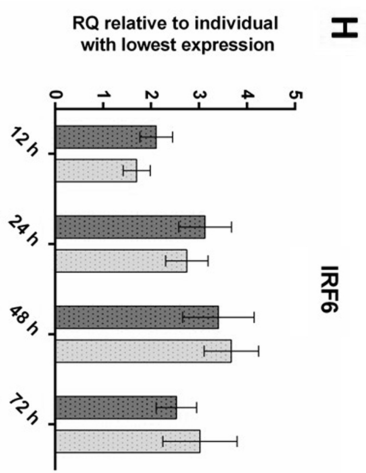
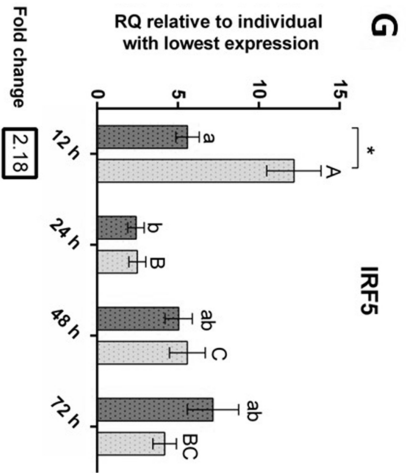
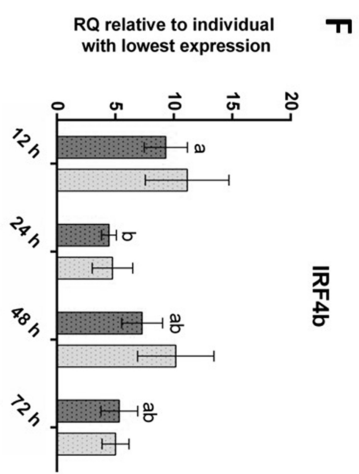
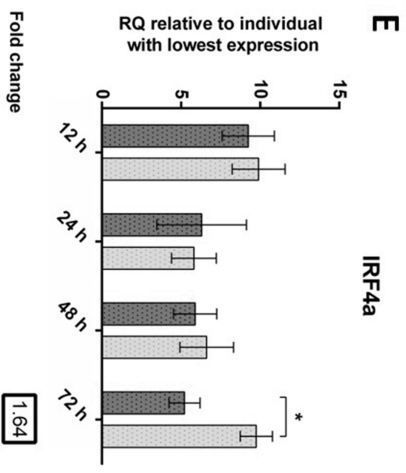
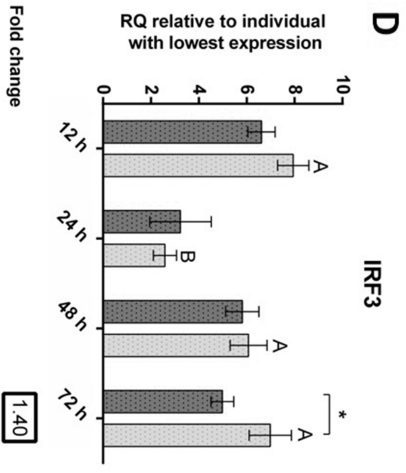
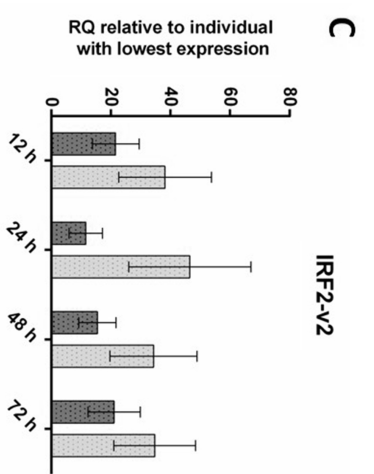
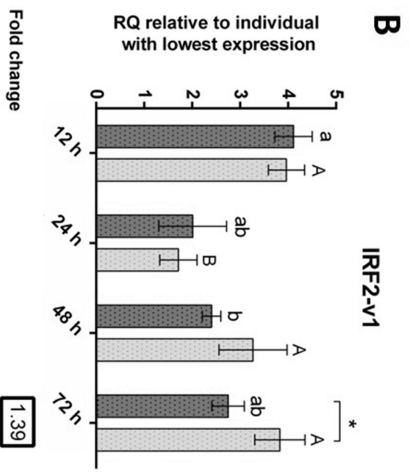
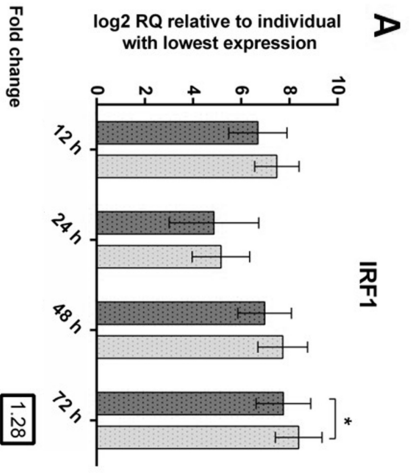




■ PBS  
 □ Poly(I:C)

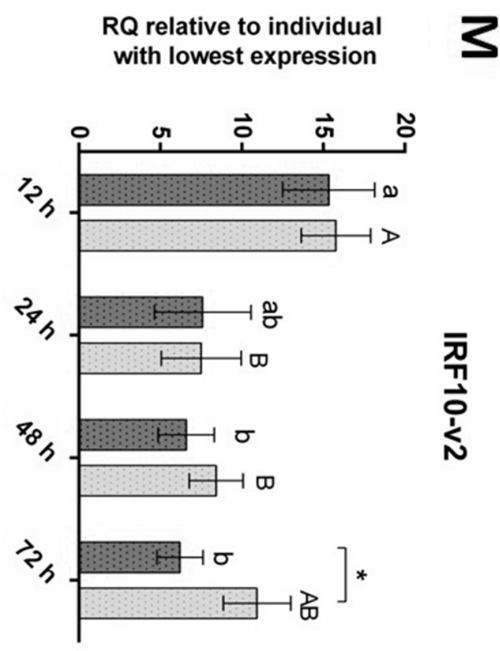
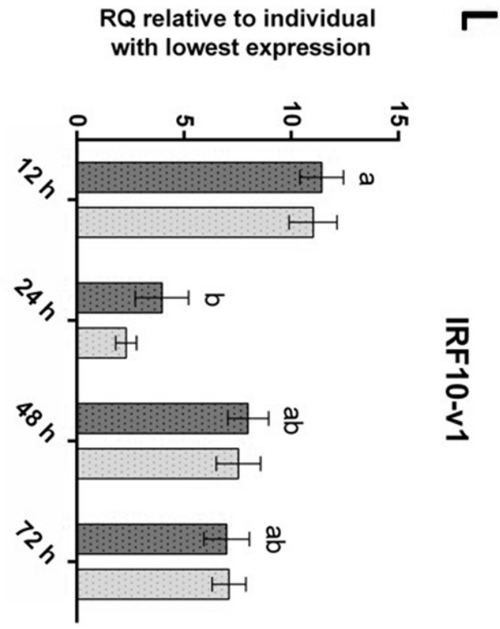
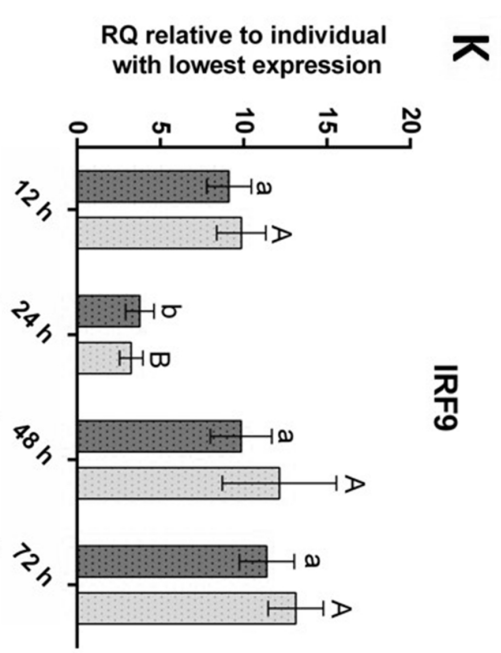
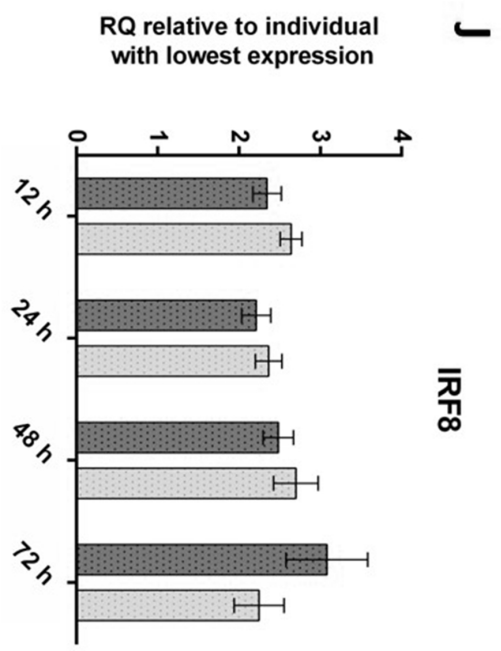
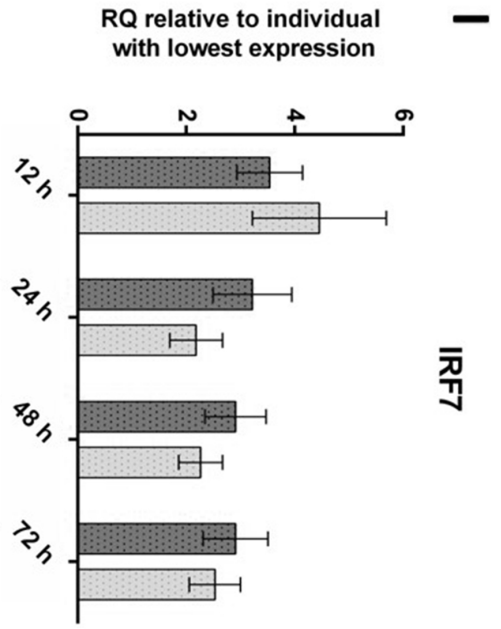


PBS  
 Poly(I:C)



■ PBS  
■ LPS





**PBS**
  
**LPS**

Table 1: Primers used in RACE for characterization of Atlantic cod *Irf2*, *Irf3*, *Irf5*, *Irf6*, and *Irf9*

Primer name	Sequence (5'-3')	Amplicon length (bp)
<i>Irf2</i> -race-5'1	GGAGGAGCTGACCATCTGCTCTT	868
<i>Irf2</i> -race-5'2	GTACGAAGACACGGGAGAGATCTG	901
<i>Irf2</i> -race-3'1	CCCAATAGGCGAGCACCTGATCAC	575 / 792
<i>Irf2</i> -race-3'2	GCTGCCATTTGTCTGCCAGACCAT	543 / 759
<i>Irf3</i> -race-5'1	GCCCCAAGGGTGCTGCGGAAGT	327
<i>Irf3</i> -race-5'2	CACCTTGATGGGGTTGGCGTTTTTC	373
<i>Irf3</i> -race-3'1	GTCAGTGTCTACAGAGGGGTGAAG	889
<i>Irf3</i> -race-3'2	CTGGTGGATAACGAAGCTGGTTTTTC	853
<i>Irf3</i> -race-mid-f	CCCTCCGTCTGGAAGAGGAACTT	302
<i>Irf3</i> -race-mid-r	CGGTTCTGCACCGAGTCGAAGGT	
<i>Irf5</i> -race-5'1	GCCTCCGTGTAGAACCTCTGCTT	874
<i>Irf5</i> -race-5'2	CAGGATCAGCCCGCGGTCCATGA	912
<i>Irf5</i> -race-3'1	CTTGCTGAGCAGCCTCCATTGAC	815
<i>Irf5</i> -race-3'2	GAAGTTCGAGTACCGCGGACGAA	789
<i>Irf6</i> -race-5'1	CTGCACCTCCAGGTCCGTCACTGA	908
<i>Irf6</i> -race-3'1	CAGCCCATCCCCCTGCTGGAGTT	972
<i>Irf9</i> -race-5'1	CAGGTCCCGCCGGGAGGAAGATG	996
<i>Irf9</i> -race-3'1	CTTCCACGTGGCGGTCAGCTACT	935

<sup>1</sup>Amplicon length was determined as the length of consensus sequence (obtained using Lasergene SeqMan Pro software – see Section 2.2), based on a minimum of six sequencing reads per amplicon.

**Table 2: Primers used in RT-PCR and QPCR studies of Atlantic cod IRF family members**

Primer name <sup>1</sup>	Sequence (5'-3')	Amplicon size (bp)	% Efficiency <sup>2</sup> (developmental/ macrophage)	Genomic location [linkage group (LG) gadMor2]
<i>Irf1</i> -fwd	AGAAGGACGCCAGTCTGTTC	101	89.1 / 91.8	LG 07
<i>Irf1</i> -rev	GCGGAAGTTGGCTTTCCATT			
<i>Irf2</i> -v1-fwd	CCAACCCGCTCATCAGTTAC	138	96.2 / 97.7	LG 10
<i>Irf2</i> -v1-rev	AGGTCAGAGGTCAGGACGAG			
<i>Irf2</i> -v2-fwd	TCCTCCCCTCTACCCTCT	127	95.2 / 96.5	
<i>Irf2</i> -v2-rev	TTTGAATCGCTTGACAGCAC			
<i>Irf3</i> -fwd	TTCTGGAGCCGCTGTAAGTT	143	105.7 / 101.2	LG 18
<i>Irf3</i> -rev	AGACAGCAGGACGGACATTT			
<i>Irf4a</i> -fwd <sup>a</sup>	TGTACCGTATCATCCCAGAGG	111	101.3 / 103.8	LG 08
<i>Irf4a</i> -rev <sup>a</sup>	AGTGGGGTATCTGGCTGTGA			
<i>Irf4b</i> -fwd <sup>a</sup>	TGGACATCACCGAACCTAC	106	91.3 / 90.3	LG 12
<i>Irf4b</i> -rev <sup>a</sup>	CATGACGAAAGCCATCTGAA			
<i>Irf5</i> -fwd	CTCTGCCAGTGCAAGGTGTA	143	94.6 / 95.1	LG 19
<i>Irf5</i> -rev	GAATCCCCCTTGTGGAACAT			
<i>Irf6</i> -fwd	GGAAGGTGAAGCTGTTCTGC	172	104.3 / 107.2	LG 13
<i>Irf6</i> -rev	ACCACCGGAATGATCTGAAC			
<i>Irf7</i> -fwd <sup>a</sup>	CATGTGCTTTGGGGAGAAGT	152	92.8 / 98.5	LG 09
<i>Irf7</i> -rev <sup>a</sup>	TCTGTAGGCTGACGTTGGTG			
<i>Irf8</i> -fwd <sup>a</sup>	TCGGGGAGGAACTACATGAC	158	90.7 / 91.1	LG 14
<i>Irf8</i> -rev <sup>a</sup>	GGCCATCTCGTCTGACATCT			
<i>Irf9</i> -fwd	GAGACGCCCAACAAGATCC	179	87.3 / 88.5	LG 23
<i>Irf9</i> -rev	AGGATGAGCTTCTGGGACTG			
<i>Irf10</i> -v1-fwd <sup>a</sup>	CCGAGAAGCCCAATAAACTG	143	97.0 / 99.3	LG 01
<i>Irf10</i> -v1-rev <sup>a</sup>	ATACTCCTCGCCAAAGCAGA			
<i>Irf10</i> -v2-fwd <sup>a</sup>	GGTCCAACGCAGTAACGATT	134	96.9 / 98.3	
<i>Irf10</i> -v2-rev <sup>a</sup>	ACTGTGGGAGACTGGCGTAT			
<i>EF1α</i> -fwd <sup>a</sup>	CCCTCCAGGACGTCTACAAG	150	88.5 / N/A	N/A
<i>EF1α</i> -rev <sup>a</sup>	GAGACTCGTGGTGCATCTCA			
<i>Tubb2</i> -fwd	GACCCACAGGAAGCTACAA	129	N/A / 89.8	N/A
<i>Tubb2</i> -rev	CATAGTGCCAGGCTCCAAGT			
<i>EIF3</i> -fwd <sup>b</sup>	AACTGTCCGTAGTCCGCAAG	125	N/A / 94.3	N/A
<i>EIF3</i> -rev <sup>b</sup>	CTGCTCAGCGAGAAACAGAA			
<i>rplp1</i> -fwd <sup>b</sup>	TCTGAAGCTAAGGCCCTCAA	141	92.5 / N/A	N/A
<i>rplp1</i> -rev <sup>b</sup>	ATCGTCTGGAGGATCAGAG			

<sup>1</sup> Primers noted with “a” were designed and first used in our previous study (Inkpen *et al.*, 2015); primers noted with “b” were previously used in Eslamloo *et al.* (2016).

<sup>2</sup> Percent amplification efficiency as in Pfaffl (2001) calculated using 7500 Fast software (Applied Biosystems).

**Table 3: Summary of IRF transcript expression responses to immune stimulation in Atlantic cod**

<b>Transcript</b>	<b>Macrophage response to poly(I:C), current study</b>	<b>Response to virus / viral PAMPs (previous studies)</b>	<b>Macrophage response to LPS, current study</b>	<b>Response to bacteria / bacterial PAMPs (previous studies)</b>
<i>GmIrf1</i>	Upregulated (12, 48 HPS).	Upregulated by poly(I:C) in spleen (Rise <i>et al.</i> , 2008; Hori <i>et al.</i> , 2012). Upregulated by nervous necrosis virus (NNV) in brain (Krasnov <i>et al.</i> , 2013)	Upregulated only at 72 HPS.	Upregulated by killed atypical <i>Aeromonas salmonicida</i> in spleen, head kidney (Feng <i>et al.</i> , 2009).
<i>GmIrf2-v1</i>	Upregulated (all time-points)	Upregulated by NNV in brain (Krasnov <i>et al.</i> , 2013).	Upregulated only at 72 HPS.	No previous studies.
<i>GmIrf2-v2</i>	Upregulated only at 24 HPS.	No previous studies.	No significant response.	
<i>GmIrf3</i>	Upregulated (all time-points).	Upregulated by poly(I:C) in larval cells; upregulated by NNV in brain (Krasnov <i>et al.</i> , 2013).	Upregulated only at 72 HPS.	No previous studies.
<i>GmIrf4a</i>	Upregulated (all time-points).	Downregulated by poly(I:C) (Inkpen <i>et al.</i> , 2015).	Upregulated only at 72 HPS.	No response to ASAL in spleen (Inkpen <i>et al.</i> , 2015).
<i>GmIrf4b</i>	Upregulated (all time-points).	Upregulated by poly(I:C) in spleen (Inkpen <i>et al.</i> , 2015).	No significant response.	Upregulated by ASAL in spleen (Inkpen <i>et al.</i> , 2015).
<i>GmIrf5</i>	Upregulated (12-24 HPS).	No previous studies.	Upregulated only at 12 HPS.	No previous studies.
<i>GmIrf6</i>	No significant response.	No previous studies.	No significant response.	No previous studies.
<i>GmIrf7</i>	Upregulated (12-48 HPS).	Upregulated by poly(I:C) in spleen (Rise <i>et al.</i> , 2008; Hori <i>et al.</i> , 2012). Upregulated by poly(I:C) in larval cells; upregulated by NNV in brain (Krasnov <i>et al.</i> , 2013).	No significant response.	Upregulated by ASAL in spleen (Inkpen <i>et al.</i> , 2015).
<i>GmIrf8</i>	No significant response.	Upregulated by poly(I:C) in spleen (Inkpen <i>et al.</i> , 2015).	No significant response.	Downregulated, then upregulated by ASAL in spleen (Inkpen <i>et al.</i> , 2015).
<i>GmIrf9</i>	Upregulated (all time-points).	Upregulated by NNV in brain (Krasnov <i>et al.</i> , 2013).	No significant response.	No previous studies.
<i>GmIrf10-v1</i>	Upregulated (all time-points).	Upregulated by poly(I:C) in spleen (Inkpen <i>et al.</i> , 2015). Upregulated by poly(I:C) in larval cells (Krasnov <i>et al.</i> , 2013).	No significant response.	Upregulated by ASAL in spleen (Inkpen <i>et al.</i> , 2015).
<i>GmIrf10-v2</i>	Upregulated (48-72 HPS).	No previous studies.	Upregulated only at 72 HPS.	No previous studies.

**Supplemental Figure 1: 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).

**Supplemental Figure 2: Nucleotide sequence of *Irf2-v1* cDNA and inferred amino acid translation.** (GenBank accession MH813456). 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 gadMor2 genome assembly. The stop codon is marked with an asterisk (\*). A putative polyadenylation signal is underlined.

**Supplemental Figure 3: Nucleotide sequence of *Irf2-v2* cDNA and inferred amino acid translation.** (GenBank accession MH813457). 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 gadMor2 genome assembly. The stop codon is marked with an

asterisk (\*). Two possible polyadenylation signals are underlined. (The use of polyadenylation signals other than AAUAAA is discussed in <sup>1</sup>MacDonald and Rodando, 2002).

**Supplemental Figure 4: Nucleotide sequence of *Irf3* cDNA and inferred amino acid translation.** (GenBank accession MH813458). 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 gadMor2 genome assembly. The stop codon is marked with an asterisk (\*).

**Supplemental Figure 5: Nucleotide sequence of *Irf5* cDNA and inferred amino acid translation.** (GenBank accession MH813459). 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 gadMor2 genome assembly. The stop codon is marked with an asterisk (\*).

**Supplemental Figure 6: Nucleotide sequence of *Irf6* cDNA and inferred amino acid translation.** (GenBank accession MH813460). 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 gadMor2 genome assembly. The stop codon is marked with an asterisk (\*). A putative polyadenylation signal is underlined.

**Supplemental Figure 7: Nucleotide sequence of *Irf9* cDNA and inferred amino acid translation.** (GenBank accession MH813461). 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 gadMor2 genome assembly. The stop codon is marked with an asterisk (\*). A putative polyadenylation signal is underlined.

**Supplemental Figure 8: Multiple sequence alignments of Atlantic cod IRF protein sequences.** Sequences were retrieved from the NCBI non-redundant protein database (see Supplemental Table 1). Alignments were carried out using the ClustalW algorithm in MEGA7 software (Kumar *et al.*, 2016). Identical amino acids are indicated by asterisks (\*); conservative substitutions are indicated by colons (:). Conserved tryptophan (W) residues are boxed.

**A) Alignment of predicted cod IRF DBD sequences with homologous sequences from other teleost fish.** IRF family subgroups are noted along the right side. **B) Alignment of predicted**

**full length cod IRF sequences.** DNA binding domain and IRF-associated domain are shaded in grey. Truncated splice variants (IRF2-v2, IRF10-v2) are omitted.

**Supplemental Figure 9: Composite agarose gel image of IRF transcript expression in 18**

**Atlantic cod tissues.** All gels are 1.7% agarose in TAE buffer. PCR was carried out using samples from two fish. Bl=blood, Ey=eye, Br=brain, Gi=gill, Hr=heart, HK=hematopoietic (head) kidney, PK=posterior (trunk) kidney, Sp=spleen, Li=liver, Go=gonad, St=stomach, PC=pyloric caecum, Mg=midgut, Hg=hindgut, DS=dorsal skin, VS=ventral skin, Mu=skeletal muscle, Fn=fin, C=no-template control.

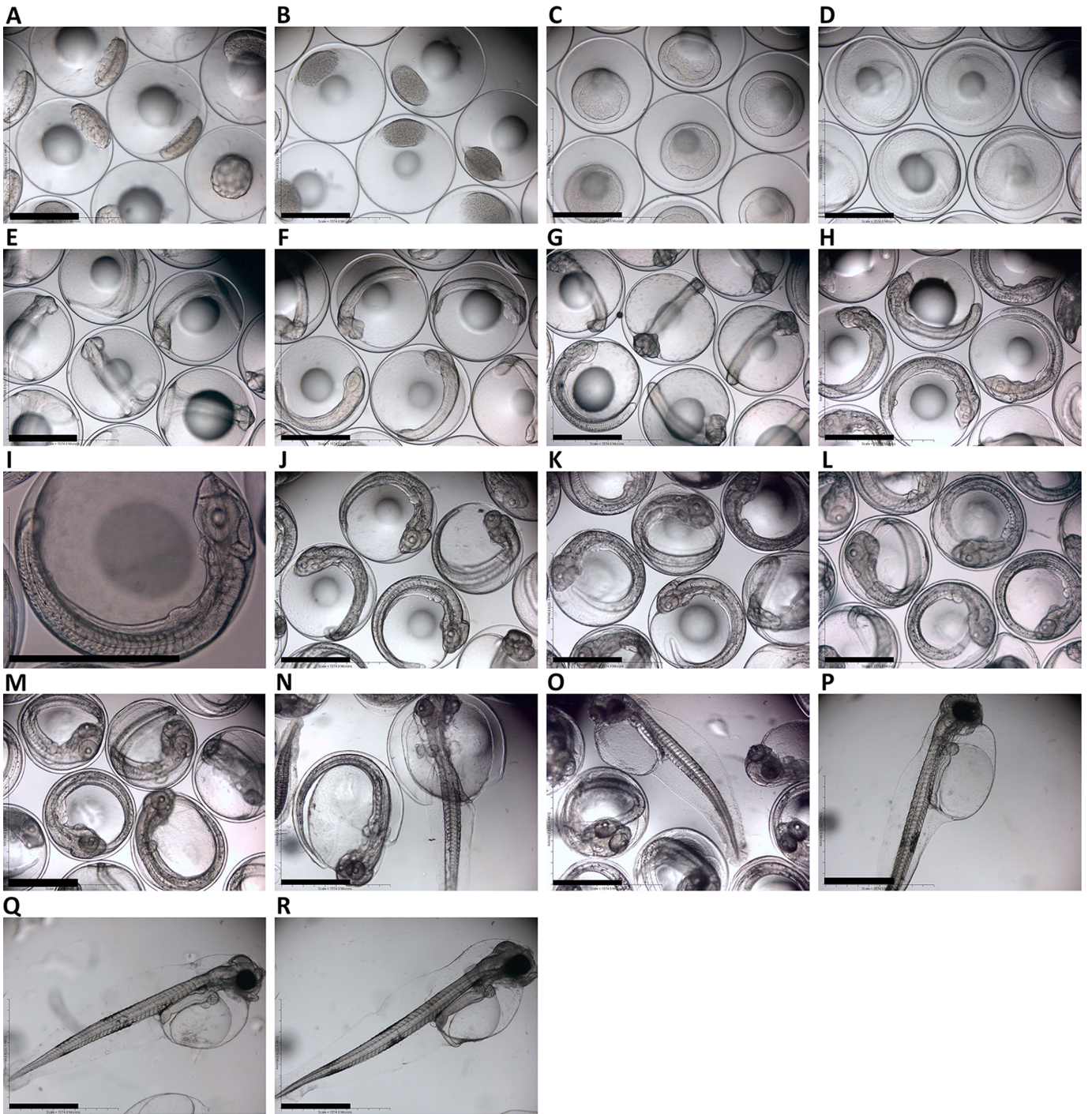
**Supplemental Figure 10: Atlantic cod macrophage transcript expression response of antiviral and antibacterial biomarkers to (A) poly(I:C) and (B) LPS, measured by QPCR.**

Data is presented as mean +/- SEM, normalized to the geometric mean of *EF1 $\alpha$*  and *rplp1* expression, with the lowest expressing sample set to RQ=1. An asterisk (\*) represents a significant difference between a poly(I:C) or LPS stimulated group and the time-matched PBS group ( $p < 0.05$ ). [\*\* =  $p < 0.01$ ; \*\*\* =  $p < 0.001$ ; \*\*\*\* =  $p < 0.0001$ ]. Capital letters and lower case letters represent similarity among PBS-treated groups and poly(I:C)/LPS-treated groups across time points, respectively. Fold change is calculated as (mean poly(I:C) or LPS)/(mean PBS).

<sup>1</sup>MacDonald CC, Redondo JL. 2002. Reexamining the polyadenylation signal: were we wrong about AAUAAA? *Mol Cell Endocrinol* 190(1-2):1-8.



Supplemental Figure 1



Supplemental Figure 2

1 tggatcaacgcagagtacatgggggcttttcttggcttgaagaggaagatgcgcgggagacaggaaagg  
 71 cggactgttaactcttttagtgctaaatgcggaaccagtggaatggttctccgctccaggtgatcgaaagg  
 141 ttctggtcgttctccgaggaactgaactcaaacctagagaataaaaatgtcgagcatagttgttggttcg  
 211 gatacattggctgcagactgctccgacagtgattctctgcaggataaccATGCCGGTAGAGAGGATGCGAA  
M P V E R M R 7  
intron 1  
 281 TGCGGCCGTGGTTGGAAGAGCAGATAAATCTTGTCTGATTCAGGACTCAAATGGGTCAACAAGGAAAA  
 M R P W L E E Q I N S C L I P G L K W V N K E K 31  
 351 GAAAAATCTTCCAGATCCCATGGATGCACGCGCGCGCCATGGCTGGGATCTGCAGAAAGACGCTCCGCTC  
 K I F Q I P W M H A A R H G W D L Q K D A P L 54  
intron 2  
 421 TTCATGAAGTGGGCCATACACACCGGTAAGTACCAGCTAGGCGTGGACCGCCCGGACCCCAAGACGTGGA  
 F M K W A I H T G K Y Q L G V D R P D P K T W 77  
 491 AGGCCAACTTCGCTGTGCGGTGAACAGCCTGCCGGACATGGAGGAGGTGAAGGACAAGAGCATCAAGAA  
 K A N F R C A V N S L P D M E E V K D K S I K K 100  
intron 3  
 561 GGGCACCAACGCCTTCAGGGTCTACAAGATGCTCTCCTCATCGGAGAGGAGCACCAAGAAAGGAAAGAAG  
 G T N A F R V Y K M L S S S E R S T K K G K K 124  
intron 4  
 631 AAGAAGGATGGGAAACCCAAGGCAGCTAAAGAAGGTGACTTCAAGGCCGAGGATGAGGGAGAGGAGCGGA  
 K K D G K P K A A K E G D F K A E D E G E E A 147  
intron 5  
 701 TGATGGGCATGGACATCAAGGAGGAAGTCTGTGAGGAAGAGGATGGCAGAGAGCCGGACGGCAACCTGGG  
 M M G M D I K E E V C E E E D G R E P D G N L G 170  
 771 TATCCACTACCAATAGCGGAGCACCTGATCACCAGCGAGCAGCTGCCATTTGTCTGCCGGACCATCGAG  
 I H Y P I G E H L I T S E Q L P F V C R T I E 194  
 841 GTGACTACTGAGAACGAAGAGCAGATGGTCAGCTCCTCCACTCCTACCCTCTCCAGATCTCTCCCGTGT  
 V T T E N E E Q M V S S S H S Y P L Q I S P V 217  
intron 6  
 911 CTTCACTACTGTGGCGGTGAGCGCGCCTGGATGCGGAACCTACAGCACGTCCATCCTCCGGATGCCGTCGG  
 S S Y C G R E R R L D A E L Q H V H P P D A V G 241  
 981 GCCCTCTGCCCGGCATGTCCAGCTTCATGGCCAACAAGCCAGCCTGCGCATCACCACCATCCAGGACCC  
 P S A R H V Q L H G Q Q A Q P A H H H H P G P 264  
 1051 CAACCCGCTCATCAGTTACCGCGCCGACAAATGGACCACGCGCTACGTCAAGTCAAAGCCCGGGACCACC  
 Q P A H Q L P R R Q M D H A L R Q V K A R D H 287  
 1121 GGGCACGAGGTGCGCGCCAGCGTCATCATGAAGACCTCCGACGTCAACTCGTCCTGACCTCTGACCTGTG  
 R A R G A R Q R H H E D L R R Q L V L T S D L C 311  
 1191 CACAGGCCGTCTGAGGTTGTTTAAACGCTGCGTAGGGAGGACGTCGACACGTCCTGTTGTCATGGTTGCTG  
 T G R L R L F K R C V G R T S T R P L S W L L 334  
 1261 TTTTGTGACACATCATCGGCGGTTGGGACGGAGTTTATAGAGGCCCTTCTGCCTAGatgtgtaata  
 F L L T H H R R L G R R F I E A P S A \* 353  
 1331 aagggaagtccttgtgaacctga<sub>(n)</sub>

Supplemental Figure 3

1 tggatcaacgcagagtacatgggggcttttgcttggcttgaagaggaagatgcgcgagacaggaaagg  
71 cggactgttaactcttttagtgctaaatgcggaaccagtggaatggttctccgctccaggtgatcgaaagc  
141 ttctggctgcttctccgaggaactgaactcaaacctagagaataaaaaatgtcgagcatagttggtggttcg  
211 gatacattggctgcagactgctccgacagtgattctctgcaggataaccATGCCGGTAGAGAGGATGCGAA  
M P V E R M R 7

*intron 1*

281 TGCGGCCGTGGTTGGAAGAGCAGATAAATTCTTGTCTGATTCCAGGACTCAAATGGGTCAACAAGGAAAA  
M R P W L E E Q I N S C L I P G L K W V N K E K 31  
351 GAAAATCTTCCAGATCCCATGGATGCACGCGGCGGCCATGGCTGGGATCTGCAGAAAGACGCTCCGCTC  
K I F Q I P W M H A A R H G W D L Q K D A P L 55

*intron 2*

421 TTCATGAAGTGGGCCATACACACCGGTAAGTACCAGCTAGGCGTGGACCGCCCGGACCCCAAGACGTGGA  
F M K W A I H T G K Y Q L G V D R P D P K T W 78  
491 AGGCCAACTTCCGCTGTGCGGTGAACAGCCTGCCGGACATGGAGGAGGTGAAGGACAAGAGCATCAAGAA  
K A N F R C A V N S L P D M E E V K D K S I K K 101

*intron 3*

561 GGGACCAACGCCTTCAGGGTCTACAAGATGCTCTCCTCATCGGAGAGGAGCACCAAGAAAGGAAAGAAG  
G T N A F R V Y K M L S S S E R S T K K G K K 125

*intron 4*

631 AAGAAGGATGGGAAACCCAAGGCAGCTAAAGAAGGTGACTTCAAGGCCGAGGATGAGGGAGAGGAGGCCGA  
K K D G K P K A A K E G D F K A E D E G E E A 148

*intron 5*

701 TGATGGGCATGGACATCAAGGAGGAAGTCTGTGAGGAAGAGGATGGCAGAGAGCCGGACGGCACCTGGG  
M M G M D I K E E V C E E E D G R E P D G N L G 171  
771 TATCCACTACCCAATAGGCGAGCACCTGATCACCAGCGAGCAGCTGCCATTTGTCTGCCGGACCATCGAG  
I H Y P I G E H L I T S E Q L P F V C R T I E 195  
841 GTGACTACTGAGAACGAAGAGCAGATGGTTCAGCTCCTCCCACTCCTACCCTCTCCAGATCTCTCCCGTGT  
V T T E N E E Q M V S S S H S Y P L Q I S P V 218  
911 CTTCTACTGTGGTGAGACCCCCACGCACAAGCACACACATATATATGTGCATACATATGTTATTATTA  
S S Y C G E T P H A Q A H T Y I C A Y I C Y Y \* 240  
981 Gtgctgtcaagcgattcaaatattttattcgcgattaatcgcatatagcaaaaaagattctatg  
1051 ctaaatatcccttgatttctttgtcccattcatttttctcattttgatgctcttatcaaatggagaagt  
1121 gcatcggcttaccttggtgcaaatgtttttttattgataacaacattggcatatactgatcaaaacaggac  
1191 gatacaaaaaaaaaagcctatagtgaattaaacgatgaacatacaacataactgccttgaacatagcagt  
1261 caggctactgcttttttggttttgaggcaaaagaaaaataataataataataataattacgttaacgcg  
1331 gataaaaaaattaacgcccgttaaaattggtttgctgtaaacgcccgttaataacgctttaaactgacagcac  
1401 tagttattggtgatattttataatgtagaaaaagacataggctcagatttttactcgctatactagggga  
1471 ttcacttctttggtgcttacttagtttagcctatgatggatactagggtggcgatagaggatatacaaa  
1541 ccgttgacgatataaaagtgtctaccggtcta(n)

## Supplemental Figure 4

1 acttgactacagcaccgccacattttctagtttaacactgttccctgtggggggtatgggcgacaacc  
 71 aggcATGGCTCAGCCCAAGCCACTCTTCATCCCCTGGTTGAAGGCCAGGATCGACAGCGGGCTGTTCCCC  
     M A Q P K P L F I P W L K A R I D S G L F P 22  
 141 GGGGTAGACTGGACCGACCCGGGCAGGACTGAGTTCTCCGTCCCCTGGAAGCACGCCCTCAGGCAGGACT  
     G V D W T D P G R T E F S V P W K H A L R Q D 45  
   intron 1  
 211 CCTCCAGCATTGACATCCTCATATTCAAGGCCTGGGCGGAGGTGAGTGGCAACGGCCAGGCCAAGGGGA  
     S S S I D I L I F K A W A E V S G N G O A Q G D 69  
 281 CCCCTCCGTCTGGAAGAGGAACCTCCGCAGCACCCCTTCGGGCCAAGAAGTTCGTCATGGTAACAAACAAC  
     P S V W K R N F R S T L R A K K F V M V T N N 92  
   intron 2  
 351 AGCAAGGAAAACGCCAACCCCATCAAGGTGTTCCGCTGGCCAAAGCAAAGCTCCTCCCCCTCCACAGAGA  
     S K E N A N P I K V F R W P K Q S S S P S T E 115  
 421 CACAGTTCAGGGTGTGAGTGTCTACAGAGGGGTGAAGGTTTCTGAGAATCTGGTGGATAACGAAGCTGG  
     T Q F R V S V F Y R G V K V S E N L V D N E A G 139  
   intron 3  
 491 TTTGAGACTTGTCTAGGAGGTGGACCACCAGTCAGGCCTGACCCTGATGACCCTGCCAACACCCCCCGTG  
     F R L V Y E V D H Q S G L T L M T L P T P P V 162  
 561 ACCTTCGACTCGGTGCAGAACCGTCTCACCCAGCAGGTGCTGGACAGTCTGGGGGCGGGGGTAGACGTGG  
     T F D S V Q N R L T Q Q V L D S L G A G V D V 185  
 631 GCGTGTGGGGCCCCGTGGTGTACAGCCATCGCCGTGGAACGCCAGAGCGTTCTGGAGCCGCTGTAAGTT  
     G V S G P V V Y S H R R G N A R A F W S R C K F 209  
 701 CGACCGCAGCCGAGAGCCGCAGCAGGTGGCCAAGATGGAGCCCCAGGCCCTCTCCGGTCCAAGACTTT  
     D R S R E P Q Q V A K M E P Q A L F R F Q D F 232  
   intron 4  
 771 GTGGGAGGGAATACAGGAATTCATCGCGGGGGGGAAATGTCCGTCCTGCTGTCTGTACATCTGTCTGGGGG  
     V G G I Q E F I A G G K C P S C C L Y I C L G 255  
   intron 5  
 841 AAGAGTGGCCGGACAGAAGGCCCTGGGAAAAGAAGCTTGTGATGGTGGAGGTGGTCTTCACGTCTCTGG  
     E E W P D R R P W E K K L V M V E V G L H V S G 279  
 911 AGTTCCTGAGGATGATGGCGGAGGATGGGGGCGCCTCCTCCCTGCAGTCAGTGGAAGTGCAGATGTCTCT  
     V P E D D G G G W G R L L P A V S G T A D V S 302  
 981 GACGGAGATGATGGAAGTGAATAGGTGACGAGAGGCCTCAGACCTCATGGGAAAAGTAGTCATTAAAT  
     D G D D G S G I G Q R E A S D L M G K V V I K 325  
 1051 GGAGTCCTAAGTCTTGGAGGGGGCGTCTTTATCATGTTTACATCTATTTTTATCATTTTTTACAAAGGAT  
     W S P K F L E G A S L S C L H L F L S F L Q R I 349  
 1121 ATTTATTTTATTGTTTAAACGGTCTTATTGTTATTGatgcaatcattattcagtagccatgaacataata  
     F I L L F N G S Y C Y \* 360  
 1261 caatggtatccatcgattatgtgaataccttcaatggttttgattataagatggtgcttttaattgattc  
 1331 tctttgattggctgtagagactgtccagttgtagtttacatgtgtattaaaaccattacacac

Supplemental Figure 5

1 gagacacccgggaaaaggggccaccagtcctctcgacatcgctctccacactggcaaccatccagctctagc  
 71 atacctgtcagctctactggcccccccccttcgagcctggcagctgcatccccccagctcctccacc  
 141 ATGAGCATCCAGCCCGCGGGATCCGGCTGAAGCCCTGGCTGTTGGCCCAGGTGAACGGAGGGCGCTACC  
 M S I Q P R R I R L K P W L L A Q V N G G R Y 23  
 211 CGGGCCTGAACTGGCTCAACCAGGAGCGGCTGTTCAGATACCCTGGAGACATGCCACACGCCACCTGCC  
 P G L N W L N Q E R L F Q I P W R H A T R H L P 47  
  
*intron 1*  
 281 CATGTCGGAGGAGGAGAACACCATCTTCAAGGCCTGGGCTTGGAGACAGGCAAATACCAGGAGGGAGTG  
 M S E E E N T I F K A W A L E T G K Y Q E G V 70  
 351 GATGAGCCCGACCCCGCCAAAGTGGAAAGGCTAACCTGCGCTGCGCCCTCAACAAGAGCCGCGAGTTCAAAC  
 D E P D P A K W K A N L R C A L N K S R E F K 93  
 421 TAATGTACGACGCGGACCAAGGAGAACCCTGTGAAGCCCTACAAAATCTACGAGGTGTGCGACCAGCCCCGG  
 L M Y D G T K E N P V K P Y K I Y E V C D Q P G 117  
  
*intron 2* *intron 3* *intron 4*  
 491 GAACGGAGATGCTGTGGATGAAGACGACGAGGAGATGGCCGAATCTTGACGATCTCACTATCGAACCCAGG  
 N G D A V D E D D E E M P N L D D L T I D P R 140  
 561 CCCAGCGACGTGTCCGCTTCCTCAACTACCACAGGGGTGAGGTGGAGATGGTCAACTTCACCCCTGCTC  
 P S D V S A F L N Y H R G E V E M V N F T P A 163  
  
*intron 5*  
 631 TCGGCGGCCCTTCGGTTCCCGGCGGTCATCCCCATGCCTCAGGTGACCGACCTGGACCTGAAGTTCGA  
 L G G P F G S P R V I P M P Q V T D L D L K F E 187  
 701 GTACCGCGGACGAAACCGCCGCTCCAGACGGTCAAGCAACCCCAAGGCTGCCGGCTGTACTACGGCCAC  
 Y R G R T A R S Q T V S N P Q G C R L Y Y G H 210  
 771 CTGGAGCCCCAGCCGGCCAGGTGGACCTGTTCCGCCCCGTCTCCCTGCAGCAGGTGCTGTTCCCGGCCA  
 L E P T P G Q V D L F G P V S L Q Q V L F P G 233  
 841 CGGCCGACCTGCAGAACGAGAAGCAGAGGTTCACACGGAGGCCCTGCTGGACGTTCATGGACCAGGGGCT  
 T A D L Q N Q K Q R F Y T E A L L D V M D R G L 257  
 911 GATCCTGGAGATCTGGGAGCAGGACATCTACGCCGTGCGCTCTGCCAGTGCAAGGTGTACTGGTCGGGG  
 I L E I W E Q D I Y A V R L C Q C K V Y W S G 280  
 981 CCCGGTGTGTCGGAGCACGGCCGCCCAACCCCATGGAGCGGGAGAAGAAGATCAAGGTGTTACAGCTCA  
 P G V S E H G P P N P M E R E K K I K V F S L 303  
  
*intron 6*  
 1051 ACCAGTTCCTGGAAAGGGCTGATCATGTTCCACAAGGGGATTCCTCCGAACCCCTCCCCTTTCGAGGTCTA  
 N Q F L E G L I M F H K G D S P N P P P F E V Y 327  
  
  
*intron 7*  
 1121 CTTCTGCTTTGGTGAAGACTGGCCGATAGGAAACCGAAGGAGAAGAAGCTCATTATTGTTTCAGTTGGTC  
 F C F G E D W P D R K P K E K K L I I V Q V V 350  
 1191 CCCGTGGTGGCCAGGATCTGACGGAGATGTTCTCCGGGAGCTCAGCTGGTCCACGGACAGCATCCGCT  
 P V V A R I L T E M F S G E L S W S T D S I R 373  
 1261 TGCAGATCTCCAATCCTGACGTCAAGGACCAGACGGTGGAGCAGTTCAAAGAGCTCCAGCGACTCCTCCA  
 L Q I S N P D V K D Q T V E Q F K E L Q R L L Q 397  
 1331 GAGCCAGCATGCTCAGGCCCTTGGGGCCGCATGTCCCTTGAggaaaaaaagtaatgacagaccagcaac  
 S Q H A Q G P W G P H V P \* 410  
 1401 cgtccaaaaaagaactacacttcttgggaactctacaagaggaacttaaagtggtgctgctgctggc  
 1471 cccagagactatattt

Supplemental Figure 6

1 gatcagaaaaaacccaactcgcaaggtgctgtttttctttaaagtcactattaaccaccgacctggtgat  
71 ttgatcctttcccccttccctggagacaaaattctgactgtagaacggatcgatagtgtagcctaacctt  
141 cggattcccgcctgtttttgggcaactaaaggagccctttcgacttcctgcttgacacagtgccagaaATC  
M 1

211 TCGGTCACTCCGCGACGCGTCCGCCATAAGCCATGGTTGGTGGCCAGGTGGACACCGGTTCGCTACCCGG  
S V T P R R V R L K P W L V A Q V D T G R Y P 24  
281 GGCTGATGTGGATCGACCGGGACGCCATGCGATTCAGGATCCCGTGGAAAGCATGCTACAAGACACAGCC  
G L M W I D R D A M R F R I P W K H A T R H T P 48

intron 1

351 TCAGCATGAAGAGGAGAACACCATTTTTAAGTCATGGGCAGTGGAGACCGGAAATTCAGGAGGGAGTT  
Q H E E E N T I F K A W A V E T G K F Q E G V 71  
421 GATGAACCTGATCCTGCAAGTGGAAAGCTCAGCTCCGATGTGCCCTCAACAAAAGCCGGGAATTC AAC  
D E P D P A K W K A Q L R C A L N K S R E F N 94  
491 TCTTCTACGATGGCAACAGGAGATCCCATGAACCCCTGAAGATATATGACGTGTGTGACATCCCCC  
L F Y D G T K E I P M N P L K I Y D V C D I P P 118

intron 2

561 GCCTCCAGCACCACAGTCCAGCAGTCCCTCCCTCACATCATGTGGTCTCCGTCCGGCTCCGAGTCTGTC  
P P S T T G S S P S P H I M W S P S G S E S S 141  
631 AACAGCCACAGAGCTGCCCTCCTTCGGTGGAGCGCTGGATGCCAAGGAGGAGCCCTGCCACGTGTGGC  
N Q P Q S C P P S V E R W M P K E E P C H V W 164  
701 CCAAAGAGGAGCCCGTGGACGTGGAGATGCAGCCCATCCCCCTGCTGGAGTTGACCCCGCCCGCCCA  
P K E E P V D V E M Q P I P L L E L T P A P P Q 188  
771 GAACCCCATCATGCTGCCGACTCCATGCCCCGACGCCATGTTCCGCTCCCCGGAGATGTGGATCAGCTCG  
N P I M L P H S M P D A M F A S P E M W I S S 211

intron 3

841 CTGCCAGTGCATTTCTCTTTTTTGTGCTCCCTTTTGGTTTGGTGTGCGTCTGCTCAGTGACGGACCTGGAGG  
L P V T F P L F V V P F W F A C R S V T D L E 234  
911 TGCAGTTCCTGTACCCGCGGGAAGGAGGTGAGCCCCCTGCTGATGGTGGAGCAACCCCGAGGGTGC CGGCT  
V Q F L Y R G K E V S P L L M V S N P Q G C R L 258  
981 GTTCTACGGGACCTGGGCCCATGGTGAACAGGAGGAGCTGTTGGCCCCGTGAGCTGGAGCAGCTG  
F Y G D L G P M V N Q E E L F G P V S L E Q L 281  
1051 CGCTTCCCAACAGGAGCACATCAACAAGCAGCAGCGCTCTTCAACAACCGCTGCTCGACGTCA  
R F P T T E H I T N D K Q R V F T N R L L D V 304  
1121 TGGACCGGGGTGATCCTGGAGGTGAGCGGCCACGACATCTACCCGTGCGACTCTGCCAGTGCAAGGT  
M D R G L I L E V S G H D I Y A V R L C Q C K V 328  
1191 GTACTGGTCCGGCCCTGCGCCCCAACCCCAACGCCCTAACCTCATAGAGCGCCAGCGGAAGGTGAAG  
Y W S G P C A P N P N A P N L I E R Q R K V K 351

intron 4

1261 CTGTTCTGCCTGGAGTCTTTTCTCAGTGGTTCATAGCGCACCGCTGGCCAGACATCCATCCCCCCCG  
L F C L E S F L S G V I A H Q R G Q T S I P P 374  
1331 ACTTTGACATCAACCTGTGCTTTGGAGAGGAGTGGCCCGACGGCAGGCCAAAGAGCGTAAACTCATCAT  
D F D I N L C F G E E W P D G R P K E R K L I M 398

intron 5

1401 GGTTCAGTTCATTCGGTGGTGGCCCGCATGATCAGCGAGATGTTCTCCGGCGACAGCAGCGGTCTCTC  
V Q I I P V V A R M I S E M F S G D S T R S F 421  
1471 GACAGCGGCAGCGTGCCTGCGAGATCTCCATCCCCGACATCAAGGACAACATAGTGACCCACCTGAAGC  
D S G S V R L Q I S I P D I K D N I V T H L K 444  
1541 AGCTCTACTGCTGCTCCTCAACACCAGGGCCAGGAGGGCTGGCCCGCGCAGCGGGCCAGAACCGCA  
Q L Y C L L L N H Q G Q E G W P A Q P G Q N Q H 468  
1611 CCTCATCTCGGCCCTGACGGGCCAGTGAagcctacccccccagacacgggtccccgtaccgaggggga  
L I S A L Q G Q \* 476  
1681 aaccagccccctagtataacagacataagctactcttgtatatttagagatatctgtaaatattttaacttc  
1751 tttgaagaaatatttttagcttctttgaaggaatccttctacttctcctggagacggtaaaaacttgta  
1821 gtgttttactgcctcaagtagtgccgagcgagaaaactgtgtgctgcacagtggtttgctgtagaactcg  
1891 ctccattgggacacagctcttctctgctcaagggaacacttgatcatttaacctagttaatgagttgag  
1961 acttaacctctggcatattatacaaaatgggtcagtggttgtgaagccgctcaccacctacgaaacct  
2031 gtgctgttaacctaggaacagagggcagtggttccctagtcttacttaccagcttttaatatgtggcgcgt  
2101 ttcctgaaataagtggtgagtgtaaacctgaaatggcaactaattatgggtcttgggtctgcaagtg  
2171 acggtccggtctcatagctatgctgtttgtcatacatttctcctcaggtaactctgccaacttttttc  
2241 aacctgactttactgaaataacctgttcttaccagttataacttttggctgaaattggattttcttggty  
2311 tctgtcttatctttgtatctgtaagctcttcttctcagggacatgactattagcctattggaat  
2381 ttttactgtcaatcctaaacaaatgtcaggtaaataatcgttttataactgtgtagaatgtcggcacttt  
2451 taatcctgtgtatttttggtaata(n)

Supplemental Figure 7

1 ggaggcttgggcttccagagaactgctccagtgcattaccgctactgggcgcacaccagctggacggggg  
71 aaacgcaactgaggtaaacgaaacaaaaaactcgcttccatttgtgtagaattcgggcgatatagcgata  
141 gggacgcctgtacttttgaacatgtatagacgaaccgacgggcacaaactttaataactatthaatccaaa  
211 aatggtgctccgatcgttttctgtttcattgtcttatgggcatgcttttagactcgcataggggtgtgt  
281 gcagagcgtcattgggtcaatttcagatatctgctgggtttacttagtgtggggtaaaataaaaaaagat  
351 ttgatattgtgacatcatgacgttgactgtctgcttgcagagccatcATGGGACCTGTAAAGCCCGCAA

M G P C K A R N 8

intron 1

421 CACGCGTAAACTGCGCTCCTGGATGGTGGAAACAGTGTGACCAGCGGTAGGTACCCGGTCTGGTGTGGGAC  
T R K L R S W M V E Q V T S G R Y P G L V W D 31  
491 GACGATGAGCAGACCATGTTCCGCATTCTTGGAAACACGCCGGGAAGCAAGACTTCAGGAGCGACGAAG  
D D E Q T M F R I P W K H A G K Q D F R S D E 54

intron 2

561 ACGCCGCCATCTTTAAGTSCGTGGGCGAGGTTCAAGGCAAGATGACGGACGAGAGCCGCAACGAGCCTGC  
D A A I F K A W A E F K G K M T D E S R N E P A 78  
631 GATCTGGAAGACGCGGCTGCGCTGCGCCCTGAACAAGAGCCTGGAGTTCGAGGAGGTGGCGGAGCGTGCC  
I W K T R L R C A L N K S L E F E E V A E R A 101  
701 CAGCTGGACATCTCCGAGCCCTCAAGGTCTACCGCCTGGTGC CGGACTCGGAGCAGGGCCTCACGGGCA  
Q L D I S E P F K V Y R L V P D S E Q G L T G 124

intron 3

771 ACGAAGAATTGACGGCCAAACGAGCGGGGGGAAGAGCAGAAGAAGAACCAAGAGGAGAGCCACCGAATC  
N E E L T A K R A G G K S R R R R T K R R A T E S 148

intron 4

841 CTCCGAGGAAGACGAGGAGCTGCCTGAGAAGCAGATGAAGGAGGAATCAGTGACGGCACCCCTGTGTGTG  
S E E D E E L P E K Q M K E E S V T A P L S V 171  
911 GAGGAGATCTTGTGCGAGGGGTGAAGCAACGCTCCAGGTGGAGGGCGGCAACCAGCAGATTTTCATATCCC  
E E I L S R G E A T L Q V E G G N Q Q I F I S 194

intron 5

intron 6

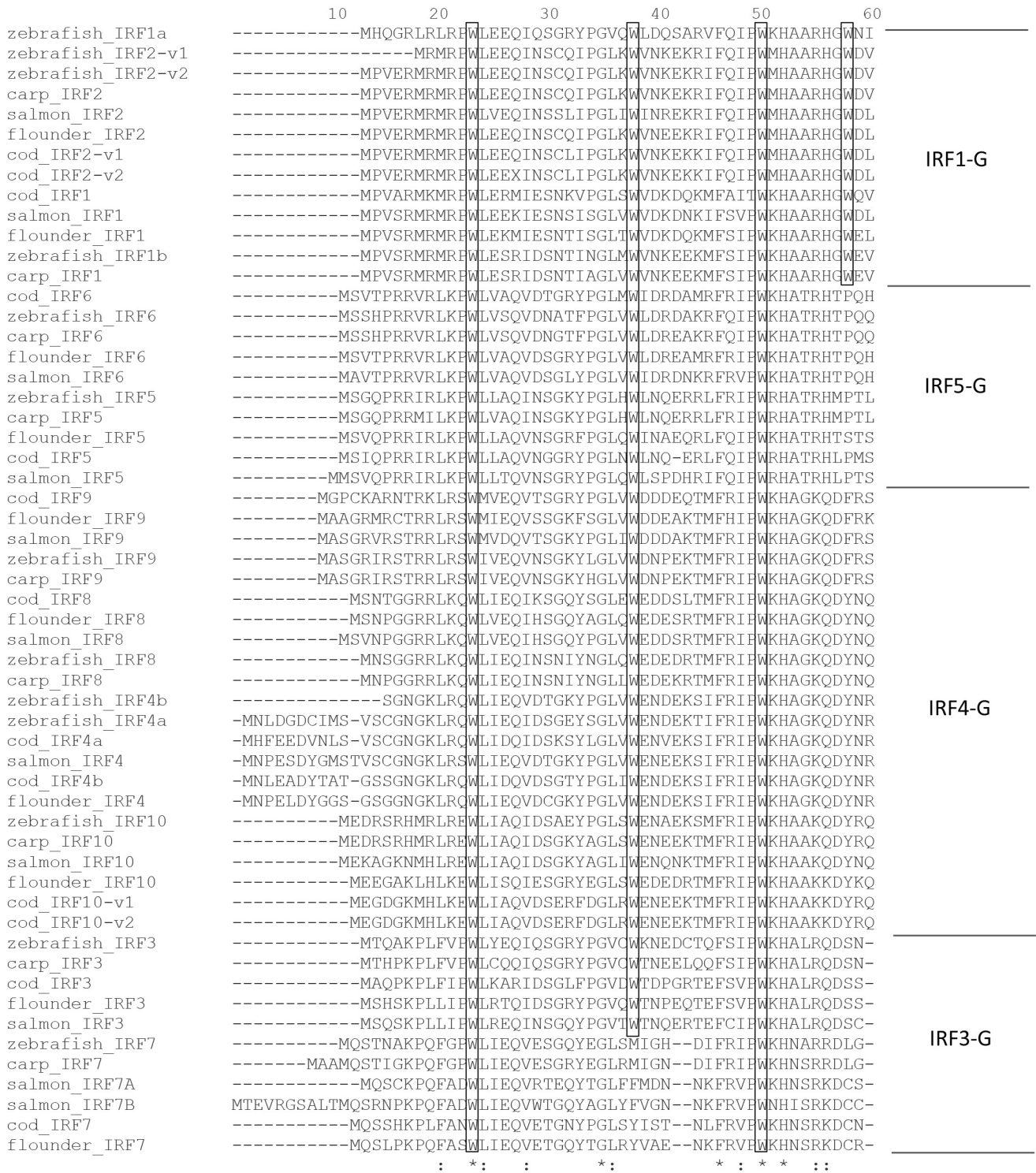
981 AAGGGAGCGGATTTGGATTTCCGGATAGAGACCAACCCTCCTCCAAGTGTGCACTCTCCAGAGGAGGA  
Q G S G I G F P D R D Q P S S K C C R S P E E D 218  
1051 CAGCTTCCACGTGGCGGTGACTACTCCGCCACGAGGTCTGTCCCGCAGGTGCAGGGCAAGGACGTC  
S F H V A V S Y S G H E V L S R E V Q G K D V 241  
1121 CGGATCAGTACCACACGCCCTCGCCGCTGCCCCCAACCCCGGCACCCCTTATGGGGGGTTCCCCGCA  
R I T Y H T P S P L P P T P A P L M G G F P R 264  
1191 TCTTCTCCCGGAGGCCCCGGCGGCTGCAGGGGGCGGGACCTGGTCCAGCTGCTGCCCTGCATGGA  
I F L P E A P A G L Q G G R D L V Q L L P C M E 288  
1261 GAAGGGCGTGGTGTGACCTCCAACCACACAGGGGTCTACCTGCGGCGTACTGCCGGGGCCGGGTGAC  
K G V V L T S N H T G V Y L R R Y C R G R V Y 311  
1331 TGGATCGGCCCGCAGCCGCCGCCCTCCCGAGACGCCCAACAAGATCCAGCGTGACGCGGAGCCCGTGC  
W I G P H A A A P P E T P N K I Q R D A E P V 334

intron 7

1401 TGATCTTCAGCAAGGAGGCGTTCGCCAGGAGCTGGAGGCCTACCAGCAACAGGGAGGAGACCCGCCGCG  
L I F S K E A F R Q E L E A Y Q Q Q G G D P P R 358  
1471 ATGCGACGTGACGCTCTGCTTCGCGCAGGAACTCGTACCACGGACACCAGTCCCAGAAGCTCATCCTC  
C D V T L C F G E E L V T T D D Q S Q K L I L 381  
1541 GTCAGGTACGGTGTGCGCACAACCCCGCCTCTCTCTCTCTCTCCCGCCGTTCAAGGGAGACCTCA  
V K V R L S R Q P R L F L F S P G R F K G D L 404  
1611 ATGCTCTCTGCTCTCTCTCACTTACGGTCAAGGGCATGTGAGGGAAGGCCCAAGTGAGGCCAGGACTTG  
N V S R S L S L H G Q G H V R E G P S E A R T \* 427  
1681 Aggcctgtgccccgacaaaattaacattaattaatggtgagatataacagcagtc aaatgcctgtatgtg  
1751 ctgctgcttcggtccagttcactgttaaaagggatagccttttgccttttcaattgagaactcgataac  
1821 attttttttttgaataaactgttttatgtagtgaagtgtggaaaatgtgataactgtgatgacctgt  
1891 gataactttttggaaacttctgtcacaataaactaaatgcaaaactggaagtaacaaaagatatt  
1961 ctgggtattcagggtgatatttcaatcta(n)

Supplemental Figure 8

A





Supplemental Figure 8

	70	80	90	100	110	120	
zebrafish_IRF1a	DKDATLFRN	WAIHTGRYKPGIDKDPDKT	W	KANFR	CALNSL	-TDVKELQDRS	IKKGHNAFR
zebrafish_IRF2-v1	EKDAPLFRN	WAIHTGKYQPG-DKDPDKT	W	KANFR	CAMNSL	-PDIEEVKDKS	NKKGTNAFR
zebrafish_IRF2-v2	EKDAPLFRN	WAIHTGKYQPG-DKDPDKT	W	KANFR	CAMNSL	-PDIEEVKDKS	NKKGTNAFR
carp_IRF2	EKDAPLFRN	WAIHTGKYQPG-DKDPDKT	W	KANFR	CAMNSL	-PDIEEVKDKS	IKKGTNAFR
salmon_IRF2	EKDAPLFMN	WAIHTGKYQLGIDKDPDKT	W	KANFR	CAMNSL	-PDIEEVKDKS	IKKGTNAFR
flounder_IRF2	EKDAPLFMR	WAIHTGKYQHGVDRPDPKT	W	KANFR	CAMNSL	-PDIEEVKDKS	IKKGTNAFR
cod_IRF2-v1	QKDAPLFMK	WAIHTGKYQLGVDRPDPKT	W	KANFR	CAVNSL	-PDMEEVKDKS	IKKGTNAFR
cod_IRF2-v2	QKDAPLFMK	WAIHTGKYQLGVDRPDPKT	W	KANFR	CAVNSL	-PDMEEVKDKS	IKKGTNAFR
cod_IRF1	EKDasLFKH	WAIHTGKFKEGVDESDDPK	W	KANFR	CAMNSL	-PDVEQVKGKNV	NKGQAVR
salmon_IRF1	NKDACLFKQ	WAMHTGKFIQGETKTDPKT	W	KANFR	CAMNSL	-PDIEEVKDKS	INRGS GAVR
flounder_IRF1	DKDASLFKK	WAIHTGKYTEGQT-SDPKT	W	KANFR	CAMNSL	-PDIEEVKDKS	SIHKGGQAVR
zebrafish_IRF1b	DKDACLFKQ	WAIHTGKYKEGVTQDPDKT	W	KANFR	CAMNSL	-PDIEEVKDKS	INKGCGAVR
carp_IRF1	DKDACLFKQ	WAIHTGKFRGVTTPDPDKT	W	KANFR	CAMNSL	-PDIEEVKDKS	INKGCGAVR
cod_IRF6	EEENTIFKA	WAVETGKFQEGVDEPDPAK	W	KAQLR	CALNKS	-REFNLFYDGTKE	I PMNPLK
zebrafish_IRF6	EEENTIFKA	WAVETGKYQEGVDEPDPAK	W	KAQLR	CALNKS	-REFNLIYDGTKE	VP MNPLK
carp_IRF6	EEENTIFKA	WAVETGKYQEGVDEPDPAK	W	KAQLR	CALNKS	-REFNLIYDGTKE	VP MNPLK
flounder_IRF6	EDEDTIFKA	WAVETGKFQEGVDEPDPAK	W	KAQLR	CALNKS	-REFNLFYDGTKE	VP MNPLK
salmon_IRF6	EEENTIFKA	WAVETGKFQEGIDDPDPAK	W	KAQLR	CALNKS	-REFNLFYDGTKE	VP MNPLK
zebrafish_IRF5	EEENTIFKA	WALETGKYQEGIDEPDPAK	W	KANLR	CALNKS	-REFRLNYDGT	KDTPVQPYK
cod_IRF5	EEENTIFKA	WALETGKYQEGVDEPDPAK	W	KANLR	CALNKS	-REFGLHYDGT	KDTPVQPYK
salmon_IRF5	EEENTIFKA	WALETGKYQEGVDEPDPAK	W	KANLR	CALNKS	-REFGLHYDGT	KDTPVQPYK
cod_IRF9	DEENTVFKA	WALETGKYQEGVDEPDPAK	W	KANLR	CALNKS	-REFQLKYDGT	KETPVRPYK
flounder_IRF9	EEENTIFKA	WALETGKYQEGVDEPDPAK	W	KANLR	CALNKS	-REFKLMYDGT	KENPVKPYK
salmon_IRF9	EEENTIFKA	WALETGKYQEGVDEPDPAK	W	KANLR	CALNKS	-REFKLYDGT	KETPVQPYK
zebrafish_IRF9	DEDAALFKA	WAEFKGKMTDE-SRNEPAT	W	KTRLR	CALNKS	-LEFEEVAERA	QLDISEPYK
carp_IRF9	DEDAALFKA	WAEFKGKLSDG-GQDNPAI	W	KTRLR	CALNKS	-PEFEEVDRA	QLDISEPYK
cod_IRF8	EEDGAIFKA	WAVFKGKLSDG-GRVDPAS	W	KTRLR	CALNKS	-PEFREVPER	SQLDISEPYK
flounder_IRF8	EEDAAIFKA	WAAEFKGLMEN-GNSDPAS	W	KTRLR	CALNKS	-PEFSEVTER	SQLDISEPYK
salmon_IRF8	EEDAAIFKA	WAEFKGKLED-GNSDPAS	W	KTRLR	CALNKS	-PEFSEVTER	SQLDISEPYK
zebrafish_IRF8	EVDASIFKA	WAVFKGKFKEG-EKAEPAT	W	KTRLR	CALNKS	-PDFEEVTDR	SQLDISEPYK
carp_IRF8	EVDASIFKA	WAVFKGKFKEG-EKAEPAT	W	KTRLR	CALNKS	-PDFEEVTDR	SQLDISEPYK
zebrafish_IRF4b	EVDASIFKA	WAVFKGKFKEG-EKAEPAT	W	KTRLR	CALNKS	-PDFEEVTDR	SQLDISEPYK
zebrafish_IRF4a	DEDAALFKA	WALFKGKFRGVDKDPPT	W	KTRLR	CALNKS	-NDFEELVDR	SQLDISEPYK
cod_IRF4a	DEDAALFKA	WALFKGKFRGVDKDPPT	W	KTRLR	CALNKS	-NDFEELVDR	SQLDISEPYK
salmon_IRF4	DEDAALFKA	WALFKGKFRGVDKDPPT	W	KTRLR	CALNKS	-NDFEELVDR	SQLDISEPYK
cod_IRF4b	DEDAALFKA	WALFKGKFRGVDKDPPT	W	KTRLR	CALNKS	-NDFEELVDR	SQLDISEPYK
flounder_IRF4	DEDAALFKA	WALFKGKFRGVDKDPPT	W	KTRLR	CALNKS	-NDFEELVDR	SQLDISEPYK
zebrafish_IRF10	NQDAALFKA	WAMYKGFQEGRDKADPST	W	KTRLR	CALNKS	-TDFQEVPER	SQLDISEPYK
carp_IRF10	NQDAALFKA	WAMYKGFQEGRDKADPST	W	KTRLR	CALNKS	-TDFQEVPER	SQLDISEPYK
salmon_IRF10	NEDAALFKA	WAVYKGYREGRDKADPT	W	KTRLR	CALNKS	-TDFQEVPER	SQLDISEPYK
flounder_IRF10	TEDAALFKA	WAVYKGYIEGRDKADPT	W	KTRLR	CALNKS	-TDFQEVPER	NQLDISEPYK
cod_IRF10-v1	QDDAALFKA	WAVYKGYKVGSDKDNPT	W	KTRLR	CALNKS	-TDFQEVPHL	NQLDISEPYK
cod_IRF10-v2	QDDAALFKA	WAVYKGYKVGSDKDNPT	W	KTRLR	CALNKS	-TDFQEVPHL	NQLDISEPYK
zebrafish_IRF3	SDDVLI FKA	WQTSAAAGDGR-LNGDPSV	W	KRNFR	SALRA	--KGFKMISDKK	-NDGADPHK
carp_IRF3	SDDVLI FKA	WQTSAAAGDGR-INGDHSV	W	KRNFR	SALRA	--KGFKMIFDNK	-NDAANPHK
cod_IRF3	SIDILIFKA	WAEVS--GNGQ-AQGDPSV	W	KRNFR	STLRA	--KKFVMVTNNS	-KENANPIK
flounder_IRF3	DTDILIFKA	WAEVS--GNGR-AHGDA	W	KRNFR	SALRS	--KGFKMVNDKK	-NETADPHK
salmon_IRF3	SDDVLI FKA	WAEVS---NGR-VQGDPSI	W	KRNFR	SALRA	--KGFKMLLDNK	-NDAANPNK
zebrafish_IRF7	DADV KIFKE	WAI VSGKINE--YPNDKAK	W	KTNFR	CALHSL	-KNFEMLEDHS	-KDPDDQHK
carp_IRF7	DEDIKIFKE	WAVVSGKINE--HPNDKAK	W	KTNFR	CALYSL	-KNFEMLEDHS	-KDPDDQHK
salmon_IRF7A	EDDRKIFRA	WAVVSGKITE--HPNDKAK	W	KTNFR	SALNSL	-CRRFKMVEDHS	-KDSNDPHK
salmon_IRF7B	EDDSKIFRA	WAVVSGKINT--HPNDKAK	W	KTNFR	CVLNNSL	-TKRFMMVEDHS	-KDSDDPHK
cod_IRF7	DEDCKIFRA	WAVASGKIHE--FPNDKAK	W	KTNFR	CALKNL	-NKRFRMSKDNS	-KNSDDPHK
flounder_IRF7	DEDSKIFRA	WAVASGKINE--FPNDKAK	W	KTNFR	CALNNSL	-SVRFKMIEDNS	-KHSDDPHK
	:	:*	**	.	:	**	.*:.*:

IRF1-G

IRF5-G

IRF4-G

IRF3-G

**Supplemental Figure 8**

**B**

atlantic_cod_IRF1	-----MPVARMKMRFLERMIESNKVPGLSIVVDDKDQKMFAITMKHAARHGQVEK	50
atlantic_cod_IRF2-v1	-----MPVERMRMRFLLEEQINSCLIPGLKWNKEKKIFQIPMMHAARHGQDLQK	50
atlantic_cod_IRF3	-----MAQPKPLFIPLKARIDSGLFPGLVWTDPGRTEFSVPMKHALRQDSS-SI	49
atlantic_cod_IRF7	-----MQSSHKPLFANWLEIQVETGNYPGLSYIST--NLFRVPMKHNSRKDCN-DE	48
atlantic_cod_IRF9	-----MGPCARNRTRKLRSMVEQVTSGRYPGLVWDDDEQTMFRIPMKHAGKQDFRSE	54
atlantic_cod_IRF8	-----MSNTGGRRLKQWLEIQIKSGQYSGLEWEDDSLTMFRIPMKHAGKQDYNQEV	51
atlantic_cod_IRF10-v1	-----MEGDGKMLKELIAQVDSERFDGLRWENEKTMFRIPMKHAACKDYRQDD	51
atlantic_cod_IRF4a	MHFEEDVNLVSCNGKLRQWLIQVDSKSYLGLVWENVEKSI FRIPMKHAGKQDYNRDE	60
atlantic_cod_IRF4b	MNLEADYTATGSSNGKLRQWLIQVDSGYTPGLWENDEKSI FRIPMKHAGKQDYNRDE	60
atlantic_cod_IRF5	-----MSIQPRRIRLKEPLLAQVNGGRYPGLVWLNQER-LFQIPWRHATRHLPMSEE	51
atlantic_cod_IRF6	-----MSVTPRRVRLKEPLVAQVDTGRYPGLWIIDRDAMRFRIPMKHATRHTPQHHEE	52



atlantic_cod_IRF1	DASLFKHVAIHTGKFKEGVDESDDPKKMKANFRCAVNSLPDVEQVKGKGNVNGQQAVRVYK	110
atlantic_cod_IRF2-v1	DAPLFMKVAIHTGKYQLGVDRDPKTKANFRCAVNSLPDMEEVKDKSIKKGTFNAFRVYK	110
atlantic_cod_IRF3	DILIFKVAEVSNGN---GQAQGDPSVWKRNRFRSTLRACK--FVMVTNNSKENANPIKVFRR	104
atlantic_cod_IRF7	DCKIFRAVAVASGKIHEF--PNDKAKWKTNFRCALKNLNKFRMSKDNSKNSDDPHKIYE	106
atlantic_cod_IRF9	DAALFKVAEAFKGMKTDDES-RNEPAIWKTRRLCALNKSLFEFEVAERAQLDISEPFKVYR	113
atlantic_cod_IRF8	DASIFKVAVAVFKGFKEGE-KAEPATWKTRRLCALNKSPDFEEVTDRLSDISEPYKVYR	110
atlantic_cod_IRF10-v1	DAALFKVAVAVYKGYKVGSDKNPTMKTRRLCALNKSTDFQEVPHLNQLDISEPYKVYR	111
atlantic_cod_IRF4a	DAALFKVAALFKDKYKEGVDPDPPTMKTRRLCALNKSNDFDELVDRLSDITEPYKVYR	120
atlantic_cod_IRF4b	DAALFKVAALFKGKFKREGIDKADPPTMKTRRLCALNKSNDFEELVDRLSDISDPYKVYR	120
atlantic_cod_IRF5	ENTIFKVALETGKYQEGVDEPDPAKWKANLRCALNKSRFEKLMYDGTENPVKPKYIYE	111
atlantic_cod_IRF6	ENTIFKVALETGKYQEGVDEPDPAKWKAQLRCALNKSRFENLFDYDGTKEIPMNPLKIYD	112

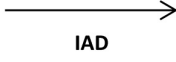
: :\* \*\* .:\* : : \*\* :\*:\* : . . . : : :

atlantic_cod_IRF1	MVEVTATKDRRTKTKDGKRRNKLTARLEETDFSDTQSCEDQH--PPHYDDTCSPQENTI	168
atlantic_cod_IRF2-v1	MLSSSERSTKKGKDKG-----KPKAAKEGDFKAEDEGEEAMMGMDIKEEVCEEEDGRE	165
atlantic_cod_IRF3	WPKQSS-----	110
atlantic_cod_IRF7	IINREAAAYQPSPEEDMVPVIYSSP--TESYPPGHEQNILEQLMTLDLLEPCQQTVGGEQ	164
atlantic_cod_IRF9	LVPDSEQGLTGNEELTAKRAGGKSRRTKRRATESSEEDE---E-----	154
atlantic_cod_IRF8	IVPEEEQKLGKTTAMV-----TTAGDIADLDCSSAEL---EELI-----	146
atlantic_cod_IRF10-v1	IESDQRAESDQTYSRV---VVVQTYASLPQSQLA---D---QWERFEERQEESHGAL	160
atlantic_cod_IRF4a	IIEGVRKPKINKVSAIFRWLSS-----	144
atlantic_cod_IRF4b	IIEGDKRRRQEDSPLSPLSYPSYPALHSQIPHCMNPPE---SGWREFYPEQ-----	170
atlantic_cod_IRF5	VCDQPGNGDAVDED-----	125
atlantic_cod_IRF6	VCDIPPPSTTGSSPSPHIMWSPSGSESSNQPPSCPPSVE---RWMPKEEPCHVWPKEE	168

**DBD**

atlantic_cod_IRF1	DST-EQDMISLP-----LSASEVPDFENVITIGNDSNNADYF-----YR-----	206
atlantic_cod_IRF2-v1	PDG--NLGIHYPIGEHLITSEQLPFVVCQTEVTTENEEMQMS-----SSHYSYPLQISPV	217
atlantic_cod_IRF3	-----	110
atlantic_cod_IRF7	WAESYQQSAIGLVYATNQATGETMHA-----MQTQPQLQPQQQAYYPVNPVPPV	215
atlantic_cod_IRF9	-----LPEKQMK-EESVTAPLVEEILSRGEATLQVEGGNQIIFISQSGSIGFPDRQPS	208
atlantic_cod_IRF8	-----KVASTDYPSAIKRSYSP-QEDGFNVQASPE	176
atlantic_cod_IRF10-v1	WR-----EHTYCGSEDSQAHSHIPL-DPSLLS-----	186
atlantic_cod_IRF4a	-----	144
atlantic_cod_IRF4b	-A--FLPELHIPQCSYPPHPWQGPPIENAXYQIKGSFYSYTHADVQP-----SAFTLDPG	222
atlantic_cod_IRF5	----DEEMPNLDDLIDRPSDVSAFLNY--HRGEVEMV-----NF-TPALGGPFGSPR	172
atlantic_cod_IRF6	PV--DVEMQPIPLLELTPAPPQNPIMLPH--SMPDAMFASPEMVISS-LPVTFFLFWVPF	223

**Supplemental Figure 8**

atlantic_cod_IRF1	-----RFEVSPEHPPEFEDAE-----	222
atlantic_cod_IRF2-v1	SSYCGRE-----RRLDAELQHVHPPDAVGPSAR-----	245
atlantic_cod_IRF3	-----SPSTETQFRVSVFYRGVKVSENLVDNEAGFRLVYEVDHQS-----	150
atlantic_cod_IRF7	L-DSGLQPSLFDLEISVHYRKVEMLKTQVSWPRVQ-----	249
atlantic_cod_IRF9	SKCCRSPEEDSFHVAVSYSGHEVLSRE-VQGKDVIRITYHTPSPL-----	251
atlantic_cod_IRF8	YWSHGSIPVFSQMMISFYGGQLMHSTVTSHPGECRISPVLPPQRAVARGYS-----	228
atlantic_cod_IRF10-v1	--PTLAI-SDFRMELTLFYRGEPMELTSSSPEGCFILQGCVPLGNE-----	230
atlantic_cod_IRF4a	-----	144
atlantic_cod_IRF4b	MRPADPL-SDLRLHVSFVSRDALVREVTISNPKGCHLIPWALEEKAYVSPGAPDLVPLPP	281
atlantic_cod_IRF5	VIPMPQV-TDLDLKFE-YR-GRTARSQTVSNPQGCRLLYYGHLEPTPGQV-----	218
atlantic_cod_IRF6	WFACRSV-TDLEVQFL-YRGKEVSPLLMVSNPQGCRLLFYGDLGPMVNGE-----	270
	 IAD	
atlantic_cod_IRF1	----ELLKLCQQLEPETNWMQSS-----SDDLSSGLHSDSNYSPHSQ----WSD	264
atlantic_cod_IRF2-v1	----HVQLHGQQAQPAHHHPGPQ----PAHQLPRRQMDHALRQVKARDHRAR----GAR	293
atlantic_cod_IRF3	-----	150
atlantic_cod_IRF7	-----	249
atlantic_cod_IRF9	-----	251
atlantic_cod_IRF8	-----	228
atlantic_cod_IRF10-v1	-----	230
atlantic_cod_IRF4a	-----	144
atlantic_cod_IRF4b	EGLTLQRMAGEEGPPSSLAMQGVRLWMTPEGLYARRQCQESVYWKEGVSPYKDKLNEMER	341
atlantic_cod_IRF5	-----	218
atlantic_cod_IRF6	-----	270
atlantic_cod_IRF1	TSSGEDLD-----MRLYTDLSTGTECYSPET-----WNMFPTPIYQQ	301
atlantic_cod_IRF2-v1	QRHHEDLRRQ-----LVLTSDLCTGPSEVV-----	318
atlantic_cod_IRF3	-----	150
atlantic_cod_IRF7	-----LHYGNEAT-----ELQARP	263
atlantic_cod_IRF9	-----PPTPAPLMGG-----FPR	264
atlantic_cod_IRF8	-----SDTMQS	234
atlantic_cod_IRF10-v1	-----RIYG-----PCSAQQ	240
atlantic_cod_IRF4a	-----	144
atlantic_cod_IRF4b	EVNCKVLDTQDFLTEIQSYGLHGRPIPPFQALLCFGDECVDTERPRRSLTVQVEPLFARQ	401
atlantic_cod_IRF5	-----DLFG-----PVSLQQ	228
atlantic_cod_IRF6	-----ELFG-----PVSLEQ	280
atlantic_cod_IRF1	INFHP-----	306
atlantic_cod_IRF2-v1	-----	318
atlantic_cod_IRF3	-----GLTLMTLPTPPVTFDSVQNRLTQQVLDLSLGGVVDVGVSGPVVYSHR	196
atlantic_cod_IRF7	ICF-----PPTDTLRDHKQVEFTNRISSIQRGLLLEVRESGLYACR	305
atlantic_cod_IRF9	IFL-----PEAPAGLQGGRDVQLLPCMEKGVVLTNSHTGVYLR	304
atlantic_cod_IRF8	VHF-----PPADLIDNERQRQVTCCKLLGHLERGVIVRANREGVFIKR	276
atlantic_cod_IRF10-v1	LSLPS-----PASLGPLEPGVARALGQLLSHLERGVLLWVAPDGLFIKR	284
atlantic_cod_IRF4a	-----	144
atlantic_cod_IRF4b	LFYYAQQTGHHYRGRYEHGVEPEHISPFEDYQRAIS--HHHHHSGMMQEITGQDIYAIR	459
atlantic_cod_IRF5	VLF-----PGTADLQNKQRFYTEALLDVMRGLILEIWEQDIYAVR	270
atlantic_cod_IRF6	LRF-----PTTEHITNDKQRFVFNRLLDVMRGLILEVSGHDIYAVR	322
atlantic_cod_IRF1	-----	306
atlantic_cod_IRF2-v1	-----	318
atlantic_cod_IRF3	RGNARAFWSRCKFDRS---REPQQVAKME-PQALFRFQDFVGGIQEFIGGK--CPSCCL	250
atlantic_cod_IRF7	QDRCHVFASADPSQA---SPDPQKLPQNT-LVELLSFEKFKELKEFKENRRG-SPEYVV	361
atlantic_cod_IRF9	YCRGRVYWIGPHAAAP--PETPNKIQRDAEPVLIFSKEAFRQLEAYQQGGD-PPRCDV	361
atlantic_cod_IRF8	LCQSRVWVSGHGEHGQHHGVPVTCCLERDA-VVKIFDTGRFLHALQLHQEQIIP-APDPTV	334
atlantic_cod_IRF10-v1	FCQGRVYWSGPLAPHT---EKPNKLERDR-TCKLLDMPVFNELQNYMQRKGP-QPNYEI	339
atlantic_cod_IRF4a	-----	144
atlantic_cod_IRF4b	LCQCKVWVSGPGVLEQ---GPPNPMEREI-RIRVFSLNKFLQALILYQKGEAPTPPPEI	515
atlantic_cod_IRF5	LCQCKVWVSGPGVSEH---GPPNPMEREK-KIKVFSLNQFLEGLIMFHKGDSNPPPPFEV	326
atlantic_cod_IRF6	LCQCKVWVSGPCAPNP---NAPNLIERQR-KVKLFCLESFLSGVIAHQRGQTSIPPDFDI	378

## Supplemental Figure 8

```
atlantic_cod_IRF1
atlantic_cod_IRF2-v1
atlantic_cod_IRF3
atlantic_cod_IRF7
atlantic_cod_IRF9
atlantic_cod_IRF8
atlantic_cod_IRF10-v1
atlantic_cod_IRF4a
atlantic_cod_IRF4b
atlantic_cod_IRF5
atlantic_cod_IRF6
```

-----

```
YICIGEEWPPDRR-WEKKLVMVEVGLHVSVP-----EDDGG-----WGRLIPAVSGT
NMCFGEKFPDGK-LEKKLIVKVVPLICRYF-----YEMAQVEGASSLDSTNVSLQISHD
TLCFGEELVTTDD-QSOKLILVKVRLSRQPLFLFSPGRFKGDLNVSRS-----LSLHGQ
TLCFGEELHDLN-AKNKILVQITAMNCQQL-----LEAVNMRVQSYNHS--PSEMS
DLCFGEIYPDAKVSATMKTITVHVVPLFAMEL-----LQRFQLEVEAEP-----DVHTP
-----
YFCFGEWPPDRK-KEKKLIVQVVPVVARIL-----TEMFSGEL--SWSTDSIRLQISNP
YFCFGEWPPDRK-KEKKLIVQVVPVVARIL-----TEMFSGEL--SWSTDSIRLQISNP
NLCFGEWPPDRK-KEKKLIMVQIIPVVARMI-----SEMFGDSTRSFDSSVRLQISIP
```

-----

```
atlantic_cod_IRF1
atlantic_cod_IRF2-v1
atlantic_cod_IRF3
atlantic_cod_IRF7
atlantic_cod_IRF9
atlantic_cod_IRF8
atlantic_cod_IRF10-v1
atlantic_cod_IRF4a
atlantic_cod_IRF4b
atlantic_cod_IRF5
atlantic_cod_IRF6
```

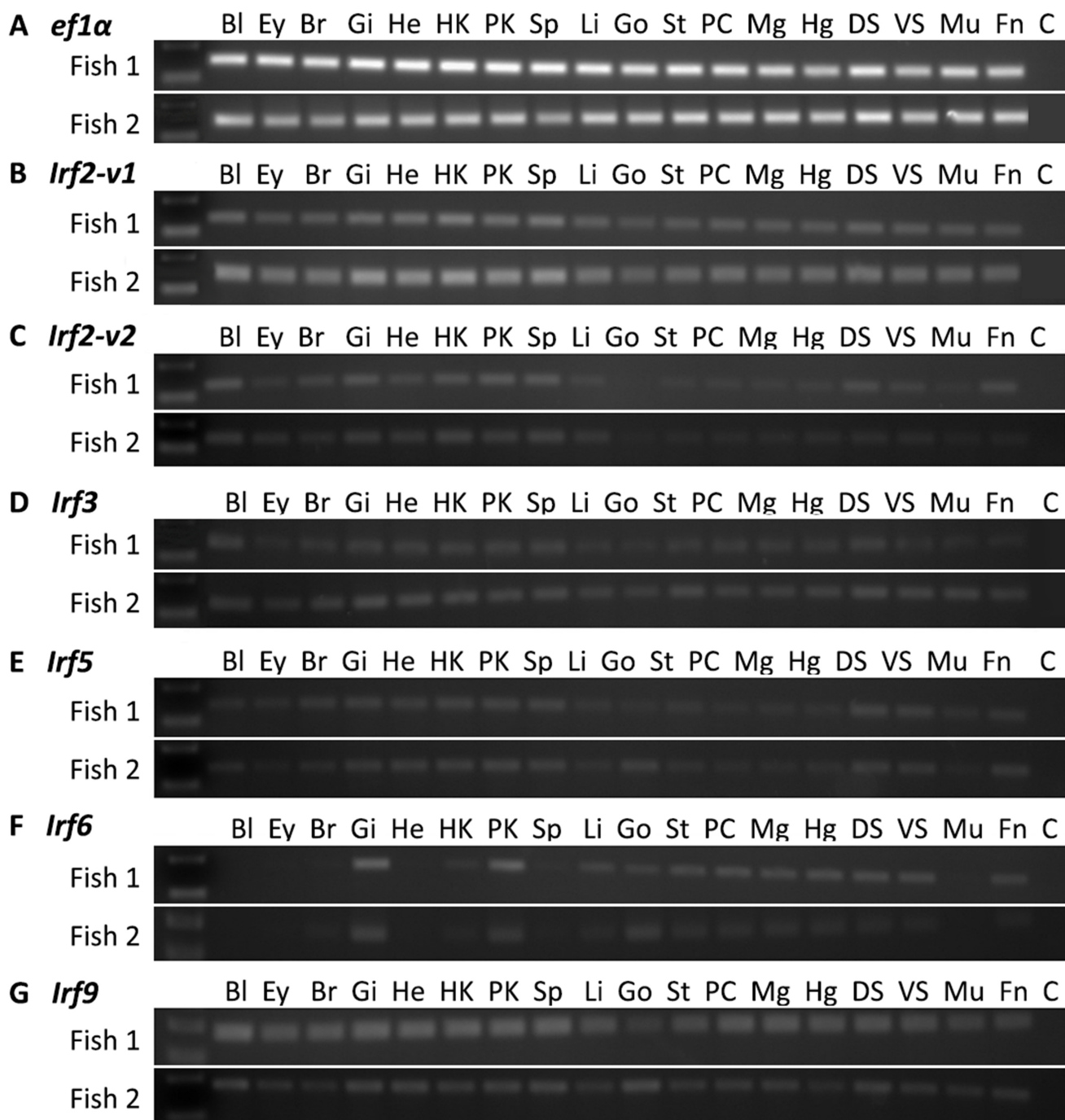
-----

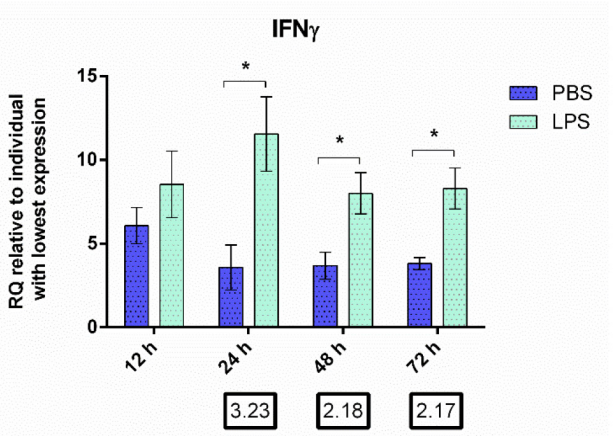
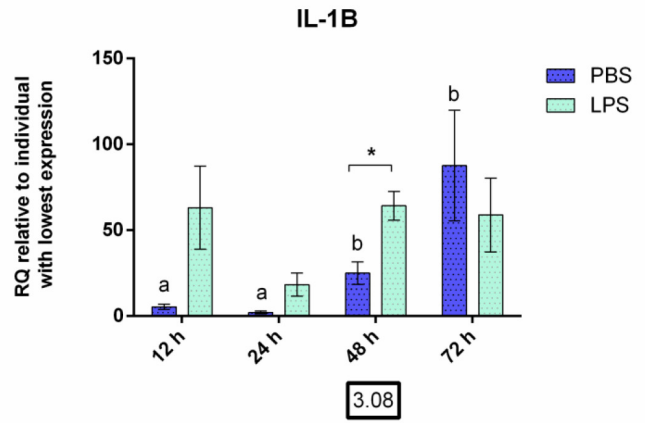
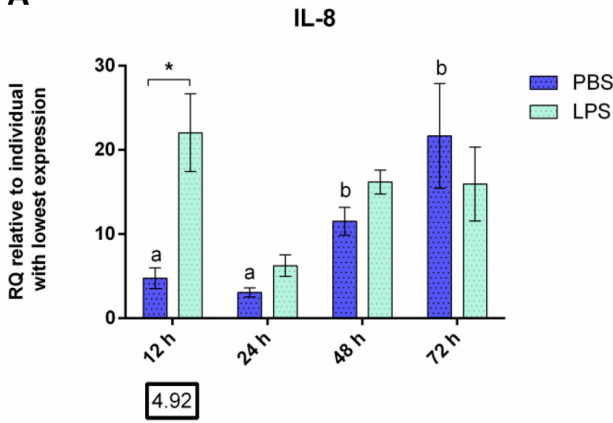
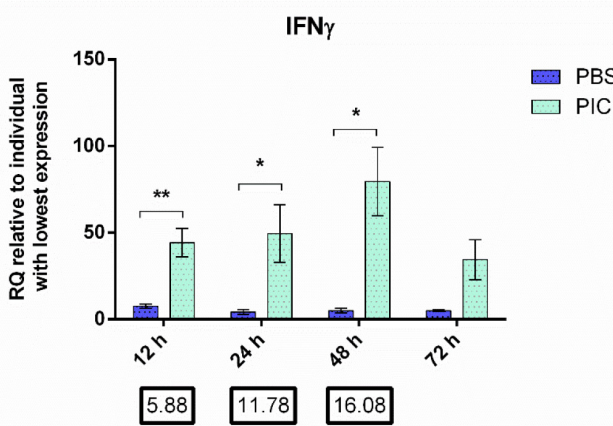
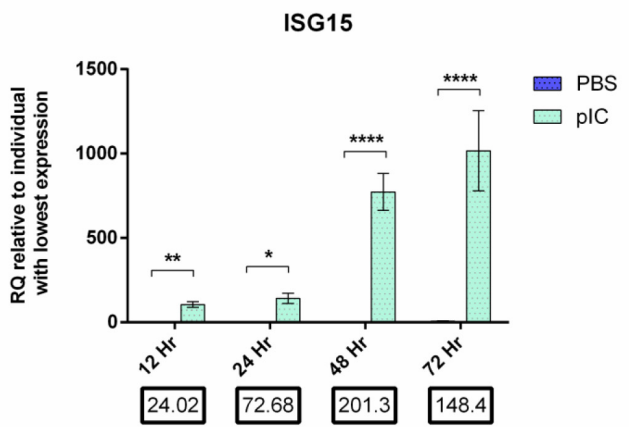
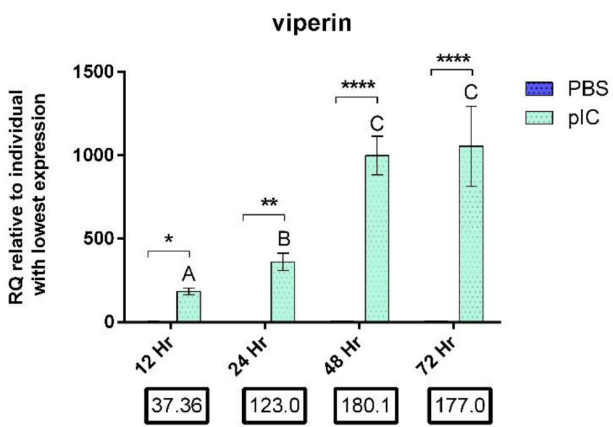
```
ADVSDGDDGSGIGQREASDLMG---KVVIKWSPKFLGASLSCLHLFLSFLQRIFILL
S-----LY---DLISSAFGLPGSQVAPQLVGHY-----
GHVREGPS--EART-----
DEMAS-----DOMARIYQDLCSYSAPQRTDCYRDNMTTA-----
KEAKDEM-----
-----
D-LKDHTV--EQFKEIHRLLQSOSA--HPNWPNTN-----
D-VKDQTV--EQFKELQRLIQSQA--QGPGPHVP-----
D-IKDNIV--THLKQLYCLLNHGQ--QEGWPAQPQGNQHLISALQGQ-----
```



IAD

Supplemental Figure 9



**A****B**

Supplemental Table 1: Teleost fish IRF amino acid sequences used in phylogenetic analysis

Protein name	Species name (common name)	GenBank accession no.
IRF1	<i>Ctenopharyngodon idella</i> (grass carp)	ADF57571
	<i>Gadus morhua</i> (Atlantic cod)	ADG85733
	<i>Oncorhynchus mykiss</i> (rainbow trout)	AAM77843
	<i>Paralichthys olivaceus</i> (Japanese flounder)	BAA83468
	<i>Salmo salar</i> (Atlantic salmon)	ACI68339
	<i>Tetraodon nigroviridis</i> (pufferfish)	CAF87280
IRF1a	<i>Danio rerio</i> (zebrafish)	AAI15341
IRF1b <sup>1</sup>		AAH85555
IRF2	<i>Ctenopharyngodon idella</i> (grass carp)	AFV99156
	<i>Danio rerio</i> (zebrafish)	AAI64907
	<i>Gadus morhua</i> (Atlantic cod)	<b>MH813456 (v1)</b>
	<i>Oncorhynchus mykiss</i> (rainbow trout)	AAK53987
	<i>Paralichthys olivaceus</i> (Japanese flounder)	ADZ96216
	<i>Salmo salar</i> (Atlantic salmon)	ACI33066
IRF3	<i>Ctenopharyngodon idella</i> (grass carp)	AHL29306
	<i>Danio rerio</i> (zebrafish)	ABY91288
	<i>Gadus morhua</i> (Atlantic cod)	<b>MH813458</b>
	<i>Oncorhynchus mykiss</i> (rainbow trout)	CAH56618
	<i>Paralichthys olivaceus</i> (Japanese flounder)	ACY69212
	<i>Salmo salar</i> (Atlantic salmon)	ACN11005
	<i>Tetraodon nigroviridis</i> (pufferfish)	CAG07572
IRF4	<i>Ctenopharyngodon idella</i> (grass carp)	AMT92193
	<i>Oncorhynchus mykiss</i> (rainbow trout)	CAH56622
	<i>Paralichthys olivaceus</i> (Japanese flounder)	AEY55358
	<i>Salmo salar</i> (Atlantic salmon)	ACI33264
	<i>Tetraodon nigroviridis</i> (pufferfish)	CAF98086
IRF4a	<i>Danio rerio</i> (zebrafish)	NP_001116182
	<i>Gadus morhua</i> (Atlantic cod)	AJR33027
IRF4b	<i>Danio rerio</i> (zebrafish)	CAI11951
	<i>Gadus morhua</i> (Atlantic cod)	AJR33026
IRF5	<i>Ctenopharyngodon idella</i> (grass carp)	ACT83675
	<i>Danio rerio</i> (zebrafish)	ABY91289
	<i>Gadus morhua</i> (Atlantic cod)	<b>MH813459</b>
	<i>Oncorhynchus mykiss</i> (rainbow trout)	CDQ64449
	<i>Paralichthys olivaceus</i> (Japanese flounder)	AEY55357
	<i>Salmo salar</i> (Atlantic salmon)	ACI33029
IRF6	<i>Tetraodon nigroviridis</i> (pufferfish)	CAF90666
	<i>Ctenopharyngodon idella</i> (grass carp)	AMT92196
	<i>Danio rerio</i> (zebrafish)	AAH56772
	<i>Gadus morhua</i> (Atlantic cod)	<b>MH813460</b>
	<i>Oncorhynchus mykiss</i> (rainbow trout)	CDQ66830
	<i>Paralichthys olivaceus</i> (Japanese flounder)	XP_019942637
	<i>Salmo salar</i> (Atlantic salmon)	XP_013988514
IRF7	<i>Tetraodon nigroviridis</i> (pufferfish)	CAG06823
	<i>Ctenopharyngodon idella</i> (grass carp)	ACS34986
	<i>Danio rerio</i> (zebrafish)	NP_956971

	<i>Gadus morhua</i> (Atlantic cod)	AJR33028
	<i>Oncorhynchus mykiss</i> (rainbow trout)	CAH56623
	<i>Paralichthys olivaceus</i> (Japanese flounder)	ACY69214
	<i>Tetraodon nigroviridis</i> (pufferfish)	CAG02387
IRF7A	<i>Salmo salar</i> (Atlantic salmon)	ACI33478
IRF7B		ACL68545
IRF8	<i>Ctenopharyngodon idella</i> (grass carp)	AMT92197
	<i>Danio rerio</i> (zebrafish)	AAH75963
	<i>Gadus morhua</i> (Atlantic cod)	AJR33029
	<i>Oncorhynchus mykiss</i> (rainbow trout)	ALS92677
	<i>Paralichthys olivaceus</i> (Japanese flounder)	AFE18694
	<i>Salmo salar</i> (Atlantic salmon)	XP_013982533
	<i>Tetraodon nigroviridis</i> (pufferfish)	CAF99526
IRF9	<i>Ctenopharyngodon idella</i> (grass carp)	AMT92198
	<i>Danio rerio</i> (zebrafish)	AAH81591
	<i>Gadus morhua</i> (Atlantic cod)	<b>MH813461</b>
	<i>Oncorhynchus mykiss</i> (rainbow trout)	CDQ76373
	<i>Paralichthys olivaceus</i> (Japanese flounder)	AHV91018
	<i>Salmo salar</i> (Atlantic salmon)	ACN11040
	<i>Tetraodon nigroviridis</i> (pufferfish)	AFR24260
IRF10	<i>Ctenopharyngodon idella</i> (grass carp)	ACT83676
	<i>Danio rerio</i> (zebrafish)	NP_998044
	<i>Gadus morhua</i> (Atlantic cod)	AJR33030 (v1)
	<i>Paralichthys olivaceus</i> (Japanese flounder)	BAI63219
	<i>Salmo salar</i> (Atlantic salmon)	XP_014000943
	<i>Tetraodon nigroviridis</i> (pufferfish)	CAG04088
IRF10a	<i>Oncorhynchus mykiss</i> (rainbow trout)	CDM74110
IRF10b		CDM74111

<sup>1</sup>Zebrafish IRF1b is also called IRF11.

\*Bolded accession numbers represent **nucleotide** sequences characterized in the current study, as amino acid sequences are not yet available.



Supplemental Table 2: Normalizer transcript Ct values for early development and macrophage QPCR studies in Atlantic cod

Embryonic / larval QPCR normalizer Ct <sup>1</sup> values				Macrophage QPCR normalizer Ct values			
Sample Name <sup>2</sup>	<i>Tubb2</i>	<i>EIF3</i>	Geometric mean	Sample Name <sup>3</sup>	<i>EF1α</i>	<i>rplp1</i>	Geometric mean
D00T1	26.190	26.125	26.157	A12C	18.419	19.033	18.724
D00T2	25.596	26.086	25.840	B12C	18.055	18.075	18.065
D00T3	27.651	24.975	26.279	C12C	17.909	17.989	17.949
D01T1	24.720	25.866	25.286	D12C	18.159	18.304	18.231
D01T2	24.911	25.558	25.233	E12C	18.408	18.475	18.441
D01T3	26.317	24.809	25.552	F12C	18.116	18.205	18.160
D02T1	22.691	25.565	24.085	A12P	18.449	18.949	18.697
D02T2	22.761	25.278	23.986	B12P	18.628	18.107	18.366
D02T3	24.514	24.098	24.305	C12P	18.781	18.473	18.626
D03T1	21.230	24.347	22.735	D12P	18.827	18.542	18.684
D03T2	21.287	24.303	22.745	E12P	19.240	18.792	19.015
D03T3	22.608	23.032	22.819	F12P	17.977	17.986	17.982
D04T1	21.946	24.040	22.969	A12L	18.219	18.908	18.561
D04T2	21.322	24.028	22.635	B12L	18.161	18.207	18.184
D04T3	22.131	22.597	22.363	C12L	18.234	18.538	18.386
D05T1	21.035	24.062	22.498	D12L	18.626	18.591	18.608
D05T2	21.808	23.958	22.857	E12L	18.567	18.722	18.644
D05T3	22.454	22.674	22.564	F12L	17.590	17.973	17.781
D06T1	21.137	23.337	22.210	B24C	17.560	17.933	17.746
D06T2	21.002	23.844	22.378	C24C	18.317	18.235	18.276
D06T3	23.220	23.646	23.432	D24C	18.665	18.926	18.795
D07T1	21.253	24.090	22.627	E24C	18.585	18.676	18.630
D07T2	21.322	23.819	22.536	F24C	17.557	17.750	17.653
D07T3	23.189	23.158	23.173	A24P	19.107	19.723	19.413
D08T1	21.041	23.544	22.257	B24P	18.412	17.972	18.191
D08T2	21.011	23.625	22.280	C24P	18.847	18.387	18.616
D08T3	23.114	23.556	23.334	D24P	19.529	19.364	19.446
D09T1	22.014	23.454	22.723	E24P	19.720	18.877	19.294
D09T2	21.870	23.233	22.541	F24P	18.051	17.896	17.974
D09T3	23.099	23.587	23.342	A24L	18.168	19.869	18.999
D10T1	22.712	24.052	23.373	B24L	18.053	18.093	18.073
D10T2	21.654	23.093	22.362	C24L	18.566	18.700	18.633
D10T3	23.622	23.974	23.797	D24L	18.940	19.508	19.222
D11T1	21.768	24.224	22.963	E24L	18.761	18.741	18.751
D11T2	22.253	23.420	22.829	F24L	17.279	17.638	17.458
D11T3	23.336	23.100	23.218	A48C	18.575	19.489	19.027
D12T1	22.436	24.033	23.221	B48C	18.281	18.526	18.403
D12T2	22.756	24.191	23.462	C48C	18.036	18.236	18.136
D12T3	23.598	23.238	23.417	D48C	18.572	18.693	18.632

D13T1	22.874	23.961	23.411	E48C	18.323	18.776	18.548
D13T2	22.775	23.276	23.024	F48C	17.145	17.912	17.524
D13T3	23.583	23.253	23.417	A48P	19.428	19.883	19.654
D14T1	21.966	23.349	22.647	B48P	18.946	18.642	18.793
D14T2	23.153	21.796	22.464	C48P	18.958	18.532	18.743
D14T3	23.961	23.541	23.750	D48P	19.387	18.995	19.190
D15T1	22.750	23.929	23.332	E48P	19.019	18.884	18.951
D15T2	23.663	24.308	23.983	F48P	18.627	18.795	18.711
D15T3	23.867	23.415	23.640	A48L	19.382	20.463	19.915
D16T1	24.072	24.384	24.228	B48L	18.091	18.271	18.181
D16T2	22.960	21.637	22.289	C48L	18.163	18.266	18.215
D16T3	23.285	22.512	22.895	D48L	18.804	18.789	18.797
D17T1	23.766	24.520	24.140	E48L	18.282	18.567	18.424
D17T2	22.375	22.787	22.580	F48L	17.750	18.521	18.131
D17T3	23.783	22.870	23.322	A72C	18.061	19.006	18.528
				B72C	18.008	18.264	18.136
				C72C	18.105	18.882	18.489
				D72C	18.423	18.748	18.585
				E72C	17.967	18.536	18.249
				F72C	17.395	18.292	17.838
				A72P	19.072	20.041	19.551
				B72P	18.711	18.600	18.656
				C72P	19.489	19.496	19.492
				D72P	17.925	17.885	17.905
				E72P	19.190	19.105	19.147
				F72P	18.660	18.995	18.827
				A72L	18.627	19.268	18.945
				B72L	18.562	18.518	18.540
				C72L	18.732	18.999	18.865
				D72L	18.779	18.572	18.676
				E72L	18.767	18.786	18.776
				F72L	17.813	18.189	18.000

<sup>1</sup> Ct values determined by Applied Biosystems 7500 Fast software; representing average of technical triplicate reactions per sample. Experimental transcript expression was normalized against the geometric mean of the two normalizers.

<sup>2</sup> Early embryonic / larval samples: D00-D17 represents 0 days post-fertilization to 17 days post-fertilization. T=tank (replicate) number.

<sup>3</sup> Macrophage samples: A-F represents individual fish. 12, 24, 48, 72 = time (h) post stimulation. C = control (PBS); P = poly(I:C); L = LPS.