Interferon-stimulated genes in zebrafish and human define an ancient arsenal of antiviral immunity

Jean-Pierre Levraud[1*], Luc Jouneau[2], Valérie Briolat[1], Valerio Laghi[1] and Pierre Boudinot[2*]

[1]Institut Pasteur, Macrophages et Développement de l'Immunité, CNRS UMR3738, Paris, France

[2]INRA, Virologie et Immunologie Moléculaires, Université Paris-Saclay, 78352 Jouy-en-Josas, France

[*] Corresponding authors :

Jean-Pierre Levraud

Phone : (+33) 1 40 61 38 42

Email : jean-pierre.levraud@pasteur.fr

Pierre Boudinot

Phone : (+33) 1 34 65 25 85

Email : Pierre.Boudinot@inra.fr

Keywords

Interferon, interferon stimulated gene, antiviral immunity, evolution, comparative transcriptomics, zebrafish

Abstract

The evolution of the interferon (IFN) system, the major innate antiviral mechanism of vertebrates, remains poorly understood. This system is believed to have appeared 500My ago, based on the detection of type IFN genes in cartilaginous fish genomes. However, the IFN system integrates many other components, most of which are encoded by IFN-stimulated genes (ISGs). To shed light on the evolution of this system, we have used deep RNA sequencing to generate a comprehensive list of ISGs of zebrafish, taking advantage of the high quality genome annotation in this species. We analyzed larvae after inoculation of recombinant zebrafish type I IFN, or infection with chikungunya virus, which induces a potent IFN response. We identified more than 400 zebrafish ISGs, defined as being either directly induced by IFN or induced by the virus in an IFN receptor-dependent manner. Their human orthologues were highly enriched in ISGs, particularly for highly-inducible genes. We identified 72 orthology groups that included ISGs in both zebrafish and human, revealing a core ancestral ISG repertoire, which includes most of the known signaling components of the IFN system. Many downstream effectors were also already present 450 My ago in the common ancestor of tetrapods and bony fish, and diversified as multi-gene families independently in the two lineages. A large proportion of the ISG repertoire is lineage-specific; around 40% of protein-coding zebrafish ISGs had no human orthologue. We identified 14 fish-specific gene families containing multiple ISGs, including finTRIMs. This work illuminates the evolution of the IFN system and provides a rich resource to explore new antiviral mechanisms.

46 **Introduction**

47

48 All living organisms are targeted by viruses, and evolution has given rise to various antiviral
49 strategies. Vertebrates possess many unique immune features, including their principal
50 innate antiviral system based on signaling by type I interferons (IFNs). Type I IFNs induce the
51 expression of hundreds of proteins encoded by ISGs (IFN-stimulated genes), making cells
52 refractory to viral infection (Schoggins and Rice, 2011). The origin of the IFN system, which
53 seems to have replaced the RNA interference antiviral system still used by plants and most
54 invertebrates (Backes et al., 2014; Guo et al., 2019), is shrouded in mystery. Cartilaginous
55 fish are the most basal clade with genomes containing recognizable type I IFN genes
56 (Secombes and Zou, 2017), suggesting that this system appeared about 500My ago with
57 jawed vertebrates. The IFN system integrates many components, most of which are encoded
58 by ISGs, which can be traced back in genomes from distant clades. However, finding the
59 orthologue(s) of a human ISG in another taxon does not imply that this gene is part of its IFN
60 system. To understand the evolution of antiviral immunity, it is therefore desirable to
61 establish how the repertoire of ISGs changed from early to modern vertebrates. This can be
62 inferred by comparing the ISGs of current living representatives of distant vertebrate taxa.
63 Bony fishes (hereafter simply called "fish") diverged from the tetrapod lineage about 450My
64 ago, and, since viral infections are a major problem in aquaculture, their IFN system has
65 been the subject of many studies, as reviewed in (Boudinot et al., 2016; Langevin et al.,
66 2013; Robertsen, 2018; Zou and Secombes, 2011). Teleost fish possess several subgroups of
67 type I IFNs (but no type III genes), with strong variation in gene numbers among fish taxa
68 (Boudinot et al., 2016). The zebrafish possess four type I IFN genes, named *ifnphi1-4*; only
69 *ifnphi1* and *ifnphi3* are active at the larval stage (Aggad et al., 2009). Their receptors have

70 been identified (Aggad et al., 2009). Even before fish IFNs were known, the first fish ISGs
71 were identified by homology cloning from cells stimulated by poly-I:C (Staeheli et al., 1989)
72 or by differential transcript analysis of cells infected by viruses (Boudinot et al., 1999;
73 O'Farrell et al., 2002). Because many virus-induced genes (*vig*) were homologous to well-
74 known mammalian ISGs, they were hypothesized to be IFN-inducible, which was often
75 confirmed by later studies, as in the case of *vig-1*, the *rsad2/viperin* orthologue (Boudinot et
76 al., 1999; Levraud et al., 2007). Similarly, upon cloning of fish IFNs, induction of *Mx* (Staeheli
77 et al., 1989) was used as a readout for their activity (Altmann et al., 2003; Lutfalla et al.,
78 2003; Robertsen et al., 2003), confirming it was an ISG. The list of fish homologues of known
79 ISGs rapidly grew with the release of new fish genomes and EST collections, allowing the
80 development of micro-arrays to study fish response to virus or recombinant type I IFNs
81 (Martin et al., 2007; Purcell et al., 2006). Candidate gene approaches were also developed,
82 testing orthologues of known mammalian ISGs in qRT-PCR assays in multiple fish infection
83 models (Levraud et al., 2007; López-Muñoz et al., 2009; Rothenburg et al., 2008). In parallel,
84 approaches without a priori identified fish ISGs that had no orthologue in mammals,
85 although they belonged to gene families involved in antiviral immunity. A large set of
86 tripartite-motif protein-encoding genes, called *fintrims* (*ftr*), distantly related to *trim25* was
87 identified in rainbow trout cells as a induced by virus infection (O'Farrell et al., 2002) and
88 later shown to form multigene families in teleosts, particularly extensive in zebrafish (van
89 der Aa et al., 2009). Similarly, a family of IFN induced ADP-ribosyltransferases named *Gig2*
90 was identified in crucian carp cells treated with UV-inactivated GCHV (grass carp
91 hemorrhage virus) (Zhang and Gui, 2004; Zhang et al., 2013). Some ISG were restricted to
92 particular fish groups such as the non-coding RNA *vig2* that is found only in salmonids
93 (Boudinot et al., 2001).

94

95 We previously established a list of zebrafish candidate ISGs using microarray analysis (Briolat
96 et al., 2014). For this, we compared the response to a poor IFN inducer, IHNV (infectious
97 hematopoietic necrosis virus) (Ludwig et al., 2011) and a strong IFN inducer, CHIKV
98 (chikungunya virus) (Palha et al., 2013). However, the array did not include the full
99 complement of zebrafish genes, and the study identified virus-induced genes which were
100 not necessarily ISGs. Here, to directly identify ISGs, we analyze the transcriptional response
101 of zebrafish larvae injected with recombinant type I IFN. We rely on deep RNA sequencing,
102 which is intrinsically quasi-exhaustive. Our approach is therefore limited mainly by the
103 quality of genome assembly and annotation, which is excellent for the zebrafish (Howe et
104 al., 2013). We complemented this analysis with a study of the response to CHIKV and its
105 dependence to expression of the zebrafish IFN receptor chains *crfb1* and *crfb2* (Aggad et al.,
106 2009). We thus established a comprehensive list of ISGs of zebrafish larvae, and performed a
107 detailed comparison with the human ISG repertoire. Our comparative analysis was
108 facilitated by a compilation of human ISGs made to perform a systematic screen (Schoggins
109 et al., 2011), and by the specialized database Interferome (Rusinova et al., 2013). We
110 identify about 70 orthology groups that include ISG in both species and thus approximate
111 the ISG repertoire of the common ancestor of all Osteichthyes. As ISGs typically evolve fast,
112 with frequent duplications and gene loss, we also identify many families of fish-specific ISGs,
113 which represent a rich resource for seeking new antiviral mechanisms. Our study provides a
114 broad overview of the evolutionary patterns of genes belonging to the type I IFN pathway,
115 and identifies gene modules induced by a viral infection independently of IFN.

116

117

118

119 ## Materials and methods

120

121 *Zebrafish husbandry*

122 Wild-type AB zebrafish, initially obtained from the Zebrafish International Resource Center
123 (Eugene, OR, USA), were raised in the Institut Pasteur facility. Animal experiments were
124 performed according to European Union guidelines for handling of laboratory animals
125 (http://ec.europa.eu/environment/chemicals/lab_animals/home_en.htm) and were
126 approved by the Institut Pasteur Animal Care and Use Committee. Eggs were obtained by
127 marble-induced spawning, cleaned by treatment with 0.003% bleach for 5 minutes, and then
128 kept in Petri dishes containing Volvic source water at 28°C. All timings in the text refer to the
129 developmental stage at the reference temperature of 28.5°C. At 3 dpf (days post
130 fertilization), shortly before injections, larvae that had not hatched spontaneously were
131 manually dechorionated. Larvae were anesthetized with 200 µg/ml tricaine (A5040, Sigma-
132 Aldrich) during the injection procedure.

133

134 *Interferon and virus inoculation*

135 Recombinant zebrafish IFNφ1 (Aggad et al., 2009), kindly provided by Rune Hartmann
136 (University of Aarhus, Denmark), was inoculated by intravenous (IV) injection in the caudal
137 cardinal vein. One nanoliter of 1mg/ml IFNφ1, or as a control, bovine serum albumin (BSA,
138 New England Biolabs) in PBS1x/10% glycerol was injected. CHIKV infections were performed
139 as described (Briolat et al., 2014; Palha et al., 2013). Briefly, ~200PFU of CHIKV115 was
140 injected IV in a volume of 1 nl at 3dpf.

141

142 *Interferon receptor knock-down*

143 Morpholino antisense oligonucleotides (Gene Tools) were injected at the one to two cells
144 stage as described (Levraud et al., 2008). Two ng of *crfb1* splice morpholino
145 (CGCCAAGATCATACCTGTAAAGTAA) was injected together with 2ng of *crfb2* splice
146 morpholino (CTATGAATCCTCACCTAGGGTAAAC), knocking down all type I IFN receptors
147 (Aggad et al., 2009). Control morphants were injected with 4 ng of control morpholino
148 (GAAAGCATGGCATCTGGATCATCGA) with no known target.

149

150 *RNA extraction*

151 For RNAseq analysis, total RNA was extracted from replicate pools of 10 injected larvae (at 6
152 hours post injection for IFNφ1 treatment, or 24 hours post injection for CHIKV infections),
153 using TRIzol (Invitrogen), following the manufacturer's protocol. The integrity of the RNA
154 was confirmed by laboratory-on-chip analysis using the 2100 Bioanalyzer (Agilent
155 Technologies), using only samples with an RNA integrity number of at least 8.

156

157 *Illumina sequencing*

158 Libraries were built using a Truseq mRNA-Seq Library Preparation Kit (Illumina, USA),
159 according to the manufacturer's recommendations. Quality control was performed on an
160 Agilent Bioanalyzer. Sequencing was performed on a HiSeq 2500 system (Illumina, USA) and
161 produced 65-base single-end reads.

162

163 *Mapping reads and gene expression counts*

164 Sequences were trimmed using cutadapt (v1.8.3). The reads quality was checked with
165 FastQC. Reads were then spliced-aligned to the zebrafish genome (GRCz10, Ensembl release
166 88) using TopHat2 (v2.0.14). The average number of mapped read per sample was
167 16.500.000. Only fragments mapping coherently and unambiguously to genes have been
168 considered for gene counts. Gene counts have been assigned using featureCounts v1.5.2
169 (Liao, 2014).
170
171 *Identification of differentially expressed genes (DEG)*
172 Differentially expressed nuclear genes between larvae treated with IFNφ1 and
173 controls, between larvae infected by CHIKV and controls, or between crfb1+2 and control
174 morphants all infected by CHIKV, were identified. DEG were identified using DESeq 1.18.0
175 (BioConductor) (Love et al., 2014) and R: 3–1-2 (R core team, 2017). Briefly, raw counts of
176 genes were subjected to a minimal pre-filtering step: genes for which the count sum, per
177 group of samples, was equal or higher than 10, in at least one group, were kept. Raw counts
178 were normalized for library size and normalized data were fitted using a negative binomial
179 general linear model. Data were adjusted for multiple testing using the Benjamini-Hochberg
180 procedure (adjusted p value). Genes with an adjusted *p* value less than 0.01 and an absolute
181 Fold Change (FC) > 2 or FC < 0.5 were considered as DEGs.
182 Sequence data were registered in the BioProject ncbi database with the SRA accession
183 number: PRJNA531581.
184
185 *Identification of human orthologues and ISGs*
186 Orthology analysis was primarily based on data from the Ensembl database
187 (www.ensembl.org), the zfin database (zfin.org), and the literature, notably our previous

188 analysis of zebrafish orthologues of human ISGs (Briolat et al., 2014). Data were
189 systematically curated manually, and conflicts resolved using a combination of literature
190 search, synteny analysis, and sequence homology analysis (two-way protein BLAST). When
191 human genes were present on the list compiled by Schoggins *et al.* (Schoggins et al., 2011),
192 they were labelled as ISGs. If absent from the list, gene names were further queried on the
193 Interferome website (http://www.interferome.org/interferome/search/showSearch.jspx)
194 which compiles the results of many transcriptomic studies on human and mouse samples
195 after IFN stimulation (Rusinova et al., 2013). We postulated that human genes present in
196 Interferome could be considered as ISG when being significantly induced at least 2-fold with
197 a stimulation for no more 12 hours by type I IFN in at least 4 datasets.
198
199 *qRTPCR*
200 RNA was extracted from individual larvae using RNeasy® Mini Kit (Qiagen). cDNA was
201 obtained using M-MLV H- reverse-transcriptase (Promega) with a dT17 primer. Quantitative
202 PCR was then performed on an ABI7300 thermocycler (Applied Biosystems) using Takyon™
203 ROX SYBR© 2x MasterMix (Eurogentec) in a final volume of 25 µl. The following pairs of
204 primers were used: ef1a (housekeeping gene used for normalization): 5'-
205 GCTGATCGTTGGAGTCAACA-3' and 5'-ACAGACTTGACCTCAGTGGT-3'; mxa: 5'-
206 GACCGTCTCTGATGTGGTTA-3' and 5'-GCATGCTTTAGACTCTGGCT-3'; ddx58: 5'-
207 ACGCCGGAGAAAGAATTTTTC-3' and 5'-TCGACAGACTCTCGATGTTG-3'
208
209
210

211

212

213

214

215 **Results**

216

217 *RNAseq analysis of IFNφ1-regulated genes*

218 To make an inventory of zebrafish ISGs, we first injected 3dpf larvae with
219 recombinant zebrafish IFNφ1, the first type I IFN to be identified in zebrafish, or BSA as a
220 negative control. Based on preliminary kinetic experiments, we chose 6 hours post injection
221 as the early plateau phase of ISG expression (Additional File 1A). RNA was extracted from
222 multiple pools of ten larvae and subjected to deep sequencing using an Illumina-based
223 platform sequencing. Reads were mapped to zebrafish genome (zv10), and the differential
224 analysis performed using the DESeq package.

225 Choosing as cutoff values adjusted *p* values <5% and fold-change (FC) >2, we
226 identified 360 IFNφ1 up-regulated genes (which are ISGs, by definition) and 75 down-
227 regulated genes (Additional File 2). As expected, genes with high basal expression levels
228 tended to display lower fold-change (Figure 1A). The top IFNφ1-upregulated genes (with FC
229 >100) comprised many well-known ISGs, many previously used as IFN signature genes in
230 zebrafish, including several *mx* genes, *rsad2, cmpk2*, several *ifit* genes, the ubiquitin-like
231 *isg15*, the helicase *dhx58* (aka *lgp2*), the kinase *pkz*, the transcription factor *stat1b*, and the
232 chemokine *ccl19a.2* (Figure 1B). To our surprise, *ddx58* (encoding RIG-I), a well-known and
233 conserved ISG (Biacchesi et al., 2009), was not found in that list. In fact, the gene model is

234 missing on the zebrafish reference genome, with only fragments of the sequence present in
235 the assembly. Therefore, we performed qRT-PCR for *ddx58* and confirmed that it is induced
236 by IFNϕ1 (Additional File 1B).

237 Gene ontology (GO) analysis was performed using David and Gorilla, and showed, as
238 expected, that IFNφ1 upregulated genes were strongly enriched in genes linked to “antiviral
239 response”, “type I interferon pathway” and "antigen processing and presentation" (not
240 shown). Similarly, enrichment analyses identified KEGG pathways for Influenza, Measles and
241 Herpes simplex infection as well as RIG-I signaling and cytosolic DNA sensing pathways. The
242 list of downregulated genes was not found associated with any particularly notable function
243 in these analyses.

244

245 *Human orthologues of zebrafish ISGs are enriched in ISGs*

246 We then searched for the human orthologues of the 360 identified zebrafish ISGs
247 (Additional File 3). All types of orthology relationships between zebrafish and human were
248 observed, from none to "many-to-many". One-to-one orthology was found for 77 genes
249 (Figure 1C). We identified one or several human orthologues for 200 zebrafish ISGs. Even
250 after exclusion of the 23 zebrafish ISGs corresponding to non-coding RNAs or transposons,
251 this proportion (200/337, 59%) is significantly lower than the 71% reported for the whole
252 genome (Howe et al., 2013) (Fischer’s exact test, $p<0.0001$). We then searched which of
253 these 196 human genes were themselves ISGs. We found 61 ISGs present in the list of 446
254 human ISGs compiled by Schoggins et al. (Schoggins et al., 2011); by querying the
255 Interferome database (Rusinova et al., 2013), we identified 11 additional human ISGs
256 (Additional File 3). In total, 97 zebrafish IFNϕ1-inducible genes were orthologous to at least
257 one human ISGs (Figure 1C). In addition, we identified a handful of genes that were not true

258 orthologues, but shared ancestry with a human ISG at the vertebrate level, such as MHC
259 class I genes (see comments on Additional File 3). As expected, human orthologues of
260 zebrafish ISGs were strongly enriched in ISGs: 37% (72 out of 196 genes), compared with 2%
261 in the entire genome based on the Schoggins list of human ISGs (Fischer's exact test, $p<$
262 0.0001). Interestingly, FC values of zebrafish ISGs were higher when they were orthologous
263 to a human ISG than when their orthologues were not ISGs (red vs blue on Figure 1C-D),
264 while FCs of zebrafish ISGs without human orthologues (grey on Figure 1C-D) were
265 intermediate (mean values, 127.7, 10.6 and 46.9, respectively; all groups significantly
266 different from each other, $p<0.001$, Kruskal-Wallis test). Thus, inducibility by type I IFNs is
267 often evolutionary conserved, making it possible to infer an ancestral set of ISGs.
268
269 *Fish-specific ISG genes and families*
270 Consistent with the expected high rate of duplication and divergence of ISGs, a
271 significant proportion of zebrafish IFNφ1-up-regulated protein-coding genes had no
272 identifiable orthologue in the human genome (137/337; 41%). Interestingly, but not unlike
273 genes with human orthologues, many of these genes belonged to multigenic families. We
274 show on Figure 2 the fish-specific gene families that contain several zebrafish ISGs.
275 This included more than 20 *fintrim (ftr)* genes, a family identified first as virus-
276 inducible genes in rainbow trout (O'Farrell et al., 2002), highly diversified in zebrafish, and
277 hypothesized to antagonize retroviruses (van der Aa et al., 2009). Of note, besides finTRIM,
278 there are two other large TRIM gene expansions in zebrafish, each with a single human
279 orthologue of unknown function (Boudinot et al., 2011). The *btr* (bloodthirsty-like TRIMs),
280 related to human TRIM39, include several ISGs. By contrast, no member of the TRIM35-like
281 family was upregulated by IFNφ1.

282 Another family had been described as virus-inducible in fish: *gig2* (GCRV-induced
283 gene 2) genes (Zhang and Gui, 2004). *gig2* genes are distantly related to the PARP family
284 (Zhang et al., 2013), which include several ISGs in humans and zebrafish. The zebrafish *gig2p*
285 and *gig2o* genes are induced by IFNφ1.

286 To our knowledge, the genes in the remaining fish-specific families had not previously
287 been described to be interferon- or virus-inducible. These families are diverse, encoding for
288 proteins expected to be membrane receptors, and presumably secreted, nuclear, or
289 cytosolic proteins (Figure 2).

290 Eight members of the very large NLR family were ISGs. These genes belong to groups
291 1, 2, 3 and 4 as defined in (Howe et al., 2016), and two of them belong to a fish specific
292 subset defined by the presence of a C-terminal B30.2 (or PRY-SPRY) domain (Laing et al.,
293 2008; Stein et al., 2007), which is most similar to the corresponding domain of finTRIM genes
294 (van der Aa et al., 2009). The specific function(s) of zebrafish NLR genes remain poorly
295 understood, but this highly expanded family may be central for inflammatory mechanisms.

296 Additionally, 3 ISGs corresponded to surface proteins with two immunoglobulin (Ig)
297 domains and a transmembrane region but not ITAM or ITIM (immunoreceptor tyrosine-
298 based activating/inhibiting motif). These genes belong to a very large family with 140
299 members, which we propose to name f2Ig (fish genes with two Ig domains).

300

301 *RNAseq analysis of CHIKV-induced genes*

302 Experimental infection of zebrafish larvae with chikungunya virus (CHIKV) induces a
303 strong type I IFN response (Palha et al., 2013). Our previous microarray-base analysis
304 indicated that the response to CHIKV was dominated by ISGs (Briolat et al., 2014). However,
305 to allow comparison with another virus with slower IFN induction kinetics, this analysis had

306 been performed at 48 hours post infection (hpi), while the peak of the IFN and ISG response,
307 as determined by qRT-PCR, is at 24 hpi (Briolat et al., 2014). Therefore, we re-analyzed here
308 the transcriptome of CHIKV-infected larvae at 24 hpi using deep RNA sequencing. Choosing
309 the same cutoff values as for the IFNφ1 analysis, we identified 466 CHIKV up-regulated genes
310 and 26 down-regulated genes (Additional File 4). Hundreds of new CHIKV-inducible genes
311 were identified, either because they were absent from the microarray, or because their
312 induction were below the cutoff of the first analysis. Among the genes significantly
313 upregulated in the microarray study, all those with a human ISG orthologue were also
314 upregulated in this new dataset, and, as expected, typically much more (Additional File 5).

315 About half of genes induced by IFNφ1 were also induced by the viral infection
316 (181/360 Figure 3A; Additional File 4, yellow), including almost all (84 out of 97) genes
317 orthologous to a human ISG, such as *mxa, b* and *e, stat1a* and *b; stat2; rasd2, isg15*, etc.
318 There was an excellent correlation of the FC values for genes induced by both IFNϕ1 and
319 CHIKV (Figure 3B). However, almost two thirds of the genes induced by CHIKV infection were
320 not significantly modulated by IFNφ1 (285/466; Figure 3A). We then asked whether this
321 CHIKV-specific response could correspond essentially to genes for which there was weak
322 induction by recombinant IFNφ1, below our arbitrary cutoff. We therefore extracted from
323 this list genes induced by IFNφ1 with FC>1.5 and with an adjusted *p* value<20%, and we
324 found 66 genes matching these conditions (Additional File 4, green): 21 genes without
325 annotation and 45 annotated genes, many of which were notoriously linked to the type I IFN
326 system. These genes notably comprised *crfb1*, encoding a type I IFN receptor subunit, and
327 two other cytokine receptors *il10Ra* and *il13Ra*; four chemokines (*ccl34, cxcl11.6, cxc18b* and
328 *cxc20*); ten additional *fintrim* and 3 other members of the *gig2* family; and two *irf*
329 transcription factors (2 and 10). It also includes the metalloreductase *steap3*, whose

330 mammalian orthologue is not an ISG, but regulates type I IFN response, CXCL10 induction
331 and iron homeostasis in mouse macrophages (Zhang et al., 2012).
332 Besides this intermediate gene set, a conservative list of 219 genes seems to be up-
333 regulated only by the virus, independently of IFNφ1 (Additional File 4, blue; Figure 3C). This
334 list contains 105 genes without annotation, but also several functional modules providing
335 interesting insights on the virus-host interactions. Functional analysis using DAVID identified
336 6 enriched KEGG pathways, namely Cytokine-cytokine receptor interaction, Cytosolic DNA-
337 sensing, Toll-like receptor signaling, RIG-I-like receptor signaling, Proteasome, Herpes
338 simplex infection.
339 Importantly, type I IFNs were induced by the infection. Consistent with our previous report
340 with another virus (Aggad et al., 2009), *ifnphi1* and *ifnphi3* were clearly dominant at this
341 larval stage, with 58±3 and 47±11 reads respectively, compared with 9±4 reads for *ifnphi2*,
342 and none detected for *ifnphi4*. Two pro-inflammatory cytokines *il1b* and *tnfb* were also
343 upregulated. Among typical sensors, *tlr3, mb21d1* (encoding cGAS) and its downstream
344 adaptor *tmem173* (encoding STING), and several kinases of the IFN signaling pathways
345 (*ripk1, tbk1*) were present. Seven proteasome subunits are induced by the virus, suggesting
346 activation of protein degradation and Ag presentation pathways. The complement pathway
347 also stands out as an important module upregulated by CHIKV infection: twelve complement
348 genes (*c1, c2*, several *c3, c7; c9; cfB; cfhl-1,-3* and *-5*) were induced by CHIKV, suggesting that
349 it is an important defense triggered in a Type I IFN independent manner. Additionally, this
350 response comprises 3 metalloaminopeptidases (*anpepb, erap1b* and *2*); the myeloid markers
351 *ncf1, mpx* and *marco*; 2 guanylate binding proteins (*gbp1* and *2*) that have well known
352 orthologues in human; the transcription factors *atf3* and *irf1b*, and with a high level of
353 expression, the enzyme *rnasel3* (an orthologue of human RNASE4, not of RNASEL, an ISG

354 with no fish counterpart). Nine *fintrim* and 3 *btr* can also be noted, underscoring the
355 importance of these TRIM with PRY/SPRY domains in virus host interactions altogether.
356 Thus, CHIKV induces a typical IFN-stimulated response of high magnitude, but also a broader
357 and less overt inflammatory response.
358
359 *IFN receptor dependence of the response to CHIKV*
360 To test the IFN-dependence of the response to CHIKV, we used morpholinos to knock down
361 in zebrafish larvae *crfb1* and *crfb2* which encode specific chains of the two types of type I IFN
362 receptors of zebrafish (Aggad et al., 2009). We previously showed that such IFNR morphant
363 larvae are hypersusceptible to CHIKV infection, dying 2 to 3 days after virus injection (Palha
364 et al., 2013). We analyzed by deep RNAseq their transcriptional response to CHIKV at 24
365 hours post-inoculation, and compared it to that of control morphant larvae. Choosing as
366 cutoff values adjusted *p* values <5% and a ratio between IFN-R morphants and controls >2,
367 we identified 187 genes for which induction was dampened by IFNR knockdown, and 10
368 genes that were upregulated in morphants (Additional File 6). Among CHIKV-induced genes
369 (Additional File 4), 181 were IFNR-dependent, representing a significant fraction (181/466;
370 39%) (Figure 3A). Predictably, the list of genes upregulated by CHIKV in a IFNR-dependent
371 manner largely - but not fully - overlapped with the gene set induced by recombinant IFNφ1
372 (129/181; 71%, see Figure 3A). This approach led us to classify 52 new zebrafish genes as
373 ISGs, being induced by CHIKV in an IFNR-dependent manner, even if they were not
374 significantly induced by recombinant IFNφ1. As previously, we searched the human
375 orthologues of these additional ISGs (Additional File 3, bottom), identifying a few more
376 human ISGs in this list, such as cGAS, NLRC5 or IFI35.

377 Together, our results provide a near-exhaustive list of zebrafish ISGs at the larval stage,
378 identified by two independent approaches, and a useful reference for future studies.
379
380 *Ancestral ISGs*
381 Assuming that the common ancestor of genes that are IFN-inducible in both human and
382 zebrafish was itself an ISG in their last common ancestor ~450 My ago, we can define a list of
383 ancestral ISGs. We identified 66 orthology groups that included an ISG on both the human
384 and the zebrafish sides (Table 1, Additional File 7). A few more ancestral ISGs were also
385 defined by pairs of ISGs with orthology relationships at the early vertebrate level. In total,
386 our list includes 72 ancestral genes (Table 1, Figure 4A).
387 Based on our orthology analysis, we propose new, more explicit names for many of the
388 zebrafish ISGs with known human orthologues (in red on Table 1 and Additional File 7). This
389 ancestral ISG core includes most ISGs with known functions. The IFN system of 450My ago
390 seems fairly similar to the present one, particularly in its signaling components (Figure 4;
391 Additional File 8). Many ancestral genes have been duplicated independently in one or both
392 lineages (Additional File 7), in addition to multiple ISGs apparently gained by either group
393 (Figure 5).
394
395 *IFN-downregulated genes*
396 Two of the most strongly IFNφ1-downregulated genes were orthologous to human genes
397 downregulated by type I IFNs, according to the Interferome database: *plin1* (perilipin 1) and
398 *acox1* (palmytoil acyl-CoA oxidase 1). This suggests that downregulation of fatty acid
399 oxidation pathway is an ancient feature of the IFN system.

400 Many IFNφ1-downregulated genes were orthologous to a human gene in a 1-to-2 manner
401 (Additional File 3), with the two zebrafish paralogues having arisen during the teleost
402 specific whole genome duplication (ohnologues). Systematically, only one of the two
403 paralogues was downregulated.
404 Remarkably, no gene was downregulated by both IFNφ1 injection and CHIKV infection
405 (Figure 3b, Additional File 4).
406
407

408 **Discussion**

409

410 The zebrafish has become an important model to study host-pathogen interactions,
411 particularly at its early life stages which are the most prone to live imaging and genetically
412 tractable. Although its antiviral interferon genes and receptors are now well identified,
413 knowledge of IFN-induced genes, or ISGs, was only partial. In this work, we used deep
414 sequencing to characterize the transcriptomic response of the 3dpf zebrafish larva to
415 recombinant IFNφ1, the first type I IFN identified in zebrafish and the most highly inducible
416 one. We analyzed in parallel the response to an alphavirus inducing a strong type I IFN
417 induction, ant the impact of IFN receptor knock-down on this response. From these different
418 datasets, we established a comprehensive list of zebrafish ISGs. This list was compared to
419 the human ISG repertoire, and a phylogenetic analysis was performed to approach the
420 ancestral ISG repertoire of early vertebrates.
421

422 *1. New insights and limitations of the work*

423 A number of studies have identified genes induced by IFN or viral infections in fish (reviewed
424 in (Langevin et al., 2013)). However, very few global descriptions after treatment with
425 recombinant type I IFN have been reported, using micro-arrays (Martin et al., 2007). Micro-
426 array analyses are limited by probe choice, and are typically biased towards genes with
427 known human homologues. RNAseq, by contrast, is mainly limited by the genome
428 annotation quality and by the analysis method, and can be reanalyzed; this approach is thus
429 more complete. Since the early zebrafish larva constitutes a reference model for
430 investigating innate immune response, drug screening as well as for modelling diseases, we
431 undertook a comprehensive description of the repertoire of ISG up-regulated at this
432 developmental stage. Importantly, we previously reported a clear transcriptional response
433 of zebrafish embryos to IFNφ1 as early as 24 hpf (Levraud et al. 2007); the responsiveness to
434 type I IFNs is thus already well established at 3 dpf. We are aware that cells present in adult
435 but not yet in larvae, notably those of the adaptive immune system such as lymphocytes and
436 dendritic cells may express additional ISGs, which should be assessed in further work.
437 There are two groups of type I IFNs in teleost fish (Zou et al., 2007) with two different
438 receptors (Aggad et al., 2009). This study only addresses the ISG repertoire induced by IFNφ1
439 (a group 1 IFN) and it is possible that group 2 IFNs (IFNφ2 and IFNφ3) induce a different ISG
440 subset. Determining this will require more studies; however, since CHIKV induces both
441 IFNφ1 and IFNφ3 while *crfb1&2* morpholinos target receptors for both type I IFN groups,
442 IFNφ3-only induced ISGs should therefore be found among CHIKV-induced, IFNR-dependent,
443 but non IFNφ1-induced genes. Such genes (listed on Table S2, bottom) constituted about
444 30% of genes for which induction by CHIKV was impacted in morphants – and only about
445 15% if one also excludes genes for which induction by IFNφ1 is almost significant (Figure 3A
446 and C).

447

448 *2. Comparative and phylogenetic analysis of zebrafish ISGs*

449 Our comparative and phylogenetic approach led to a tentative reconstruction of the
450 ISG repertoire of the last common ancestor of teleosts and tetrapods (LCATT) that lived
451 ~450My ago and probably resembled the fossil osteichtyan *Ligulalepis* (Clement et al., 2018).
452 To do so, we looked for human (co-)ortholog(s) of all zebrafish ISGs identified in our analysis.
453 Based on available data compilations (Rusinova et al., 2013; Schoggins et al., 2011), we then
454 determined which one(s) of these human orthologs were themselves induced by type I IFN.
455 In such cases, we considered that they most likely originated from an "ancestral" ISG,
456 present in the LCATT. It is generally believed that the type I IFN system emerged during the
457 early evolution of jawed vertebrates, since Chondrichtyans (rays, sharks and chimeras) but
458 not Agnathans (lampreys and hagfish) possess typical type I IFN genes (Secombes and Zou,
459 2017; Venkatesh et al., 2014). Hence, it is important to note that the IFN system had already
460 evolved, expanded and standardized for more than 50My before our last common ancestor
461 with zebrafish.

462 Approximately half of what we defined as ancestral ISGs are represented by 1-to-1 co-
463 orthologs in zebrafish and human (Additional File 7, top rows) – a situation of practical
464 interest, as the likelihood of conservation of gene function is highest in this case. These are
465 either isolated genes (e.g. *RSAD2*, *ISG15* or *cGAS*), or members of "old" families already
466 stabilized in the LCATT (e.g., *IRF7* and *IRF9*) (Briolat et al 2014). The situation is relatively
467 similar for a few ancestral genes such as *STAT1* or *SOCS1*, with one human orthologue and
468 two zebrafish co-orthologues that arose during the teleost-specific whole genome
469 duplication and were retained. In contrast, many other “young” families have clearly been
470 subjected to further duplication during later evolution of fish or tetrapods, leading to

471 orthology groups containing multiple ISG both in zebrafish and human, the most spectacular
472 examples being the *ISG12*, *IFIT* and *IFI44* families.
473 In addition to the ancestral genes with true zebrafish and human orthologs, we added to this
474 list a few genes with a more complex history, with a human and zebrafish ISGs that shared
475 an ancestor at the basal vertebrate level (Table 1 and Additional File 7, bottom). These
476 ancestral genes must have been duplicated in the LCATT genome; the teleost and tetrapod
477 lineages then retained distinct paralogues. This comprises some genes whose evolutionary
478 history is extremely difficult to trace due to multiple copies and extensive polymorphism,
479 such as MHC class I genes. Here, only *mhc1zba* was found to be a zebrafish ISG, but this does
480 not necessarily imply that other zebrafish MHC class I genes are not ISGs, as they may have
481 been missed due to mapping issues; the strain we used (AB) is not the same as the one of
482 the reference genome (Tü), and strain-specific divergences are considerable between strains
483 for MHC class I, with deep evolutionary roots (McConnell et al., 2016). Importantly, we did
484 not define ancestral ISGs for zebrafish-human ISG pairs that appeared to be related at first
485 glance, but, upon further analysis, were too distant; for example, zebrafish *vamp5* and
486 human *VAMP8* are both ISGs, but share their last common ancestor at the Opistokontha
487 level, before the split of fungi and animals, very long before the emergence of IFNs.
488 Nevertheless, the type I IFN system also includes very old genes that were already present in
489 basal metazoans. The RNAseL/OAS module is a good example of such cases, being found
490 across metazoans from mammals to sponges (Päri et al., 2007) - but lost in the fish branch.
491 Another striking example is the cGAS-STING module recently identified in cnidaria
492 (Kranzusch et al., 2015). The implication of these genes in the antiviral immunity of basal
493 branches of animals is unknown but certainly worth investigating. The main models for
494 invertebrate immunity are flies and mosquitoes but they largely rely on RNAi mechanisms to

495 contain viruses (Guo et al., 2019). Central signaling modules of the vertebrate IFN system
496 such as TLR/NFκB and JAK/STAT, are also present in insects and in more distant metazoans,
497 but with different outcomes (Liongue et al., 2016).
498 Additionally, a few important genes do not meet our criteria for "ancestral" ISG because
499 they are not typically inducible either in zebrafish or in human (Figure 5). For example, *irf3* is
500 an ISG in fish but not in human, while it is the reverse for *JAK2*. Hence, our list of ancestral
501 ISG is likely not complete, but it provides a core repertoire pointing to most fundamental
502 factors of the vertebrate innate antiviral arsenal.
503 A relatively large number of ISGs have no orthologue in the other lineage, such as human
504 APOBEC3, RNASEL, OAS, AIM2 (Figure 5). Similarly, many fish-specific ISG likely have been
505 co-opted by the IFN pathway during fish evolution. In this case, they do not have clear
506 orthologs in human and other tetrapods (as for finTRIMs and nlr-B30.2), or their ortholog(s)
507 have no link with the type I IFN system. The finTRIM family contains the largest number of
508 zebrafish ISGs of any family, ancestral or not. Interestingly, ISGs are found only among the
509 recently diversified, species-specific finTRIMs; the most basal members (*ftr82-84*), well-
510 conserved among fish, were not found here to be induced by IFNφ1 or by CHIKV, consistent
511 with previous studies (Langevin et al., 2019). Nevertheless, *ftr83* appears to mediate
512 protection especially in the gill region by stimulating local *ifnphi1* expression (Langevin et al.,
513 2017). The diversity and evolution under positive selection of the IFN-inducible finTRIMs
514 evoke viral recognition (van der Aa et al., 2009), yet their functions remain unclear.
515 The co-optation of new genes in the ISG repertoire may be operated quickly and in a group-
516 specific manner, by introduction of sequence motifs in the regulatory sequences, for
517 example via retroviral insertion (Arnold and Kunte, 2017; Chuong et al., 2016). However, we
518 cannot exclude that these branch-specific ISGs are in fact ancestral, but lost in one of the

519 two lineages; this is the case for *Gig2* genes, which are present in the Coelacanth genome as
520 well as in fish, and thus were lost in tetrapods. Thus, our repertoire of ancestral ISGs is
521 underestimated because we cannot include the lineage-specific losses.
522 Ancestral ISGs encode very diverse proteins in localization and function (Figure 4). We
523 provide an extended discussion of their classification in Additional File 8.
524
525 *4. Characterization of the IFN independent response to CHIKV infection*
526 Knowing the repertoire of ISG also offers the possibility to identify genes that are induced by
527 viral infection independently of the type I IFN pathway. While a subset of ISGs can be
528 induced via IFN-dependent and -independent pathways in human and fish - for example
529 *rsad2* (Boudinot et al., 1999). Thus, IRF3 dependent, type I IFN independent induction of
530 many ISG by particular viruses has been described (Ashley et al., 2019).
531 However, about half of the genes upregulated by CHIKV were not induced by IFNϕ1
532 injection, and most were not affected by IFNR knockdown. Notably, three gene sets stand
533 out in this list: (1) components of the complement cascade, that are known to play a role in
534 antiviral defense; (2) cytokines including some CC and CXC chemokines as well as the type I
535 IFN themselves, which do not appear to be strongly auto/cross inducible and (3) many *btr*
536 and *ftr* TRIM E3 ligases as well as multiple proteasome components. Interestingly too, *irf1b*,
537 the zebrafish orthologue of IRF1 (a human ISG), is CHIKV-inducible, but not in a IFNR-
538 dependent manner – consistent with previous work (Briolat et al., 2014) - and was not
539 induced by IFNϕ1. Many other genes of unknown function also share the same induction
540 pattern, and would certainly be worth investigating. A strong redundancy of antiviral
541 pathways has certainly been selected during evolution, since viruses have developed
542 multiple strategies of immune subversion.

543

544

545 *Conclusions*

546 Antiviral genes are well known to evolve very fast, as postulated by the Red Queen
547 hypothesis, under strong pressure from pathogens. This is indeed illustrated by the large
548 number of ISGs that are either fish- or mammal-specific. Nevertheless, our data define a
549 surprisingly stable set of core ISGs, that were apparently co-opted into the new IFN system
550 of early vertebrates about 500My ago, and have been maintained for the last 450My both in
551 fish and tetrapods. The full list of zebrafish ISG provides a powerful reference to characterize
552 the subtle interactions between viruses and the host response, including redundancy of
553 immune pathways and viral subversion mechanisms. It also constitutes a valuable resource
554 for the study of autoinflammatory disease using the emerging zebrafish model.

555

556 Competing interests

557 The authors declare that they have no competing interests.

558

559 Author's contribution

560 JPL and PB designed the project, acquired funding, generated figures and wrote the paper.
561 JPL performed microinjections. LJ made the bioinformatics analysis. VB managed fish and
562 extracted RNA. VB and VL performed qRT-PCR and drafted figures.

563

564 Acknowledgements

565 We thank Jean-Yves Coppee and Caroline Proux from the transcriptomics platform of Institut
566 Pasteur, for generation and sequencing of RNA libraries. We are indebted to Rune Hartmann
567 (Aarhus, Denmark) for recombinant zebrafish interferon and stimulating discussion. We
568 thank Emma Colucci and Pedro Hernandez-Cerda for critical reading of the manuscript.

569

570 ## Declaration

571 Sequence data is publicly available in the BioProject ncbi database with the SRA accession
572 number: PRJNA531581

573

574 ## Funding

575 ANR grant Fish-RNAVAX (ANR-16-CE20-0002-02 to PB, ANR-16-CE20-0002-03 to JPL)
576 European Union's Horizon 2020 research and innovation programme, Marie Sklodowska-
577 Curie grant agreement No 721537 "ImageInLife" to JPL and VL

578

579

580 ## References

581

582 van der Aa, L.M., Levraud, J.-P., Yahmi, M., Lauret, E., Briolat, V., Herbomel, P., Benmansour,
583 A., and Boudinot, P. (2009). A large new subset of TRIM genes highly diversified by
584 duplication and positive selection in teleost fish. BMC Biol. *7*, 7.
585 Aggad, D., Mazel, M., Boudinot, P., Mogensen, K.E., Hamming, O.J., Hartmann, R., Kotenko,
586 S., Herbomel, P., Lutfalla, G., and Levraud, J.-P. (2009). The two groups of zebrafish virus-
587 induced interferons signal via distinct receptors with specific and shared chains. J. Immunol.

588 *183*, 3924–3931.

589 Altmann, S.M., Mellon, M.T., Distel, D.L., and Kim, C.H. (2003). Molecular and Functional

590 Analysis of an Interferon Gene from the Zebrafish, Danio rerio. J. Virol. *77*, 1992–2002.

591 Arnold, M.L., and Kunte, K. (2017). Adaptive Genetic Exchange: A Tangled History of

592 Admixture and Evolutionary Innovation. Trends Ecol. Evol. *32*, 601–611.

593 Ashley, C.L., Abendroth, A., McSharry, B.P., and Slobedman, B. (2019). Interferon-

594 Independent Upregulation of Interferon-Stimulated Genes during Human Cytomegalovirus

595 Infection is Dependent on IRF3 Expression. Viruses *11*, E246.

596 Backes, S., Langlois, R.A., Schmid, S., Varble, A., Shim, J. V., Sachs, D., and TenOever, B.R.

597 (2014). The Mammalian Response to Virus Infection Is Independent of Small RNA Silencing.

598 Cell Rep. *8*, 114–125.

599 Biacchesi, S., LeBerre, M., Lamoureux, A., Louise, Y., Lauret, E., Boudinot, P., and Brémont,

600 M. (2009). Mitochondrial antiviral signaling protein plays a major role in induction of the fish

601 innate immune response against RNA and DNA viruses. J. Virol. *83*, 7815–7827.

602 Boudinot, P., Massin, P., Blanco, M., Riffault, S., and Benmansour, A. (1999). vig-1, a new fish

603 gene induced by the rhabdovirus glycoprotein, has a virus-induced homologue in humans

604 and shares conserved motifs with the MoaA family. J. Virol. *73*, 1846–1852.

605 Boudinot, P., Salhi, S., Blanco, M., and Benmansour, A. (2001). Viral haemorrhagic

606 septicaemia virus induces vig -2, a new interferon-responsive gene in rainbow trout. Fish

607 Shellfish Immunol. *11*, 383–397.

608 Boudinot, P., van der Aa, L.M., Jouneau, L., Pasquier, L., Pontarotti, P., Briolat, V.,

609 Benmansour, A., and Levraud, J.P. (2011). Origin and evolution of TRIM proteins: New

610 insights from the complete TRIM repertoire of zebrafish and pufferfish. PLoS One *6*, e22022.

611 Boudinot, P., Langevin, C., Secombes, C.J., and Levraud, J.-P. (2016). The peculiar

612 characteristics of fish type I interferons. Viruses *8*.

613 Briolat, V., Jouneau, L., Carvalho, R., Palha, N., Langevin, C., Herbomel, P., Schwartz, O.,
614 Spaink, H.P., Levraud, J.-P., and Boudinot, P. (2014). Contrasted Innate Responses to Two
615 Viruses in Zebrafish: Insights into the Ancestral Repertoire of Vertebrate IFN-Stimulated
616 Genes. J. Immunol. *192*, 4328–4341.

617 Chuong, E.B., Elde, N.C., and Feschotte, C. (2016). Regulatory evolution of innate immunity
618 through co-option of endogenous retroviruses. Science (80-.). *351*, 1083–1087.

619 Clement, A.M., King, B., Giles, S., Choo, B., Ahlberg, P.E., Young, G.C., and Long, J.A. (2018).
620 Neurocranial anatomy of an enigmatic Early Devonian fish sheds light on early osteichthyan
621 evolution. Elife *7*, e34349.

622 Guo, Z., Li, Y., and Ding, S.W. (2019). Small RNA-based antimicrobial immunity. Nat. Rev.
623 Immunol. *19*, 31–44.

624 Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C., Muffato, M., Collins, J.E.,
625 Humphray, S., McLaren, K., Matthews, L., et al. (2013). The zebrafish reference genome
626 sequence and its relationship to the human genome. Nature *496*, 498–503.

627 Howe, K., Schiffer, P.H., Zielinski, J., Wiehe, T., Laird, G.K., Marioni, J.C., Soylemez, O.,
628 Kondrashov, F., and Leptin, M. (2016). Structure and evolutionary history of a large family of
629 NLR proteins in the zebrafish. Open Biol. *6*, 160009.

630 Kranzusch, P.J., Wilson, S.C., Lee, A.S.Y., Berger, J.M., Doudna, J.A., and Vance, R.E. (2015).
631 Ancient Origin of cGAS-STING Reveals Mechanism of Universal 2',3' cGAMP Signaling. Mol.
632 Cell *59*, 891–903.

633 Laing, K.J., Purcell, M.K., Winton, J.R., and Hansen, J.D. (2008). A genomic view of the NOD-
634 like receptor family in teleost fish: Identification of a novel NLR subfamily in zebrafish. BMC
635 Evol. Biol. *8*, 42.

636 Langevin, C., Aleksejeva, E., Passoni, G., Palha, N., Levraud, J.P., and Boudinot, P. (2013). The
637 antiviral innate immune response in fish: Evolution and conservation of the IFN system. J.
638 Mol. Biol. *425*, 4904–4920.
639 Langevin, C., Aleksejeva, E., Houel, A., Briolat, V., Torhy, C., Lunazzi, A., Levraud, J.-P., and
640 Boudinot, P. (2017). FTR83, a member of the large fish-specific finTRIM family, triggers IFN
641 pathway and counters viral infection. Front. Immunol. *8*, 617.
642 Langevin, C., Levraud, J.P., and Boudinot, P. (2019). Fish antiviral tripartite motif (TRIM)
643 proteins. Fish Shellfish Immunol. *86*, 724–733.
644 Levraud, J.-P., Boudinot, P., Colin, I., Benmansour, A., Peyrieras, N., Herbomel, P., and
645 Lutfalla, G. (2007). Identification of the zebrafish IFN receptor: implications for the origin of
646 the vertebrate IFN system. J. Immunol. *178*, 4385–4394.
647 Levraud, J.-P., Colucci-Guyon, E., Redd, M.J., Lutfalla, G., and Herbomel, P. (2008). In vivo
648 analysis of zebrafish innate immunity. Methods Mol. Biol. *415*, 337–363.
649 Liongue, C., Sertori, R., and Ward, A.C. (2016). Evolution of Cytokine Receptor Signaling. J.
650 Immunol. *197*, 11–18.
651 López-Muñoz, A., Roca, F.J., Meseguer, J., and Mulero, V. (2009). New insights into the
652 evolution of IFNs: zebrafish group II IFNs induce a rapid and transient expression of IFN-
653 dependent genes and display powerful antiviral activities. J. Immunol. *182*, 3440–3449.
654 Ludwig, M., Palha, N., Torhy, C., Briolat, V., Colucci-Guyon, E., Brémont, M., Herbomel, P.,
655 Boudinot, P., and Levraud, J.P. (2011). Whole-body analysis of a viral infection: Vascular
656 endothelium is a primary target of Infectious Hematopoietic Necrosis Virus in zebrafish
657 larvae. PLoS Pathog. *7*, e1001269.
658 Lutfalla, G., Crollius, H.R., Stange-Thomann, N., Jaillon, O., Mogensen, K., and Monneron, D.
659 (2003). Comparative genomic analysis reveals independent expansion of a lineage-specific

660 gene family in vertebrates: The class II cytokine receptors and their ligands in mammals and
661 fish. BMC Genomics *4*, 29.
662 Martin, S.A.M., Taggart, J.B., Seear, P., Bron, J.E., Talbot, R., Teale, A.J., Sweeney, G.E.,
663 Høyheim, B., Houlihan, D.F., Tocher, D.R., et al. (2007). Interferon type I and type II
664 responses in an Atlantic salmon (Salmo salar) SHK-1 cell line by the salmon TRAITS/SGP
665 microarray. Physiol. Genomics *32*, 33–44.
666 McConnell, S.C., Hernandez, K.M., Wcisel, D.J., Kettleborough, R.N., Stemple, D.L., Yoder,
667 J.A., Andrade, J., and de Jong, J.L.O. (2016). Alternative haplotypes of antigen processing
668 genes in zebrafish diverged early in vertebrate evolution. Proc. Natl. Acad. Sci. *113*, E5014–
669 E5023.
670 O'Farrell, C., Vaghefi, N., and Cantonnet, M. (2002). Survey of Transcript Expression in
671 Rainbow Trout Leukocytes Reveals a Major Contribution of Interferon-Responsive Genes in
672 the Early Response to a Rhabdovirus Infection. J. Virol. *76*, 8040–8049.
673 Palha, N., Guivel-Benhassine, F., Briolat, V., Lutfalla, G., Sourisseau, M., Ellett, F., Wang, C.H.,
674 Lieschke, G.J., Herbomel, P., Schwartz, O., et al. (2013). Real-Time Whole-Body Visualization
675 of Chikungunya Virus Infection and Host Interferon Response in Zebrafish. PLoS Pathog. *9*,
676 e1003619.
677 Päri, M., Kuusksalu, A., Lopp, A., Reintamm, T., Justesen, J., and Kelve, M. (2007). Expression
678 and characterization of recombinant 2′,5′- oligoadenylate synthetase from the marine
679 sponge Geodia cydonium. FEBS J. *274*, 3462–3474.
680 Purcell, M.K., Nichols, K.M., Winton, J.R., Kurath, G., Thorgaard, G.H., Wheeler, P., Hansen,
681 J.D., Herwig, R.P., and Park, L.K. (2006). Comprehensive gene expression profiling following
682 DNA vaccination of rainbow trout against infectious hematopoietic necrosis virus. Mol.
683 Immunol. *43*, 2089–2106.

684 Robertsen, B. (2018). The role of type I interferons in innate and adaptive immunity against
685 viruses in Atlantic salmon. Dev. Comp. Immunol. *80*, 41–52.
686 Robertsen, B., Bergan, V., Røkenes, T., Larsen, R., and Albuquerque, A. (2003). Atlantic
687 Salmon Interferon Genes: Cloning, Sequence Analysis, Expression, and Biological Activity. J.
688 Interf. Cytokine Res. *23*, 601–612.
689 Rothenburg, S., Deigendesch, N., Dey, M., Dever, T.E., and Tazi, L. (2008). Double-stranded
690 RNA-activated protein kinase PKR of fishes and amphibians: varying the number of double-
691 stranded RNA binding domains and lineage-specific duplications. BMC Biol. *6*, 12.
692 Rusinova, I., Forster, S., Yu, S., Kannan, A., Masse, M., Cumming, H., Chapman, R., and
693 Hertzog, P.J. (2013). INTERFEROME v2.0: An updated database of annotated interferon-
694 regulated genes. Nucleic Acids Res. *41*, 1040–1046.
695 Schoggins, J.W., and Rice, C.M. (2011). Interferon-stimulated genes and their antiviral
696 effector functions. Curr. Opin. Virol. *1*, 519–525.
697 Schoggins, J.W., Wilson, S.J., Panis, M., Murphy, M.Y., Jones, C.T., Bieniasz, P., and Rice, C.M.
698 (2011). A diverse range of gene products are effectors of the type I interferon antiviral
699 response. Nature *472*, 481–485.
700 Secombes, C.J., and Zou, J. (2017). Evolution of interferons and interferon receptors. Front.
701 Immunol. *8*, 209.
702 Staeheli, P., Yu, Y.X., Grob, R., and Haller, O. (1989). A double-stranded RNA-inducible fish
703 gene homologous to the murine influenza virus resistance gene Mx. Mol. Cell. Biol. *9*, 3117–
704 3121.
705 Stein, C., Caccamo, M., Laird, G., and Leptin, M. (2007). Conservation and divergence of gene
706 families encoding components of innate immune response systems in zebrafish. Genome
707 Biol. *8*, R251.

708 Venkatesh, B., Lee, A.P., Ravi, V., Maurya, A.K., Lian, M.M., Swann, J.B., Ohta, Y., Flajnik,
709 M.F., Sutoh, Y., Kasahara, M., et al. (2014). Elephant shark genome provides unique insights
710 into gnathostome evolution. Nature *505*, 174–179.
711 Zhang, Y.B., and Gui, J.F. (2004). Identification and expression analysis of two IFN-inducible
712 genes in crucian carp (Carassius auratus L.). Gene *60*, 1–9.
713 Zhang, F., Tao, Y., Zhang, Z., Guo, X., An, P., Shen, Y., Wu, Q., Yu, Y., and Wang, F. (2012).
714 Metalloreductase steap3 coordinates the regulation of iron homeostasis and inflammatory
715 responses. Haematologica *97*, 1826–1835.
716 Zhang, Y.B., Liu, T.K., Jiang, J., Shi, J., Liu, Y., Li, S., and Gui, J.F. (2013). Identification of a
717 Novel Gig2 Gene Family Specific to Non-Amniote Vertebrates. PLoS One *8*, e60588.
718 Zou, J., and Secombes, C.J. (2011). Teleost fish interferons and their role in immunity. Dev.
719 Comp. Immunol. *35*, 1376–1387.
720 Zou, J., Tafalla, C., Truckle, J., and Secombes, C.J. (2007). Identification of a second group of
721 type I IFNs in fish sheds light on IFN evolution in vertebrates. J. Immunol. *179*, 3859–3871.
722
723
724 **Figure legends**
725
726 **Figure 1. Zebrafish larva transcriptional response to type I IFN** (A) Fold change (FC)/basal
727 expression representation with all genes detected in the analysis. Basal expression was the
728 average of read numbers mapped onto a given gene in control (BSA-injected) larvae. FC is
729 the ratio of reads numbers in IFNφ1-injected larvae divided by basal expression. pVal
730 corresponds to adjusted p-Value. (B) FC/basal expression representation, limited to ISGs

731 identified in IFNφ1-injected larvae (e.g. red dots in panel A); key genes commented in the
732 text are indicated with gene families identified by different colors. (C) FC/basal expression
733 representation for zebrafish ISGs represented according to the type of orthology with
734 human genes, and if human orthologues include at least one human ISG. (D) Fold change
735 distribution for zebrafish ISG with different types of orthology relationship to human genes,
736 using the same color code as in C but pooling genes with single and multiple human
737 orthologues. Zebrafish ISGs with a human ISG within a vertebrate-level orthology groupes
738 were not included in this analysis.

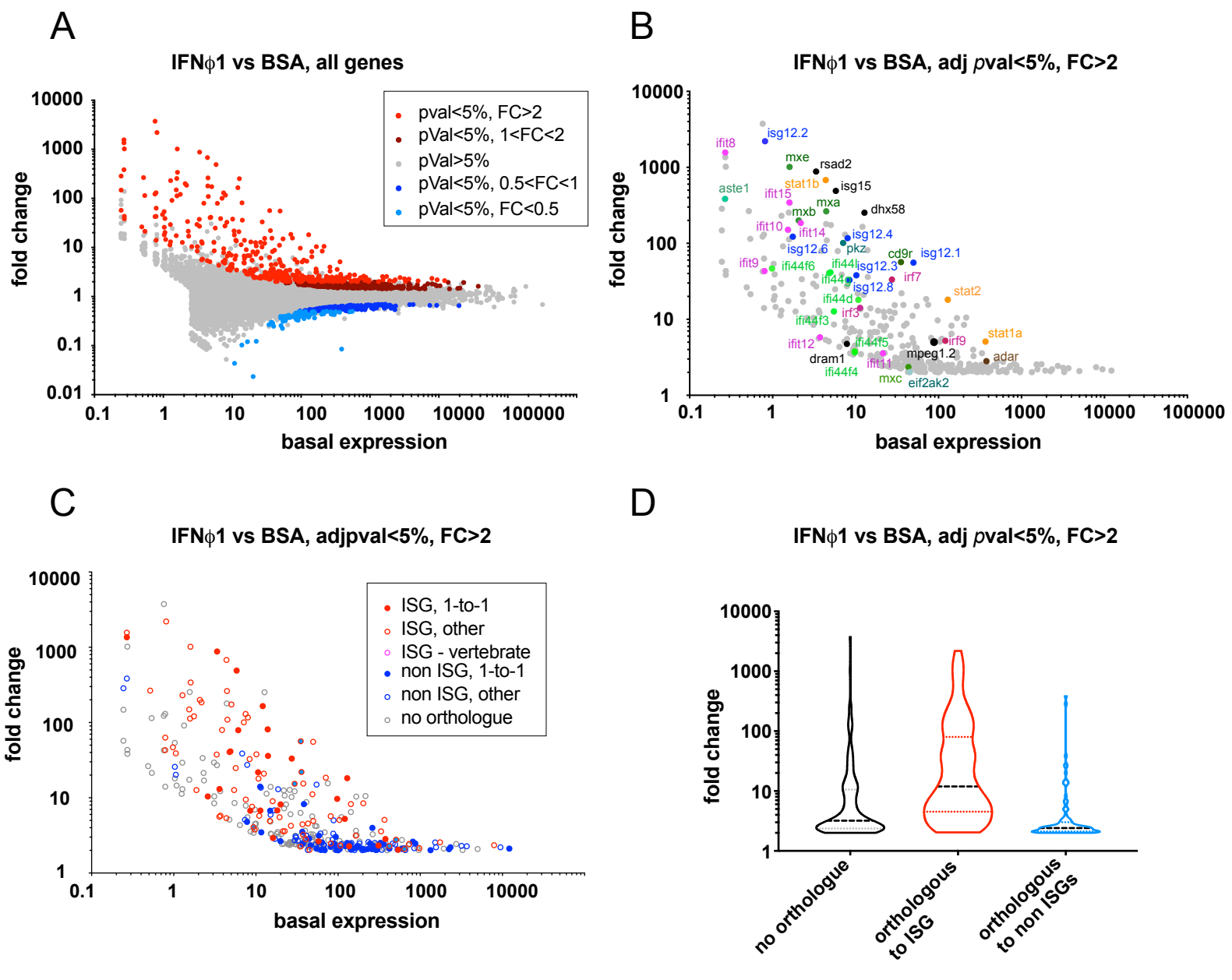
739

740 **Figure 2. Main families of fish-specific ISG and their domain organisation.** Each panel show
741 the typical domain organization of a family of fish-specific ISGs (e.g., with no human
742 orthologue), as determined by SMART analysis (http://smart.embl-heidelberg.de/). The
743 accession numbers of IFNφ1-induced genes (FC>2, pVal<5%) and the fraction they represent
744 within the family are indicated.

745

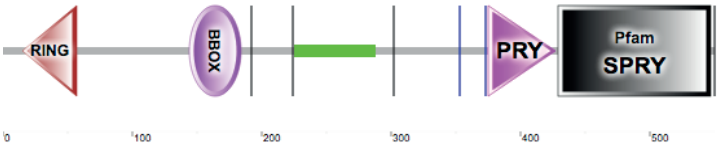
746 **Figure 3. Comparison of ISG repertoires identified by IFNφ1 induction and by CHIKV**
747 **infection.** (A) Venn analysis of genes up- or down-regulated in zebrafish larva by IFNφ1
748 treatment (IFNφ1), CHIKV infection (ChikV), and CHIKV infection in the context of type I IFN
749 receptor knock-down ("MoCrfb"). ISG are identified either by their responsiveness to IFNφ1
750 or by a CHIKV induction abolished in crfb1+2 morphants (venn diagram at the top). Genes
751 induced by CHIKV infection in an IFNφ1 independent way are analysed in the Venn diagram
752 at the bottom of the panel. (B) FC/FC representation of transcriptome response to IFNφ1
753 and CHIKV. Color code identifies genes significantly induced by IFNφ1 and/or CHIKV, and
754 genes which are significantly induced in one condition (FC>2, adj p val< 0.05) while almost

755 significantly induced in the other ("quasi Sig"; thresholds FC>1.5 and/or adj p value < 0.2).
756 (C) Venn analysis of genes significantly induced by IFNφ1 and/or CHIKV. Gene subsets
757 corresponding to genes induced quasi-significantly is represented within dotted lines.
758
759 **Figure 4. Graphic overview of the ancestral ISG repertoire**, organized by functional
760 modules.
761
762 **Figure 5. Evolution of the ISG repertoire** since the LCATT. Genes names are given as
763 examples, with no attempt to be exhaustive.
764
765 **Additional File 1. A. Kinetics of response to recombinant IFNϕ1**, measured by RT-qPCR. Data
766 normalized on the mean of expression of control (samples), displayed as mean ± SD. **B.**
767 ***ddx58,* encoding RIG-I, is an zebrafish ISG**, as measured by qRT-PCR in zebrafish larvae 6
768 hours after IFNϕ1 injection; quantification of *MXA* also shown as a control.
769
770 **Additional File 5. Transcriptomic response to CHIKV at 24 and 48 hours post infection.** 48h
771 data from (Briolat et al., 2014), 24h data from this study; only genes identified as
772 upregulated in the 2014 microarray study are displayed here.


A
IFNϕ1 vs BSA, all genes
pval<5%, FC>2
pVal<5%, 1<FC<2
pVal>5%
pVal<5%, 0.5<FC<1
pVal<5%, FC<0.5
fold change
basal expression
B
IFNϕ1 vs BSA, adj pval<5%, FC>2
isg12.2
ifit8
mxe
rsad2
stat1b
isg15
aste1
ifit15
mxa
mxb
dhx58
ifit10
ifit14
isg12.4
isg12.6
pkz
cd9r
isg12.1
ifit9
ifi44f6
ifi44l
isg12.3
irf7
ifi44
isg12.8
stat2
ifi44d
irf3
ifi44f3
stat1a
ifit12
ifi44f5
irf9
dram1
mpeg1.2
adar
ifit11
ifi44f4
mxc
eif2ak2
fold change
basal expression
C
IFNϕ1 vs BSA, adjpval<5%, FC>2
ISG, 1-to-1
ISG, other
ISG - vertebrate
non ISG, 1-to-1
non ISG, other
no orthologue
fold change
basal expression
D
IFNϕ1 vs BSA, adj pval<5%, FC>2
fold change
no orthologue
orthologous to ISG
orthologous to non ISGs

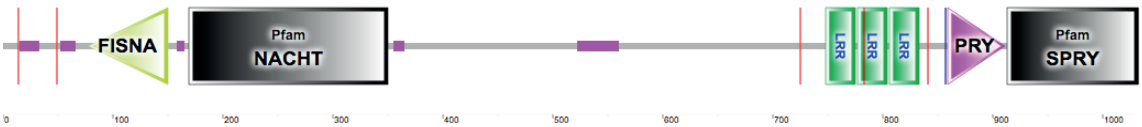
finTRIMs 21 ISGs out of ~80 paralogues

ENSDARG00000067661 ENSDARG00000095678 ENSDARG00000052971 ENSDARG00000026704 ENSDARG00000053293 ENSDARG00000030954 ENSDARG00000095345
ENSDARG00000038646 ENSDARG00000078065 ENSDARG00000098004 ENSDARG00000032043 ENSDARG00000095298 ENSDARG00000099940 ENSDARG00000074360
ENSDARG00000041220 ENSDARG00000100250 ENSDARG00000097679 ENSDARG00000079537 ENSDARG00000092961 ENSDARG00000095032 ENSDARG00000039852



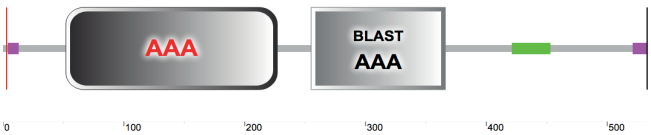
fish-specific NLRs 6 ISGs out of ~350 paralogues (several subfamilies)

ENSDARG00000102440 ENSDARG00000077488 ENSDARG00000077673 ENSDARG00000090033 ENSDARG00000078912 ENSDARG00000087597



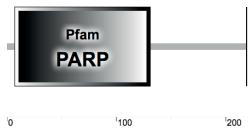
septin-related 2 ISGs out of 15 paralogues

ENSDARG00000053761 ENSDARG00000078093



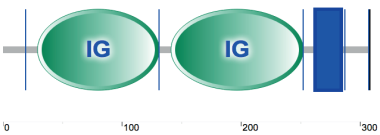
Gig2 family 2 ISGs of 11 paralogues

ENSDARG00000088260 ENSDARG00000086903



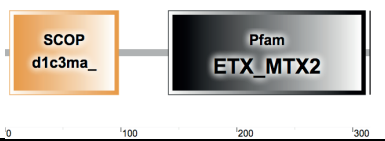
Fish two-Ig domain 3 ISGs out of ~140 paralogues

ENSDARG00000099577 ENSDARG00000095643 ENSDARG00000071657



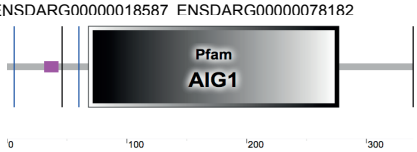
lectin 2 ISGs out of 14 paralogues

ENSDARG00000087318 ENSDARG00000039099



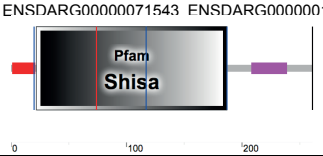
AIG1 domain 2 ISGs out of 4 paralogues

ENSDARG00000018587 ENSDARG00000078182



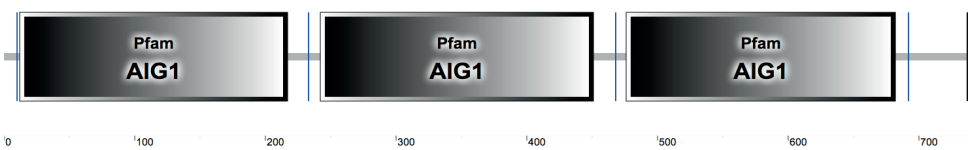
Shisa-related 2 ISGs out of 2 paralogues

ENSDARG00000071543 ENSDARG00000013681



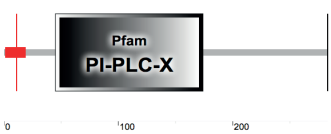
multi AIG1 domains 2 ISGs out of 10 paralogues

ENSDARG00000011581 ENSDARG00000101671



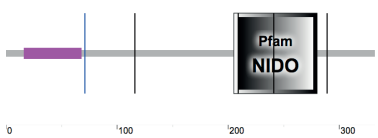
phospholipase 2 ISGs out of 5 paralogues

ENSDARG00000075622 ENSDARG00000099715



NIDO domain 2 ISGs of ouf 6 paralogues

ENSDARG00000103973 ENSDARG00000101930

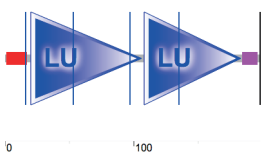


tandem ZNFs 2 ISGs out of ~260 genes

ENSDARG00000091243 ENSDARG00000099872

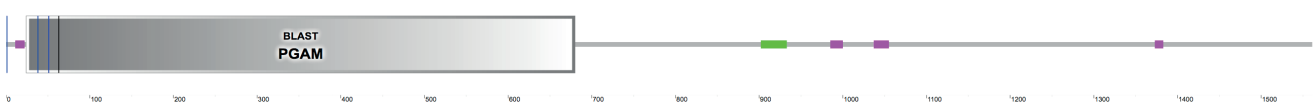
Ly-related, GPI anchor 2 ISGs out of 9 paralogues

ENSDARG00000104775 ENSDARG00000089123

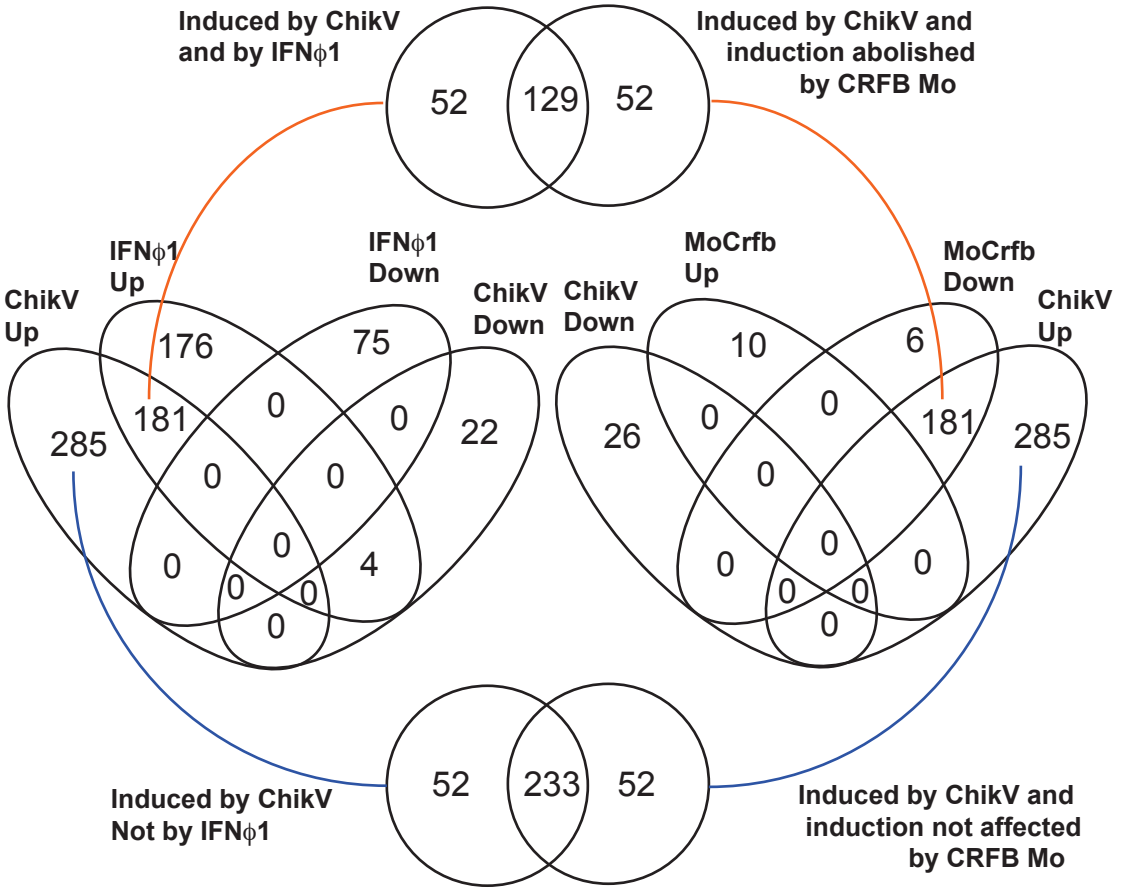


large, URGCP-related 3 ISGs out of 15 paralogues

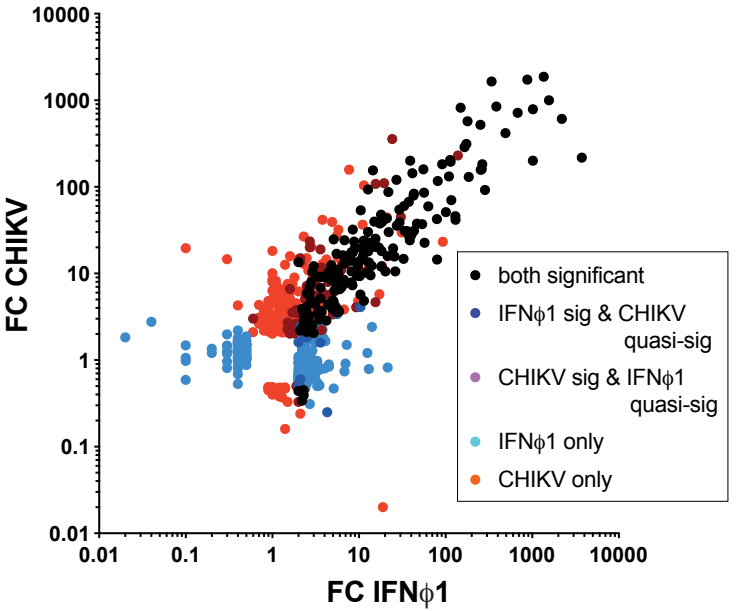
ENSDARG00000087575 ENSDARG00000101189 ENSDARG00000078832



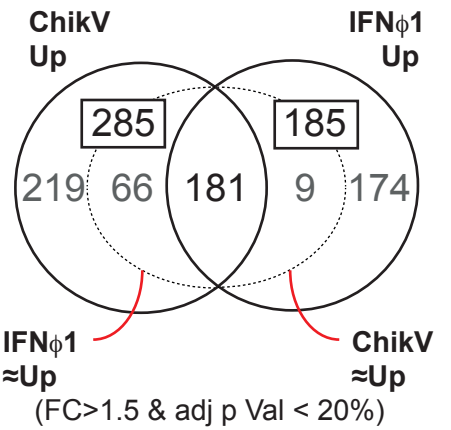
A.

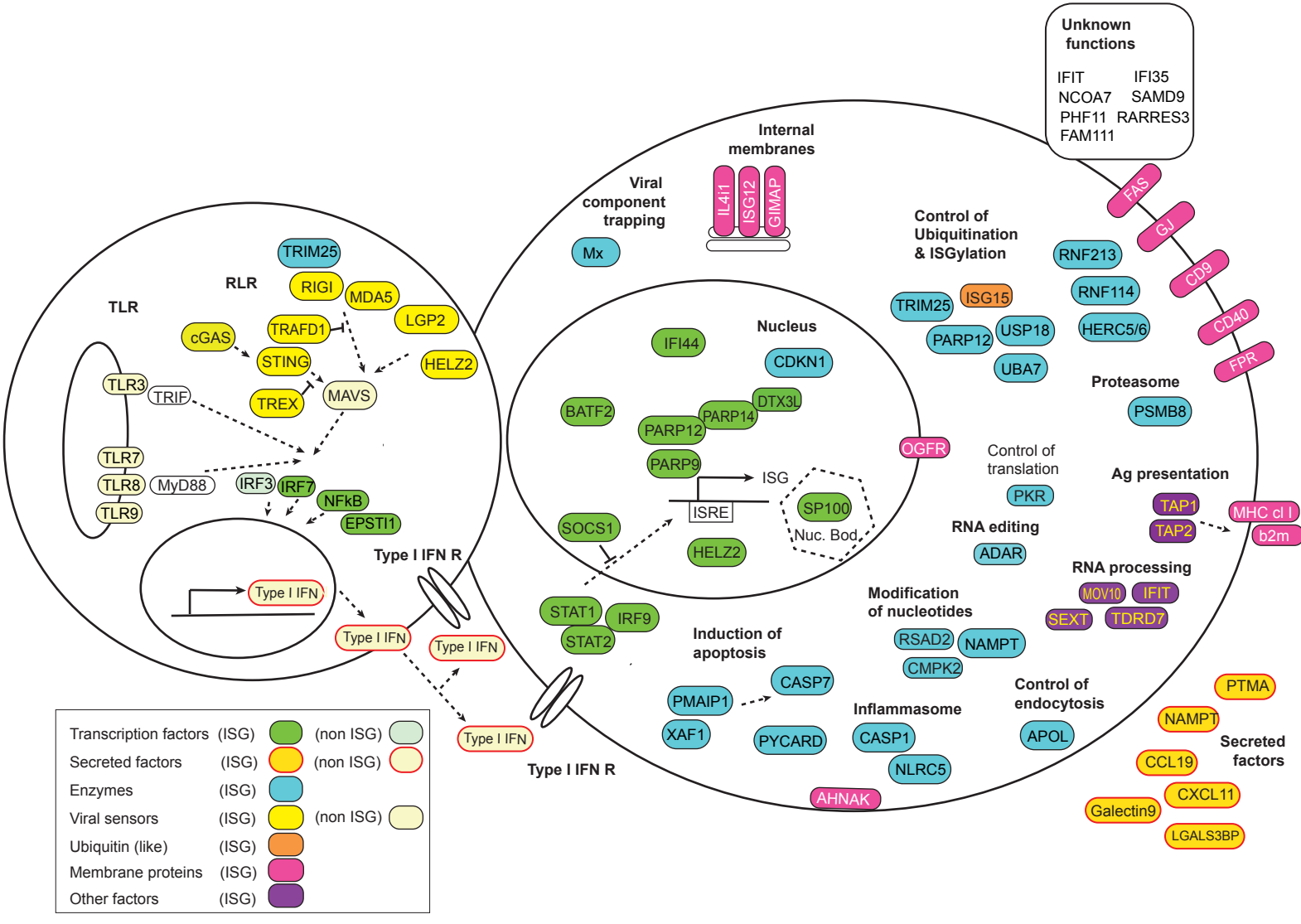

Induced by ChikV and by IFNϕ1
52
129
52
Induced by ChikV and induction abolished by CRFB Mo
ChikV Up
IFNϕ1 Up
IFNϕ1 Down
ChikV Down
ChikV Down
MoCrfb Up
MoCrfb Down
ChikV Up
285
181
176
0
75
0
22
0
0
0
0
0
0
4
0
26
0
10
0
6
181
285
0
0
0
0
0
0
0
Induced by ChikV Not by IFNϕ1
52
233
52
Induced by ChikV and induction not affected by CRFB Mo

B.

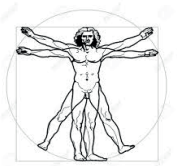

FC CHIKV
FC IFNϕ1
10000
1000
100
10
1
0.1
0.01
0.01
0.1
1
10
100
1000
10000
both significant
IFNϕ1 sig & CHIKV quasi-sig
CHIKV sig & IFNϕ1 quasi-sig
IFNϕ1 only
CHIKV only

C.


ChikV Up
IFNϕ1 Up
285
185
219
66
181
9
174
IFNϕ1 ≈Up
ChikV ≈Up
(FC>1.5 & adj p Val < 20%)


Unknown functions
IFIT
IFI35
NCOA7
SAMD9
PHF11
RARRES3
FAM111
Internal membranes
IL4i1
ISG12
GIMAP
Viral component trapping
Mx
Control of Ubiquitination & ISGylation
RNF213
RNF114
HERC5/6
TRIM25
ISG15
PARP12
USP18
UBA7
FAS
GJ
CD9
CD40
FPR
Proteasome
PSMB8
Nucleus
IFI44
CDKN1
BATF2
DTX3L
PARP14
PARP12
PARP9
ISG
ISRE
SP100
Nuc. Bod.
SOCS1
HELZ2
OGFR
Control of translation
PKR
Ag presentation
TAP1
TAP2
MHC cl I
b2m
RNA editing
ADAR
RNA processing
MOV10
IFIT
SEXT
TDRD7
Modification of nucleotides
RSAD2
NAMPT
CMPK2
STAT1
IRF9
STAT2
Induction of apoptosis
CASP7
PMAIP1
XAF1
PYCARD
Inflammasome
CASP1
NLRC5
AHNAK
Control of endocytosis
APOL
PTMA
NAMPT
CCL19
CXCL11
Galectin9
LGALS3BP
Secreted factors
TLR
RLR
TRIM25
RIGI
MDA5
LGP2
cGAS
TRAFD1
STING
HELZ2
TLR3
TRIF
TREX
MAVS
TLR7
TLR8
MyD88
TLR9
IRF3
IRF7
NFkB
EPSTI1
Type I IFN R
Type I IFN
Type I IFN
Type I IFN
Type I IFN
Type I IFN R
Transcription factors (ISG) (non ISG)
Secreted factors (ISG) (non ISG)
Enzymes (ISG)
Viral sensors (ISG) (non ISG)
Ubiquitin (like) (ISG)
Membrane proteins (ISG)
Other factors (ISG)

LCATT



Conserved ISG

Ancestral ISG maintained in teleosts as

single copy:

rsad2, trim25, isg15, apol

or multiple paralogues:

ifi44, mx, isg2,herc5/6, sp100

Ancestral ISG maintained in tetrapods as

single copy:

RSAD2,TRIM25, ISG15

or multiple paralogues:

APOL,SP100,MX

Loss and gain of ISG in fish or tetrapods

Genes present in teleosts and tetrapods but IFN-inducible only in teleosts

ctla4	*dram1*
il4r	*irf3*
nuggc	*rel, relb*
btr	*vamp8*
steap3	*casp3*

Loss or gain of inducibility

Old genes present in teleosts and tetrapods but IFN-inducible only in tetrapods

GZMB	*GPX2*
IFNGR1	*IRF1*
JAK2	*JUNB*
MyD88	*OPTN*
SMAD3	*IFI30*

ISG present in teleosts, not in tetrapods

gig2
fintrim
fish specific nlr
fish 2-Ig domains
phospholipases
pkz

Gene loss or gene gain either in fish or in tetrapods

ISG present in tetrapods, not in teleosts

OAS	*DDX60*
APOBEC3	*IFNLR1*
CLEC4	*IL15 & IL15RA*
RNAseL	*ISG2*
IDO1	*AIM2/IFI16*
BST2	*TRIM5, 19, 21, 22*

Table 1. Orthology groups defining ancestral ISGs

ancestral gene	zebrafish	human
	OSTEICHTYAN-LEVEL ORTHOLOGUES	
ADAR	adar	ADAR
AHNAK	ahnak	AHNAK
APOL	apol	APOL1, 2, 3, 6
B2M	b2m,b2ml	B2M
BATF2	batf2	BATF2
CASP1	caspb	CASP1, CASP4
CASP7	casp7	CASP7
CCL19	ccl19a.1, a.2	CCL19
CD40	cd40	CD40
cGAS	cgas	CGAS
CMPK2	cmpk2	CMPK2
CXCL11	cxcl11.3	CXCL11
DDX58 (RIGI)	ddx58	DDX58
DHX58 (LGP2)	dhx58	DHX58
DTX3L	dtx3lb1,2,3	DTX3L
EPSTI1	epsti1	EPSTI1
FAM111	fam111.1	FAM111A
FAS	fas	FAS
FPR	fpr	FPR2
GIMAP	gimapb1	GIMAP2
HELZ2	helz2a,b	HELZ2
HERC5/6	herc56.1, 2, 3, 4	HERC5, HERC6
IFI35	ifi35	IFI35
IFI44	ifi44a1, a5, c2, d, f3-6, g	IFI44, IFI44L
IFIH1 (MDA-5)	ifih1	IFIH1
IFIT	ifit8-12, 14-16	IFIT1-3, 5
IL4i1	il4i1	IL4I1
IRF7	irf7	IRF7
IRF9	irf9	IRF9
ISG12	isg12.1-4, 6-7	IFI6, IFI27
ISG15	isg15	ISG15
LGALS9	lgals9l1, 3	LGALS9, 9C
MOV10	mov10a, b.1, b.2	MOV10
MX	mxa, b, c, e	MX1, MX2
NAMPT	nampta, namptb	NAMPT
NCOA7	ncoa7a	NCOA7
NLRC5	nlrc5	NLRC5
OGFR	ogfrl1,2	OGFR
PARP12	parp12a,b	PARP12
PARP14	parp14a,c	PARP14
PARP9	parp9	PARP9
PHF11	phf11	PHF11
PKR	pkz, eif2ak2	EIF2AK2
PMAIP1	pmaip1	PMAIP1
PSMB8	psmb8a	PSMB8
PTMA	ptmaa	PTMA
PYCARD (ASC)	pycard	PYCARD
RARRES3	rarres3	RARRES3
RNF114	rfn114	RNF114
RNF213	rnf213a,b	RNF213
RSAD2	rsad2	RSAD2
SOCS1	socs1a,b	SOCS1
SP100	sp100.1, sp100.3, sp100.4	SP100, 110, 140, 140L
STAT1	stat1a,b	STAT1
STAT2	stat2	STAT2
TAP1	tap1	TAP1
TAP2	tap2a,t	TAP2
TDRD7	tdrd7b	TDRD7
TMEM173 (STING)	tmem173	TMEM173
TRAFD1	trafd1	TRAFD1
TREX	trex3	TREX1
TRIM25	trim25	TRIM25
UBA7	uba7	UBA7
USP18	usp18	USP18
XAF1	xaf1	XAF1
ZNFX1	znfx1	ZNFX1
	VERTEBRATE-LEVEL ORTHOLOGUES	
CD9	cd9r	CD9
CDKN1	cdkn1d	CDKN1A
GJ	cx30.3	GJA4
MHC class I	mhc1zba	HLA-A, -B, -C
SAMD9	samd9r	SAMD9
TABPB	tapbpr1, tapbpr2	TAPBP

Orthology groups that include ISGs in both zebrafish and human, and therefore define an ancestral IDSG in their common ancestor. This is a condensed version of Additional File 7, which also includes non-IFN inducible genes which belong to the osteichtyan-level orthology groups, and provides Ensembl Gene IDs. In red, our gene re-naming proposals.