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

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

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

[†] 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

Funding: ANR grant Fish-RNAVAX (ANR-16-CE20-0002-02 and ANR-16-CE20-0002-03), European Union's Horizon 2020 research and innovation programme, Marie Sklodowska-Curie grant agreement No 721537 "ImageInLife"

Running title:

Comparing ISG repertoires of zebrafish and human

Abstract

The evolution of the interferon (IFN) system, the major innate antiviral mechanism of vertebrates, remains poorly understood. According to the detection of type I IFN genes in cartilaginous fish genomes, the system appeared 500My ago. However, the IFN system integrates many other components, most of which are encoded by IFN-stimulated genes (ISGs). To shed light on its evolution, 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, a potent IFN inducer. 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 containing 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.

Key points

- We established an exhaustive list of larval zebrafish ISGs.
- Orthologous ISGs in fish and human identify a large ancestral ISG repertoire.

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 (1). The origin of the IFN system, which seems to have replaced
53 the RNA interference antiviral system still used by plants and most invertebrates (2, 3), is
54 shrouded in mystery. Cartilaginous fish are the most basal clade with genomes containing
55 recognizable type I IFN genes (4), suggesting that this system appeared about 500My ago
56 with jawed vertebrates. The IFN system integrates many components, most of which are
57 encoded by ISGs, which can be traced back in genomes from distant clades. However,
58 finding the orthologue(s) of a human ISG in another taxon does not imply that this gene is
59 part of its IFN system. To understand the evolution of antiviral immunity, it is therefore
60 desirable to establish how the repertoire of ISGs changed from early to modern vertebrates.
61 This can be inferred by comparing the ISGs of current living representatives of distant
62 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 (5–8). Teleost fish possess several
66 subgroups of type I IFNs (but no type III genes), with strong variation in gene numbers
67 among fish taxa (7). The zebrafish possess four type I IFN genes, named *ifnphi1-4*; only
68 *ifnphi1* and *ifnphi3* are active at the larval stage (9). Their receptors have been identified (9).
69 Even before fish IFNs were known, the first fish ISGs were identified by homology cloning

70 from cells stimulated by poly-I:C (10) or by differential transcript analysis of cells infected by
71 viruses (11, 12). Because many virus-induced genes (*vig*) were homologous to well-known
72 mammalian ISGs, they were hypothesized to be IFN-inducible, which was often confirmed by
73 later studies, as in the case of *vig-1*, the *rsad2/viperin* orthologue (11, 13). Similarly, upon
74 cloning of fish IFNs, induction of *Mx* (10) was used as a readout for their activity (14–16),
75 confirming it was an ISG. The list of fish homologues of known ISGs rapidly grew with the
76 release of new fish genomes and EST collections, allowing the development of micro-arrays
77 to study fish response to virus or recombinant type I IFNs (17, 18). Candidate gene
78 approaches were also developed, testing orthologues of known mammalian ISGs in qRT-PCR
79 assays in multiple fish infection models (13, 19, 20). In parallel, approaches without a priori
80 identified fish ISGs that had no orthologue in mammals, although they belonged to gene
81 families involved in antiviral immunity. A large set of tripartite-motif protein-encoding
82 genes, called *fintrims* (*ftr*), distantly related to *trim25* was identified in rainbow trout cells as
83 a induced by virus infection (12) and later shown to form multigene families in teleosts,
84 particularly extensive in zebrafish (21). Similarly, a family of IFN induced ADP-
85 ribosyltransferases named *Gig2* was identified in crucian carp cells treated with UV-
86 inactivated GCHV (grass carp hemorrhage virus) (22, 23). Some ISG were restricted to
87 particular fish groups such as the non-coding RNA *vig2* that is found only in salmonids (24).
88
89 We previously established a list of zebrafish candidate ISGs using microarray analysis (25).
90 For this, we compared the response to a poor IFN inducer, IHNV (infectious hematopoietic
91 necrosis virus) (26) and a strong IFN inducer, CHIKV (chikungunya virus) (27). However, the
92 array did not include the full complement of zebrafish genes, and the study identified virus-
93 induced genes which were not necessarily ISGs. Here, to directly identify ISGs, we analyze

94 the transcriptional response of zebrafish larvae injected with recombinant type I IFN. We
95 rely on deep RNA sequencing, which is intrinsically quasi-exhaustive. Our approach is
96 therefore limited mainly by the quality of genome assembly and annotation, which is
97 excellent for the zebrafish (28). We complemented this analysis with a study of the response
98 to CHIKV and its dependence to expression of the zebrafish IFN receptor chains *crfb1* and
99 *crfb2* (9). We thus established a comprehensive list of ISGs of zebrafish larvae, and
100 performed a detailed comparison with the human ISG repertoire. Our comparative analysis
101 was facilitated by a compilation of human ISGs made to perform a systematic screen (29),
102 and by the specialized database Interferome (30). We identify about 70 orthology groups
103 that include ISG in both species and thus approximate the ISG repertoire of the common
104 ancestor of all Osteichthyes. As ISGs typically evolve fast, with frequent duplications and
105 gene loss, we also identify many families of fish-specific ISGs, which represent a rich
106 resource for seeking new antiviral mechanisms. Our study provides a broad overview of the
107 evolutionary patterns of genes belonging to the type I IFN pathway, and identifies gene
108 modules induced by a viral infection independently of IFN.
109
110
111

112 **Materials and methods**

113

114 *Zebrafish husbandry*

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

126

127 *Interferon and virus inoculation*

128 Recombinant zebrafish IFNφ1 (9), kindly provided by Rune Hartmann (University of Aarhus,
129 Denmark), was inoculated by intravenous (IV) injection in the caudal cardinal vein. One
130 nanoliter of 1mg/ml IFNφ1, or as a control, bovine serum albumin (BSA, New England
131 Biolabs) in PBS1x/10% glycerol was injected. CHIKV infections were performed as described
132 (25, 27). Briefly, ~200PFU of CHIKV115 was injected IV in a volume of 1 nl at 3dpf.

133

134 *Interferon receptor knock-down*

135 Morpholino antisense oligonucleotides (Gene Tools) were injected at the one to two cells
136 stage as described (31). Two ng of *crfb1* splice morpholino
137 (CGCCAAGATCATACCTGTAAAGTAA) was injected together with 2ng of *crfb2* splice
138 morpholino (CTATGAATCCTCACCTAGGGTAAAC), knocking down all type I IFN receptors (9).
139 Control morphants were injected with 4 ng of control morpholino
140 (GAAAGCATGGCATCTGGATCATCGA) with no known target.
141
142 *RNA extraction*
143 For RNAseq analysis, total RNA was extracted from replicate pools of 10 injected larvae (at 6
144 hours post injection for IFNφ1 treatment, or 24 hours post injection for CHIKV infections),
145 using TRIzol (Invitrogen), following the manufacturer's protocol. The integrity of the RNA
146 was confirmed by laboratory-on-chip analysis using the 2100 Bioanalyzer (Agilent
147 Technologies), using only samples with an RNA integrity number of at least 8.
148
149 *Illumina sequencing*
150 Libraries were built using a Truseq mRNA-Seq Library Preparation Kit (Illumina, USA),
151 according to the manufacturer's recommendations. Quality control was performed on an
152 Agilent Bioanalyzer. Sequencing was performed on a HiSeq 2500 system (Illumina, USA) and
153 produced 65-base single-end reads.
154
155 *Mapping reads and gene expression counts*
156 Sequences were trimmed using cutadapt (v1.8.3). The reads quality was checked with
157 FastQC. Reads were then spliced-aligned to the zebrafish genome (GRCz10, Ensembl release
158 88) using TopHat2 (v2.0.14). The average number of mapped read per sample was

159 16.500.000. Only fragments mapping coherently and unambiguously to genes have been
160 considered for gene counts. Gene counts have been assigned using featureCounts v1.5.2
161 (Liao, 2014).
162
163 *Identification of differentially expressed genes (DEG)*
164 Differentially expressed nuclear genes between larvae treated with IFNφ1 and
165 controls, between larvae infected by CHIKV and controls, or between crfb1+2 and control
166 morphants all infected by CHIKV, were identified. DEG were identified using DESeq 1.18.0
167 (BioConductor) (Love et al., 2014) and R: 3–1-2 (R core team, 2017). Briefly, raw counts of
168 genes were subjected to a minimal pre-filtering step: genes for which the count sum, per
169 group of samples, was equal or higher than 10, in at least one group, were kept. Raw counts
170 were normalized for library size and normalized data were fitted using a negative binomial
171 general linear model. Data were adjusted for multiple testing using the Benjamini-Hochberg
172 procedure (adjusted p value). Genes with an adjusted *p* value less than 0.01 and an absolute
173 Fold Change (FC) > 2 or FC < 0.5 were considered as DEGs.
174 Sequence data were registered in the BioProject ncbi database with the SRA accession
175 number: PRJNA531581.
176
177 *Identification of human orthologues and ISGs*
178 Orthology analysis was primarily based on data from the Ensembl database
179 (www.ensembl.org), the zfin database (zfin.org), and the literature, notably our previous
180 analysis of zebrafish orthologues of human ISGs (25). Data were systematically curated
181 manually, and conflicts resolved using a combination of literature search, synteny analysis,
182 and sequence homology analysis (two-way protein BLAST). When human genes were

183 present on the list compiled by Schoggins *et al.* (29), they were labelled as ISGs. If absent
184 from the list, gene names were further queried on the Interferome website
185 (http://www.interferome.org/interferome/search/showSearch.jspx) which compiles the
186 results of many transcriptomic studies on human and mouse samples after IFN stimulation
187 (30). We postulated that human genes present in Interferome could be considered as ISG
188 when being significantly induced at least 2-fold with a stimulation for no more 12 hours by
189 type I IFN in at least 4 datasets.
190
191 *qRTPCR*
192 RNA was extracted from individual larvae using RNeasy® Mini Kit (Qiagen). cDNA was
193 obtained using M-MLV H- reverse-transcriptase (Promega) with a dT17 primer. Quantitative
194 PCR was then performed on an ABI7300 thermocycler (Applied Biosystems) using Takyon™
195 ROX SYBR© 2x MasterMix (Eurogentec) in a final volume of 25 µl. The following pairs of
196 primers were used: *ef1a* (housekeeping gene used for normalization): 5'-
197 GCTGATCGTTGGAGTCAACA-3' and 5'-ACAGACTTGACCTCAGTGGT-3'; *mxa*: 5'-
198 GACCGTCTCTGATGTGGTTA-3' and 5'-GCATGCTTTAGACTCTGGCT-3'; *ddx58*: 5'-
199 ACGCCGGAGAAAGAATTTTTC-3' and 5'-TCGACAGACTCTCGATGTTG-3'; *aqp9a*: 5'-
200 CTGTACTACGACGCCTTCAT -3' and 5'-GAGAATACAGAGCACCAGCA -3'.
201
202
203

204 **Results**

205

206 *RNAseq analysis of IFNφ1-regulated genes*

207 To make an inventory of zebrafish ISGs, we first injected 3dpf larvae with
208 recombinant zebrafish IFNφ1, the first type I IFN to be identified in zebrafish, or BSA as a
209 negative control. Based on preliminary kinetic experiments, we chose 6 hours post injection
210 as the early plateau phase of ISG expression (Figure S1A). RNA was extracted from multiple
211 pools of ten larvae and subjected to deep sequencing using an Illumina-based platform
212 sequencing. Reads were mapped to zebrafish genome (zv10), and the differential analysis
213 performed using the DESeq package.

214 Choosing as cutoff values adjusted *p* values <5% and fold-change (FC) >2, we
215 identified 360 IFNφ1 up-regulated genes (which are ISGs, by definition) and 75 down-
216 regulated genes (Table S1). As expected, genes with high basal expression levels tended to
217 display lower fold-change (Figure 1A). The top IFNφ1-upregulated genes (with FC >100)
218 comprised many well-known ISGs, many previously used as IFN signature genes in zebrafish,
219 including several *mx* genes, *rsad2, cmpk2*, several *ifit* genes, the ubiquitin-like *isg15*, the
220 helicase *dhx58* (aka *lgp2*), the kinase *pkz*, the transcription factor *stat1b*, and the chemokine
221 *ccl19a.2* (Figure 1B). To our surprise, *ddx58* (encoding RIG-I), a well-known and conserved
222 ISG (32), was not found in that list. In fact, the gene model is missing on the zebrafish
223 reference genome, with only fragments of the sequence present in the assembly. Therefore,
224 we performed qRT-PCR for *ddx58* and confirmed that it is induced by IFNϕ1 (Figure S1B). We
225 searched the list of zebrafish orthologues of human ISGs (25) for more genes not annotated
226 on the reference genome and found only one besides *ddx58*: *aqp9a*. By qRT-PCR, this gene

227 appeared to be moderately (0.75 fold) downregulated by IFNϕ1 (Figure S1B), and thus was
228 not an ISG.
229 Among the 360 zebrafish ISGs identified by RNAseq, 23 corresponded to non-coding
230 ISGs or transposons with no clear homologues in mammals, and were excluded from further
231 phylogenetic analyses.
232 Gene ontology (GO) analysis was performed using David and Gorilla, and showed, as
233 expected, that IFNφ1 upregulated genes were strongly enriched in genes linked to “antiviral
234 response”, “type I interferon pathway” and "antigen processing and presentation" (not
235 shown). Similarly, enrichment analyses identified KEGG pathways for Influenza, Measles and
236 Herpes simplex infection as well as RIG-I signaling and cytosolic DNA sensing pathways. The
237 list of downregulated genes was not found associated with any particularly notable function
238 in these analyses.
239
240 *Human orthologues of zebrafish ISGs are enriched in ISGs*
241 We then searched for the human orthologues of the 337 identified protein-coding
242 zebrafish ISGs (Table S2). All types of orthology relationships between zebrafish and human
243 were observed, from none to "many-to-many". One-to-one orthology was found for 77
244 genes (Figure 1C). We identified one or several human orthologues for 200 zebrafish ISGs.
245 This proportion (200/337, 59%) is significantly lower than the 71% reported for the whole
246 genome (28) (Fisher’s exact test, $p<0.0001$).
247 We then searched which of these human genes were themselves ISGs. We found 61
248 ISGs present in the list of 446 human ISGs compiled by Schoggins et al. (29); by querying the
249 Interferome database (30), we identified 11 additional human ISGs (Table S2). In total, 97
250 zebrafish IFNϕ1-inducible genes were orthologous to at least one human ISGs (Figure 1C). In

251 addition, we identified a handful of genes that were not true orthologues, but shared
252 ancestry with a human ISG at the vertebrate level, such as MHC class I genes (see comments
253 on Table S2).

254 As expected, human orthologues of zebrafish ISGs were strongly enriched in ISGs:
255 while there are 446 human ISGs out of 20454 genes in the genome (i.e., 2%), we found 72
256 ISGs among the 196 human orthologues to zebrafish ISGs (i.e., 37%) (Fisher's exact test, $p<$
257 0.0001). Interestingly, FC values of zebrafish ISGs were higher when they were orthologous
258 to a human ISG than when their orthologues were not ISGs (red vs blue on Figure 1C-D),
259 while FCs of zebrafish ISGs without human orthologues (grey on Figure 1C-D) were
260 intermediate (mean values, 127.7, 10.6 and 46.9, respectively; all groups significantly
261 different from each other, $p<0.001$, Kruskal-Wallis test). Thus, inducibility by type I IFNs is
262 often evolutionary conserved, making it possible to infer an ancestral set of ISGs.

263

264 *Fish-specific ISG genes and families*

265 Consistent with the expected high rate of duplication and divergence of ISGs, a
266 significant proportion of zebrafish IFNφ1-up-regulated protein-coding genes had no
267 identifiable orthologue in the human genome (137/337; 41%). Interestingly, but not unlike
268 genes with human orthologues, many of these genes belonged to multigenic families. We
269 show on Figure 2 the fish-specific gene families that contain several zebrafish ISGs, with
270 domains identified by the SMART tool (33).

271 The genes listed on Figure 2 included more than 20 *fintrim (ftr)* genes, a family
272 identified first as virus-inducible genes in rainbow trout (12), highly diversified in zebrafish,
273 and hypothesized to antagonize retroviruses (21). Of note, besides finTRIM, there are two
274 other large TRIM gene expansions in zebrafish, each with a single human orthologue of

275 unknown function (34). The *btr* (bloodthirsty-like TRIMs), related to human TRIM39, include
276 several ISGs. By contrast, no member of the TRIM35-like family was upregulated by IFNφ1.
277 Another family had been described as virus-inducible in fish: *gig2* (GCRV-induced
278 gene 2) genes (22). *gig2* genes are distantly related to the PARP family (23), which include
279 several ISGs in humans and zebrafish. The zebrafish *gig2p* and *gig2o* genes are induced by
280 IFNφ1.
281 To our knowledge, the genes in the remaining fish-specific families had not previously
282 been described to be interferon- or virus-inducible. These families are diverse, encoding
283 proteins expected to be membrane receptors, and presumably secreted, nuclear, or
284 cytosolic proteins (Figure 2).
285 Eight members of the very large NLR family were ISGs. These genes belong to groups
286 1, 2, 3 and 4 as defined in (35), and two of them belong to a fish specific subset defined by
287 the presence of a C-terminal B30.2 (or PRY-SPRY) domain (36, 37), which is most similar to
288 the corresponding domain of finTRIM genes (21). The specific function(s) of zebrafish NLR
289 genes remain poorly understood, but this highly expanded family may be central for
290 inflammatory mechanisms.
291 Additionally, 3 ISGs corresponded to membrane proteins with two immunoglobulin
292 (Ig) domains and a transmembrane region but not ITAM or ITIM (immunoreceptor tyrosine-
293 based activating/inhibiting motif). These genes belong to a very large family with 140
294 members, which we propose to name f2Ig (fish genes with two Ig domains).
295
296 *RNAseq analysis of CHIKV-induced genes*
297 Experimental infection of zebrafish larvae with chikungunya virus (CHIKV) induces a
298 strong type I IFN response (27). Our previous microarray-based analysis indicated that the

299 response to CHIKV was dominated by ISGs (25). However, to allow comparison with another
300 virus with slower IFN induction kinetics, this analysis had been performed at 48 hours post
301 infection (hpi), while the peak of the IFN and ISG response, as determined by qRT-PCR, is at
302 24 hpi (25). Therefore, we re-analyzed here the transcriptome of CHIKV-infected larvae at
303 24 hpi using deep RNA sequencing. Choosing the same cutoff values as for the IFNφ1
304 analysis, we identified 466 CHIKV up-regulated genes and 26 down-regulated genes (Table
305 S3). Hundreds of new CHIKV-inducible genes were identified, either because they were
306 absent from the microarray, or because their induction were below the cutoff of the first
307 analysis. Among the genes significantly upregulated in the microarray study, all those with a
308 human ISG orthologue were also upregulated in this new dataset, and, as expected, typically
309 much more (Figure S2).

310 About half of genes induced by IFNφ1 were also induced by the viral infection
311 (181/360 Figure 3A; Table S3, yellow), including almost all (84 out of 97) genes orthologous
312 to a human ISG, such as *mxa, b* and *e, stat1a* and *b; stat2; rasd2, isg15*, etc. There was a
313 clear correlation of the FC values for genes induced by both IFNϕ1 and CHIKV (Figure 3B).
314 However, almost two thirds of the genes induced by CHIKV infection were not significantly
315 modulated by IFNφ1 (285/466; Figure 3A). We then asked whether this CHIKV-specific
316 response could correspond essentially to genes for which there was weak induction by
317 recombinant IFNφ1, below our arbitrary cutoff. We therefore extracted from this list genes
318 induced by IFNφ1 with FC>1.5 and with an adjusted *p* value<20%, and we found 66 genes
319 matching these conditions (Table S3, green): 21 genes without annotation and 45 annotated
320 genes, many of which were notoriously linked to the type I IFN system. These genes notably
321 comprised *crfb1*, encoding a type I IFN receptor subunit, and two other cytokine receptors
322 *il10Ra* and *il13Ra*; four chemokines (*ccl34, cxcl11.6, cxc18b* and *cxc20*); ten additional *fintrim*

323 and 3 other members of the *gig2* family; and two *irf* transcription factors (2 and 10). It also
324 includes the metalloreductase *steap3*, whose mammalian orthologue is not an ISG, but
325 regulates type I IFN response, CXCL10 induction and iron homeostasis in mouse
326 macrophages (38).
327 Besides this intermediate gene set, a conservative list of 219 genes seems to be up-
328 regulated only by the virus (FC>2 and adjusted *p* value <5%), independently of IFNφ1 (FC<1.5
329 or adjusted *p* value >20%) (Table S3, blue; Figure 3C). This list contains 105 genes without
330 annotation, but also several functional modules providing interesting insights on the virus-
331 host interactions. Functional analysis using DAVID identified 6 enriched KEGG pathways,
332 namely Cytokine-cytokine receptor interaction, Cytosolic DNA-sensing, Toll-like receptor
333 signaling, RIG-I-like receptor signaling, Proteasome, Herpes simplex infection.
334 Importantly, type I IFNs were induced by the infection. Consistent with our previous report
335 with another virus (9), *ifnphi1* and *ifnphi3* were clearly dominant at this larval stage, with
336 58±3 and 47±11 reads respectively, compared with 9±4 reads for *ifnphi2*, and none detected
337 for *ifnphi4*. Two pro-inflammatory cytokines *il1b* and *tnfb* were also upregulated. Among
338 typical sensors, *tlr3, mb21d1* (encoding cGAS) and its downstream adaptor *tmem173*
339 (encoding STING), and several kinases of the IFN signaling pathways (*ripk1, tbk1*) were
340 present. Seven proteasome subunits are induced by the virus, suggesting activation of
341 protein degradation and Ag presentation pathways. The complement pathway also stands
342 out as an important module upregulated by CHIKV infection: twelve complement
343 component genes (*c1, c2*, several *c3, c7; c9; cfB; cfhl-1,-3* and *-5*) were induced by CHIKV,
344 suggesting that it is an important defense triggered in a Type I IFN independent manner.
345 Additionally, this response comprises 3 metalloaminopeptidases (*anpepb, erap1b* and *2*); the
346 myeloid markers *ncf1, mpx* and *marco*; 2 guanylate binding proteins (*gbp1* and *2*) that have

347 well known orthologues in human; the transcription factors *atf3* and *irf1b*, and with a high
348 level of expression, the enzyme *rnasel3* (an orthologue of human RNASE4, not of RNASEL, an
349 ISG with no fish counterpart). Nine *fintrim* and 3 *btr* can also be noted, underscoring the
350 importance of these TRIM with PRY/SPRY domains in virus host interactions altogether.
351 Thus, CHIKV induces a typical IFN-stimulated response of high magnitude, but also a broader
352 and less overt inflammatory response.
353
354 *IFN receptor dependence of the response to CHIKV*
355 To test the IFN-dependence of the response to CHIKV, we used morpholinos to knock down
356 in zebrafish larvae *crfb1* and *crfb2* which encode specific chains of the two types of type I IFN
357 receptors of zebrafish (9). We previously showed that such IFNR morphant larvae are
358 hypersusceptible to CHIKV infection, dying 2 to 3 days after virus injection (27). We analyzed
359 by deep RNAseq their transcriptional response to CHIKV at 24 hours post-inoculation, and
360 compared it to that of control morphant larvae. Choosing as cutoff values adjusted *p* values
361 <5% and a ratio between IFN-R morphants and controls >2, we identified 187 genes for
362 which induction was dampened by IFNR knockdown, and 10 genes that were upregulated in
363 morphants (Table S4). Among CHIKV-induced genes (Table S3), 181 were IFNR-dependent,
364 representing a significant fraction (181/466; 39%) (Figure 3A). Predictably, the list of genes
365 upregulated by CHIKV in a IFNR-dependent manner largely - but not fully - overlapped with
366 the gene set induced by recombinant IFNφ1 (129/181; 71%, see Figure 3A). This approach
367 led us to classify 52 new zebrafish genes as ISGs, being induced by CHIKV in an IFNR-
368 dependent manner, even if they were not significantly induced by recombinant IFNφ1. As
369 previously, we searched the human orthologues of these additional ISGs (Table S2, bottom),
370 identifying a few more human ISGs in this list, such as cGAS, NLRC5 or IFI35.

371 Together, our results provide a near-exhaustive list of zebrafish ISGs at the larval stage,
372 identified by two independent approaches, and a useful reference for future studies.
373
374 *Ancestral ISGs*
375 Assuming that the common ancestor of genes that are IFN-inducible in both human and
376 zebrafish was itself an ISG in their last common ancestor ~450 My ago, we can define a list of
377 ancestral ISGs. We identified 66 orthology groups that included an ISG on both the human
378 and the zebrafish sides (Table I, Table S5). A few more ancestral ISGs were also defined by
379 pairs of ISGs with orthology relationships at the early vertebrate or gnathostome level –
380 meaning that the zebrafish gene is not directly orthologous to a human ISG, but is
381 paralogous (with an ancestral taxonomy level labelled in Ensembl as "vertebrates" or "jawed
382 vertebrates") to another gene itself orthologous to a human ISG. In total, our list includes 72
383 ancestral genes (Table I, Figure 4A).
384 Based on our orthology analysis, we propose new, more explicit names for many of the
385 zebrafish ISGs with known human orthologues (in red on Table S5). This ancestral ISG core
386 includes most ISGs with known functions. The IFN system of 450My ago seems fairly similar
387 to the present one, particularly in its signaling components (Figure 4). Many ancestral genes
388 have been duplicated independently in one or both lineages (Table S5), in addition to
389 multiple ISGs apparently gained by either group (Figure 5).
390
391 *IFN-downregulated genes*
392 Two of the most strongly IFNφ1-downregulated genes (Table S1, bottom) were orthologous
393 to human genes downregulated by type I IFNs, according to the Interferome database: *plin1*

394 (perilipin 1) and *acox1* (palmytoil acyl-CoA oxidase 1). This suggests that downregulation of
395 fatty acid oxidation pathway is an ancient feature of the IFN system.
396 Many IFNφ1-downregulated genes were orthologous to a human gene in a 1-to-2 manner,
397 with the two zebrafish paralogues having arisen during the teleost specific whole genome
398 duplication (ohnologues). Systematically, only one of the two paralogues was
399 downregulated.
400 Remarkably, no gene was downregulated by both IFNφ1 injection and CHIKV infection
401 (Figure 3b, Table S3).
402
403

404 **Discussion**

405

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

417

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

419 A number of studies have identified genes induced by IFN or viral infections in fish (reviewed
420 in (6)). However, very few global descriptions after treatment with recombinant type I IFN
421 have been reported, using micro-arrays (18). Micro-array analyses are limited by probe
422 choice, and are typically biased towards genes with known human homologues. RNAseq, by
423 contrast, is mainly limited by the genome annotation quality and by the analysis method,
424 and can be reanalyzed; this approach is thus more complete. Since the early zebrafish larva
425 constitutes a reference model for investigating innate immune response, drug screening as
426 well as for modelling diseases, we undertook a comprehensive description of the repertoire
427 of ISG up-regulated at this developmental stage. Importantly, we previously reported a clear

428 transcriptional response of zebrafish embryos to IFNφ1 as early as 24 hpf (Levraud et al.
429 2007); the responsiveness to type I IFNs is thus already well established at 3 dpf. We are
430 aware that cells present in adult but not yet in larvae, notably those of the adaptive immune
431 system such as lymphocytes and dendritic cells may express additional ISGs, which should be
432 assessed in further work.
433 There are two groups of type I IFNs in teleost fish (39) with two different receptors (9). This
434 study only addresses the ISG repertoire induced by IFNφ1 (a group 1 IFN) and it is possible
435 that group 2 IFNs (IFNφ2 and IFNφ3) induce a different ISG subset. Determining this will
436 require more studies; however, since CHIKV induces both IFNφ1 and IFNφ3 while *crfb1&2*
437 morpholinos target receptors for both type I IFN groups, IFNφ3-only induced ISGs should
438 therefore be found among CHIKV-induced, IFNR-dependent, but non IFNφ1-induced genes.
439 Such genes (listed on Table S2, bottom) constituted about 30% of genes for which induction
440 by CHIKV was impacted in morphants – and only about 15% if one also excludes genes for
441 which induction by IFNφ1 is almost significant (Figure 3A and C). A previous report by López-
442 Muñoz *et al* suggests differences in ISG induction, notably in kinetics, by different IFNφs (19).
443
444 *2. Comparative and phylogenetic analysis of zebrafish ISGs*
445 Our comparative and phylogenetic approach led to a tentative reconstruction of the
446 ISG repertoire of the last common ancestor of teleosts and tetrapods (LCATT) that lived
447 ~450My ago and probably resembled the fossil osteichtyan *Ligulalepis* (40). To do so, we
448 looked for human (co-)orthologue(s) of all zebrafish ISGs identified in our analysis. Based on
449 available data compilations (29, 30), we then determined which one(s) of these human
450 orthologues were themselves induced by type I IFN. In such cases, we considered that they
451 most likely originated from an "ancestral" ISG, present in the LCATT. It is generally believed

452 that the type I IFN system emerged during the early evolution of jawed vertebrates, since
453 Chondrichtyans (rays, sharks and chimeras) but not Agnathans (lampreys and hagfish)
454 possess typical type I IFN genes (4, 41). Hence, it is important to note that the IFN system
455 had already evolved, expanded and standardized for more than 50My before our last
456 common ancestor with zebrafish.

457 Approximately half of what we defined as ancestral ISGs are represented by 1-to-1
458 orthologues in zebrafish and human (Table S5, top rows) – a situation of practical interest, as
459 the likelihood of conservation of gene function is highest in this case. These are either
460 isolated genes (e.g. *RSAD2*, *ISG15* or *cGAS*), or members of "old" families already stabilized in
461 the LCATT (e.g., *IRF7* and *IRF9*) (25). The situation is relatively similar for a few ancestral
462 genes such as *STAT1* or *SOCS1*, with one human orthologue and two zebrafish co-
463 orthologues that arose during the teleost-specific whole genome duplication and were
464 retained. In contrast, many other "young" families have clearly been subjected to further
465 duplication during later evolution of fish or tetrapods, leading to orthology groups
466 containing multiple ISG both in zebrafish and human, the most spectacular examples being
467 the *ISG12*, *IFIT* and *IFI44* families.

468 The frequency of orthology with a human gene is lower for ISGs (59%) than for the entire
469 genome (71%). This is probably a consequence of the stronger evolutionary pressure of
470 genes involved in the arms race with pathogens, as postulated by the Red Queen hypothesis
471 (42). Similar mechanisms also explain the frequent and extensive gene duplications, as well
472 as gene losses if some virus disappears, removing the corresponding selective pressure on a
473 given ISG. Possibly, a greater diversity of aquatic viruses could further favor ISG retention
474 and divergence after duplication, but few direct evidences are available.

475 In addition to the ancestral genes with true zebrafish and human orthologues, we added to
476 this list a few genes with a more complex history, with a human and zebrafish ISGs that
477 shared an ancestor at the basal vertebrate level (Table I and Table S5, bottom). These
478 ancestral genes must have been duplicated in the LCATT genome; the teleost and tetrapod
479 lineages then retained distinct paralogues. This comprises some genes whose evolutionary
480 history is extremely difficult to trace due to multiple copies and extensive polymorphism,
481 such as MHC class I genes. Here, only *mhc1zba* was found to be a zebrafish ISG, but this does
482 not necessarily imply that other zebrafish MHC class I genes are not ISGs, as they may have
483 been missed due to mapping issues; the strain we used (AB) is not the same as the one of
484 the reference genome (Tü), and strain-specific divergences are considerable between strains
485 for MHC class I, with deep evolutionary roots (43). Importantly, we did not define ancestral
486 ISGs for zebrafish-human ISG pairs that appeared to be related at first glance, but, upon
487 further analysis, were too distant; for example, zebrafish *vamp5* and human *VAMP8* are
488 both ISGs, but share their last common ancestor at the Opistokontha level, before the split
489 of fungi and animals, very long before the emergence of IFNs.
490 Nevertheless, the type I IFN system also includes very old genes that were already present in
491 basal metazoans. The RNAseL/OAS module is a good example of such cases, being found
492 across metazoans from mammals to sponges (44) - but lost in the fish branch. Another
493 striking example is the cGAS-STING module recently identified in cnidaria (45). The
494 implication of these genes in the antiviral immunity of basal branches of animals is unknown
495 but certainly worth investigating. The main models for invertebrate immunity are flies and
496 mosquitoes but they largely rely on RNAi mechanisms to contain viruses (3). Central
497 signaling modules of the vertebrate IFN system such as TLR/NFκB and JAK/STAT, are also

498 present in insects and in more distant metazoans, but they induce different set of genes with
499 other functions (46).
500 Additionally, a few important genes do not meet our criteria for "ancestral" ISG because
501 they are not typically inducible either in zebrafish or in human (Figure 5). For example, *irf3* is
502 an ISG in fish but not in human, while it is the reverse for *JAK2*. Hence, our list of ancestral
503 ISG is likely not complete, but it provides a core repertoire pointing to most fundamental
504 factors of the vertebrate innate antiviral arsenal.
505 A relatively large number of ISGs have no orthologue in the other lineage, such as human
506 APOBEC3, RNASEL, OAS, AIM2 (Figure 5). Similarly, many fish-specific ISG likely have been
507 co-opted by the IFN pathway during fish evolution. In this case, they do not have clear
508 orthologues in human and other tetrapods (as for finTRIMs and nlr-B30.2), or their
509 orthologue(s) have no link with the type I IFN system. The finTRIM family contains the
510 largest number of zebrafish ISGs of any family, ancestral or not. Interestingly, ISGs are found
511 only among the recently diversified, species-specific finTRIMs; the most basal members
512 (*ftr82-84*), well-conserved among fish, were not found here to be induced by IFNφ1 or by
513 CHIKV, consistent with previous studies (47). Nevertheless, *ftr83* appears to mediate
514 protection especially in the gill region by stimulating local *ifnphi1* expression (48). The
515 diversity and evolution under positive selection of the IFN-inducible finTRIMs evoke viral
516 recognition (21), yet their functions remain unclear.
517 The co-optation of new genes in the ISG repertoire may be operated quickly and in a group-
518 specific manner, by introduction of sequence motifs in the regulatory sequences, for
519 example via retroviral insertion (49, 50). However, we cannot exclude that these branch-
520 specific ISGs are in fact ancestral, but lost in one of the two lineages; this is the case for *Gig2*
521 genes, which are present in the Coelacanth genome as well as in fish, and thus were lost in

522 tetrapods. Thus, our repertoire of ancestral ISGs is underestimated because we cannot
523 include the lineage-specific losses.
524 Do ancestral ISGs identified here define a minimal but complete set of response elements
525 from recognition to elimination of invading viruses? Probably not, as this ancestral core
526 group of ISGs was backed up by more ISGs is any species, including the LCATT. For example,
527 the absence of the well-known OAS/RNAseL module genes in fish (and therefore in our list of
528 conserved ancestral ISGs) is puzzling, and one could predict that other fish genes have taken
529 over similar functions. Similarly, APOBEC3 genes are absent in fish, and maybe their RNA-
530 editing mechanisms are mediated by other genes, possibly by ADAR1.
531 Ancestral ISGs encode very diverse proteins in localization and function (Figure 4). We
532 provide an extended discussion of their classification below.
533
534 *4. Characterization of the IFN independent response to CHIKV infection*
535 Knowing the repertoire of ISG also offers the possibility to identify genes that are induced by
536 viral infection independently of the type I IFN pathway. While a subset of ISGs can be
537 induced via IFN-dependent and -independent pathways in human and fish - for example
538 *rsad2* (11). Thus, IRF3 dependent, type I IFN independent induction of many ISG by particular
539 viruses has been described (51).
540 However, about half of the genes upregulated by CHIKV were not induced by IFNϕ1
541 injection, and most were not affected by IFNR knockdown. Notably, three gene sets stand
542 out in this list: (1) components of the complement cascade, that are known to play a role in
543 antiviral defense; (2) cytokines including some CC and CXC chemokines as well as the type I
544 IFN themselves, which do not appear to be strongly auto/cross inducible and (3) many *btr*
545 and *ftr* TRIM E3 ligases as well as multiple proteasome components. Interestingly too, *irf1b*,

546 the zebrafish orthologue of IRF1 (a human ISG), is CHIKV-inducible, but not in a IFNR-
547 dependent manner – consistent with previous work (25) - and was not induced by IFNϕ1.
548 Many other genes of unknown function also share the same induction pattern, and would
549 certainly be worth investigating. A strong redundancy of antiviral pathways has certainly
550 been selected during evolution, since viruses have developed multiple strategies of immune
551 subversion.
552 Contrary to what was observed with upregulated genes, there was no overlap between gene
553 sets downregulated by IFNϕ1 and by CHIKV. This remarkable difference could be due to the
554 alternative inflammatory response induced by the virus besides type I IFNs, or to kinetic
555 differences.
556
557 *5. Classification of ancestral ISGs*
558 The ancestral ISG presented in Table I can be classified based on molecular functions:
559 sensors, transcription factors and other signal transduction factors, secreted factors,
560 enzymes including ubiquitination factors, and membrane receptors, which we discuss below.
561 The antiviral mechanisms described in human or in other mammalian systems also provide
562 hints about the likely conserved mode of action of these factors.
563 *Transcription factors.*
564 Many members of the list appear to have DNA binding capacity and may be classified as
565 transcription factors.
566 * Three IRF (3, 7, 9) and two STAT(1, 2) constitute fundamental components of the type I
567 pathway signaling, and were already present in the LCA of fish and mammals.

568 * BATF2 is a member of the AP-1/ATF family transcription factors that controls the
569 differentiation of immune cells and play key regulatory roles in immune responses. BATF2
570 promotes TLR7-induced Th1 responses (52).
571 * SP100 is a tumor suppressor and a major constituent of the PML bodies controlling
572 transcription and/or chromatin conformation.
573 * HELZ2 is a helicase that acts as a transcriptional coactivator for a number of nuclear
574 receptors, including AHR, a nuclear receptor regulating lipid metabolism and the
575 susceptibility to dengue virus (53).
576 * transcription coactivators. PARPs can act as transcriptional co-activators and potentiate
577 induction of many ISG (54). The multiple *ifi44* zebrafish genes counts 7 ISG among 19
578 members, but their two human co-orthologues are induced by type I IFN. Located in the
579 nucleus, IFI44 binds and blocks the HIV1 LTR promoter (55). However, the numerous
580 zebrafish *ifi44* probably have sub-functionalized, and mediate multiple antiviral mechanisms.
581
582 *Sensors and related genes.*
583 The helicases RIG-I, LPGP2 and IFIH1 (aka MDA5) stand as primary ISGs encoding viral
584 sensors. Besides, as a cytoplasmic helicase HELZ2 might also play a sensor role. In keeping
585 with this, TREX proteins have a 3'-to-5' DNA exonuclease activity that is important to block
586 the sting-dependent initiation of IFN responses by DNA fragments from endogenous
587 retroviruses and elements (56).
588
589 *Enzymes.*
590 Besides transcription factors, enzymes are the most important category of ancestral ISG.
591 They may play a role in signaling or have a direct antiviral activity.

592 * Poly-ADP-ribose polymerase (PARP) are involved in many cellular processes, from
593 regulation of chromatin conformation to transcription control, and several PARP also are
594 induced by infection and inflammation. PARP are represented by *parp9*, *parp*12 and *parp*14
595 among ancestral ISG. Strikingly, these three PARP are part of a nuclear complex, with the E3
596 ubiquitin ligase encoded by *dtx3l* that is also an ancestral ISG that promoting RNA Pol II
597 recruitment at IRF3-dependent promoters (54). Our data showing that key components of
598 this complex are part of the essential type I IFN system underscore its importance in the core
599 antiviral response. Besides, other activities of PARP may be involved in antiviral mechanisms;
600 for example PARP12 mediates ADP-ribosylation of Zika virus NS1 and NS3, leading to their
601 degradation by the proteasome (57). The ADP-ribo-hydrolase encoded by the Chikungunya
602 virus, that is required for its virulence, is another hint of the central importance of these
603 enzymes in antiviral defense (58).
604 * Several E3 Ubiquitin ligases were found among ancestral ISG, including *trim25, usp18,*
605 *rnf114, dtx3l*. The mechanisms through which they exert antiviral activity or regulate the
606 response are not fully resolved. The critical role of *trim25* in RIGI activation, and its capacity
607 of ISGylation (59) have been well documented in fish and mammals. *isg15*, an ubiquitin like
608 protein, is also an ancestral ISG playing a central role in the type I IFN pathway in fish and
609 mammals (6, 59) via multiple mechanisms.
610 * The pro-apoptotic caspase *casp7* possess type I IFN induced orthologues in zebrafish and
611 human. Interestingly, ancestral ISG also comprise *pmaip1* that promotes caspase activation
612 and apoptosis via modifications of the mitochondrial membrane, and *xiaf1*, a negative
613 regulator of members of inhibitor of apoptosis proteins. Taken together, these observations
614 indicate that the ancestral type I IFN system comprised a pro-apoptotic module.
615 * The pro-inflammatory caspase *casp1*, is also an ancestral ISG, as is *pycard* which encodes

616 ASC, the major scaffold protein of the canonical inflammasome. Induction of the
617 inflammasome is thus an ancestral property of the IFN response. Many upstream sensors of
618 the inflammasome are IFN-inducible, but they are generally divergent in the two lineages,
619 *nlrc5* being the only ancestral ISG.

620 * *Rsad2* (aka *viperin*) is an enzyme with a direct antiviral function, that catalyzes the
621 conversion of CTP to a completely new ribonucleoside, the 3'-deoxy-3',4'-didehydro-CTP
622 (ddhCTP) acting as a terminator of RNA synthesis (60). Interestingly, both ancestral ISG *rsad2*
623 and the nucleotide modifier *cpmk2* are located very close to each other in the genome in
624 fish as well as in mammals; likely forming a conserved functional antiviral unit.

625 * Adenosine deaminases acting on double-stranded RNA (ADARs) deaminate adenosine to
626 produce inosine in double-stranded RNA structures, regulating the inflammation induced by
627 such molecules (61). Accordingly, loss of function of *adar* in zebrafish larvae leads to brain
628 inflammation in a model of Aicardi-Goutières syndrome, suggesting a key regulatory role of
629 this gene during type I IFN response (62).

630 * Protein Kinase R (PKR, encoded by *eif2ak2*) is activated by dsRNA (and thus could have
631 been listed above as a sensor), leading phosphorylation of EIF2α, and to inhibition of protein
632 synthesis and viral replication. Many viruses encode PKR inhibitors of this cornerstone
633 antiviral factor that also affects transcription factors like IRF1, STATs, and NF-kappaB and
634 upregulates many genes including *β2microglobulin* and *isg15* (63). Interestingly, the other
635 ancestral ISG *epsti1* can activate PKR promoters and induce PKR-dependent genes in human
636 (64), questioning whether *pkr* and *epsti1* may have been functionally coupled from the
637 LCATT. Fish possess a lineage specific paralogue of PKR called PKZ, which detects Z-DNA (65).

638

639 *Secreted factors.*

640 * In humans and mice, ccl19 is implicated in lymphocyte migration and is important to
641 define compartments within lymphoid tissues. In rainbow trout, one of the six ccl19
642 paralogues present in the genome participate to antiviral immunity though promotion of
643 mucosal and central CD8+ T cell response (66).
644 * Some of the fish homologues of murine and human IFN inducible CXC chemokines – ie,
645 CXCL9-11, which bind CXCR3, a receptor expressed by various leukocyte, including some T
646 cells, macrophages, and dendritic cell subsets – are also up-regulated by IFNφ in zebrafish
647 larvae. These genes have been largely expanded in fish, and two lineages of CXCL11 have
648 been recently distinguished, both closely related to the mammalian CXCL9-11 (67, 68). The
649 up-regulated *cxcl11.3* (aka *cxc66* or *cxcl11ac*) identified in this work belongs to the lineage
650 1. The zebrafish has three *cxcr3* paralogues, and receptor-ligand binding, tested for three
651 other zebrafish cxcl11 ligands, does not follow ligand lineage (69), so the receptor(s) of this
652 ISG remains to be identified experimentally.
653 * Another soluble factor up-regulated by type I IFN and viral infection in fish is galectin9 (this
654 work, and (12)). In mammals, Galectin9 is involved in multiple mechanisms of antiviral
655 immunity. For example, it is a potent factor against HCMV because it blocks the entry of the
656 virus in target cells (70). Galectin9 can also regulates HIV transcription, and induces the
657 expression of the deaminase APOBEC3G, a potent antiviral factor (71). Besides, the galectin-
658 9 receptor TIM3 is implicated in the control of Th1 cells (72).
659
660 *Membrane proteins*
661 * While zebrafish and human *mhc class I* are not direct orthologues, *mhc class I* genes are in
662 the list, with *β2microglobulin* and the peptide transporters *tap-1* and *tap-2*, as well as
663 homologues of TAPBP and proteasome subunits, indicating that this pathway is a

664 fundamental component of the type I IFN system.

665 * Other important membrane proteins in the list are tetraspanins of the CD9 family, that
666 regulate degranulation of myeloid subsets and secretion of cytokines, hence constitute key
667 players in inflammation (73).

668 * Zebrafish possess eight *isg12* genes located in tandem, of which six were highly inducible
669 by IFNϕ1 and by CHIKV. Their human ISG orthologues IFI6 and IFI27 (aka ISG12A) are internal
670 membrane proteins stabilizing ER membrane and preventing the formation of flavivirus-
671 induced ER membrane invaginations (74) or destabilize mitochondrial membrane and
672 promote apoptosis (75). In fact, IFI27 can also recruit a E3 ubiquitin ligase and targets HCV
673 NS5 protein to degradation (76), illustrating the potential diversity of antiviral mechanisms
674 mediated by members of this family.

675 *APOL1 affects endocytosis and promotes an expansion of the lysosomal compartment,
676 favoring for example the degradation of the HIV-1 protein Vif (77).

677

678 *ISG with unknown functions or unknown antiviral mechanisms*

679 Even in human and mice, the basis of antiviral activity of certain ISGs remains completely
680 unknown. For example, the effects of PHF11, RNF114, or SAMD9 are elusive. In the latter, a
681 DNA/RNA-binding AlbA, a NTPase, and a OB domain with predicted RNA-binding properties
682 suggest a link with nucleic acid metabolism or sensing (78). These very old ISG with
683 counterparts found across Metazoa and even in procaryotes, are key restriction factors of
684 poxviruses (79).

685

686 *6. Conclusions*

687 Antiviral genes are well known to evolve very fast, as postulated by the Red Queen
688 hypothesis, under strong pressure from pathogens. This is indeed illustrated by the large
689 number of ISGs that are either fish- or mammal-specific. Nevertheless, our data define a
690 surprisingly stable set of core ISGs, that were apparently co-opted into the new IFN system
691 of early vertebrates about 500My ago, and have been maintained for the last 450My both in
692 fish and tetrapods. The full list of zebrafish ISG provides a powerful reference to characterize
693 the subtle interactions between viruses and the host response, including redundancy of
694 immune pathways and viral subversion mechanisms. It also constitutes a valuable resource
695 for the study of autoinflammatory disease using the emerging zebrafish model.
696
697
698

Acknowledgements

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

708 **References**

709

710 1. Schoggins, J. W., and C. M. Rice. 2011. Interferon-stimulated genes and their antiviral
711 effector functions. *Curr. Opin. Virol.* 1: 519–525.
712 2. Backes, S., R. A. Langlois, S. Schmid, A. Varble, J. V. Shim, D. Sachs, and B. R. TenOever.
713 2014. The Mammalian Response to Virus Infection Is Independent of Small RNA Silencing.
714 *Cell Rep.* 8: 114–125.
715 3. Guo, Z., Y. Li, and S. W. Ding. 2019. Small RNA-based antimicrobial immunity. *Nat. Rev.*
716 *Immunol.* 19: 31–44.
717 4. Secombes, C. J., and J. Zou. 2017. Evolution of interferons and interferon receptors. *Front.*
718 *Immunol.* 8: 209.
719 5. Zou, J., and C. J. Secombes. 2011. Teleost fish interferons and their role in immunity. *Dev.*
720 *Comp. Immunol.* 35: 1376–1387.
721 6. Langevin, C., E. Aleksejeva, G. Passoni, N. Palha, J. P. Levraud, and P. Boudinot. 2013. The
722 antiviral innate immune response in fish: Evolution and conservation of the IFN system. *J.*
723 *Mol. Biol.* 425: 4904–4920.
724 7. Boudinot, P., C. Langevin, C. J. Secombes, and J.-P. Levraud. 2016. The peculiar
725 characteristics of fish type I interferons. *Viruses* 8.
726 8. Robertsen, B. 2018. The role of type I interferons in innate and adaptive immunity against
727 viruses in Atlantic salmon. *Dev. Comp. Immunol.* 80: 41–52.
728 9. Aggad, D., M. Mazel, P. Boudinot, K. E. Mogensen, O. J. Hamming, R. Hartmann, S.
729 Kotenko, P. Herbomel, G. Lutfalla, and J.-P. Levraud. 2009. The two groups of zebrafish virus-
730 induced interferons signal via distinct receptors with specific and shared chains. *J. Immunol.*
731 183: 3924–3931.

732 10. Staeheli, P., Y. X. Yu, R. Grob, and O. Haller. 1989. A double-stranded RNA-inducible fish
733 gene homologous to the murine influenza virus resistance gene Mx. *Mol. Cell. Biol.* 9: 3117–
734 21.
735 11. Boudinot, P., P. Massin, M. Blanco, S. Riffault, and A. Benmansour. 1999. vig-1, a new fish
736 gene induced by the rhabdovirus glycoprotein, has a virus-induced homologue in humans
737 and shares conserved motifs with the MoaA family. *J. Virol.* 73: 1846–1852.
738 12. O'Farrell, C., N. Vaghefi, and M. Cantonnet. 2002. Survey of Transcript Expression in
739 Rainbow Trout Leukocytes Reveals a Major Contribution of Interferon-Responsive Genes in
740 the Early Response to a Rhabdovirus Infection. *J. Virol.* 76: 8040–8049.
741 13. Levraud, J.-P., P. Boudinot, I. Colin, A. Benmansour, N. Peyrieras, P. Herbomel, and G.
742 Lutfalla. 2007. Identification of the zebrafish IFN receptor: implications for the origin of the
743 vertebrate IFN system. *J. Immunol.* 178: 4385–4394.
744 14. Lutfalla, G., H. R. Crollius, N. Stange-Thomann, O. Jaillon, K. Mogensen, and D.
745 Monneron. 2003. Comparative genomic analysis reveals independent expansion of a lineage-
746 specific gene family in vertebrates: The class II cytokine receptors and their ligands in
747 mammals and fish. *BMC Genomics* 4: 29.
748 15. Altmann, S. M., M. T. Mellon, D. L. Distel, and C. H. Kim. 2003. Molecular and Functional
749 Analysis of an Interferon Gene from the Zebrafish, Danio rerio. *J. Virol.* 77: 1992–2002.
750 16. Robertsen, B., V. Bergan, T. Røkenes, R. Larsen, and A. Albuquerque. 2003. Atlantic
751 Salmon Interferon Genes: Cloning, Sequence Analysis, Expression, and Biological Activity. *J.*
752 *Interf. Cytokine Res.* 23: 601–612.
753 17. Purcell, M. K., K. M. Nichols, J. R. Winton, G. Kurath, G. H. Thorgaard, P. Wheeler, J. D.
754 Hansen, R. P. Herwig, and L. K. Park. 2006. Comprehensive gene expression profiling
755 following DNA vaccination of rainbow trout against infectious hematopoietic necrosis virus.

756 *Mol. Immunol.* 43: 2089–2106.

757 18. Martin, S. A. M., J. B. Taggart, P. Seear, J. E. Bron, R. Talbot, A. J. Teale, G. E. Sweeney, B.
758 Høyheim, D. F. Houlihan, D. R. Tocher, J. Zou, and C. J. Secombes. 2007. Interferon type I and
759 type II responses in an Atlantic salmon (Salmo salar) SHK-1 cell line by the salmon
760 TRAITS/SGP microarray. *Physiol. Genomics* 32: 33–44.

761 19. López-Muñoz, A., F. J. Roca, J. Meseguer, and V. Mulero. 2009. New insights into the
762 evolution of IFNs: zebrafish group II IFNs induce a rapid and transient expression of IFN-
763 dependent genes and display powerful antiviral activities. *J. Immunol.* 182: 3440–3449.

764 20. Rothenburg, S., N. Deigendesch, M. Dey, T. E. Dever, and L. Tazi. 2008. Double-stranded
765 RNA-activated protein kinase PKR of fishes and amphibians: varying the number of double-
766 stranded RNA binding domains and lineage-specific duplications. *BMC Biol.* 6: 12.

767 21. van der Aa, L. M., J.-P. Levraud, M. Yahmi, E. Lauret, V. Briolat, P. Herbomel, A.
768 Benmansour, and P. Boudinot. 2009. A large new subset of TRIM genes highly diversified by
769 duplication and positive selection in teleost fish. *BMC Biol.* 7: 7.

770 22. Zhang, Y. B., and J. F. Gui. 2004. Identification and expression analysis of two IFN-
771 inducible genes in crucian carp (Carassius auratus L.). *Gene* 60: 1–9.

772 23. Zhang, Y. B., T. K. Liu, J. Jiang, J. Shi, Y. Liu, S. Li, and J. F. Gui. 2013. Identification of a
773 Novel Gig2 Gene Family Specific to Non-Amniote Vertebrates. *PLoS One* 8: e60588.

774 24. Boudinot, P., S. Salhi, M. Blanco, and A. Benmansour. 2001. Viral haemorrhagic
775 septicaemia virus induces vig -2, a new interferon-responsive gene in rainbow trout. *Fish*
776 *Shellfish Immunol.* 11: 383–397.

777 25. Briolat, V., L. Jouneau, R. Carvalho, N. Palha, C. Langevin, P. Herbomel, O. Schwartz, H. P.
778 Spaink, J.-P. Levraud, and P. Boudinot. 2014. Contrasted Innate Responses to Two Viruses in
779 Zebrafish: Insights into the Ancestral Repertoire of Vertebrate IFN-Stimulated Genes. *J.*

780 *Immunol.* 192: 4328–41.

781 26. Ludwig, M., N. Palha, C. Torhy, V. Briolat, E. Colucci-Guyon, M. Brémont, P. Herbomel, P.
782 Boudinot, and J. P. Levraud. 2011. Whole-body analysis of a viral infection: Vascular
783 endothelium is a primary target of Infectious Hematopoietic Necrosis Virus in zebrafish
784 larvae. *PLoS Pathog.* 7: e1001269.

785 27. Palha, N., F. Guivel-Benhassine, V. Briolat, G. Lutfalla, M. Sourisseau, F. Ellett, C. H. Wang,
786 G. J. Lieschke, P. Herbomel, O. Schwartz, and J. P. Levraud. 2013. Real-Time Whole-Body
787 Visualization of Chikungunya Virus Infection and Host Interferon Response in Zebrafish. *PLoS*
788 *Pathog.* 9: e1003619.

789 28. Howe, K., M. D. Clark, C. F. Torroja, J. Torrance, C. Berthelot, M. Muffato, J. E. Collins, S.
790 Humphray, K. McLaren, L. Matthews, S. McLaren, I. Sealy, M. Caccamo, C. Churcher, C. Scott,
791 J. C. Barrett, R. Koch, G.-J. Rauch, S. White, W. Chow, B. Kilian, L. T. Quintais, J. a Guerra-
792 Assunção, Y. Zhou, Y. Gu, J. Yen, J.-H. Vogel, T. Eyre, S. Redmond, R. Banerjee, J. Chi, B. Fu, E.
793 Langley, S. F. Maguire, G. K. Laird, D. Lloyd, E. Kenyon, S. Donaldson, H. Sehra, J. Almeida-
794 King, J. Loveland, S. Trevanion, M. Jones, M. Quail, D. Willey, A. Hunt, J. Burton, S. Sims, K.
795 McLay, B. Plumb, J. Davis, C. Clee, K. Oliver, R. Clark, C. Riddle, D. Elliot, D. Eliott, G.
796 Threadgold, G. Harden, D. Ware, S. Begum, B. Mortimore, B. Mortimer, G. Kerry, P. Heath, B.
797 Phillimore, A. Tracey, N. Corby, M. Dunn, C. Johnson, J. Wood, S. Clark, S. Pelan, G. Griffiths,
798 M. Smith, R. Glithero, P. Howden, N. Barker, C. Lloyd, C. Stevens, J. Harley, K. Holt, G.
799 Panagiotidis, J. Lovell, H. Beasley, C. Henderson, D. Gordon, K. Auger, D. Wright, J. Collins, C.
800 Raisen, L. Dyer, K. Leung, L. Robertson, K. Ambridge, D. Leongamornlert, S. McGuire, R.
801 Gilderthorp, C. Griffiths, D. Manthravadi, S. Nichol, G. Barker, S. Whitehead, M. Kay, J.
802 Brown, C. Murnane, E. Gray, M. Humphries, N. Sycamore, D. Barker, D. Saunders, J. Wallis, A.
803 Babbage, S. Hammond, M. Mashreghi-Mohammadi, L. Barr, S. Martin, P. Wray, A. Ellington,

804 N. Matthews, M. Ellwood, R. Woodmansey, G. Clark, J. D. Cooper, J. Cooper, A. Tromans, D.
805 Grafham, C. Skuce, R. Pandian, R. Andrews, E. Harrison, A. Kimberley, J. Garnett, N. Fosker,
806 R. Hall, P. Garner, D. Kelly, C. Bird, S. Palmer, I. Gehring, A. Berger, C. M. Dooley, Z. Ersan-
807 Ürün, C. Eser, H. Geiger, M. Geisler, L. Karotki, A. Kirn, J. Konantz, M. Konantz, M.
808 Oberländer, S. Rudolph-Geiger, M. Teucke, C. Lanz, G. Raddatz, K. Osoegawa, B. Zhu, A.
809 Rapp, S. Widaa, C. Langford, F. Yang, S. C. Schuster, N. P. Carter, J. Harrow, Z. Ning, J.
810 Herrero, S. M. J. Searle, A. Enright, R. Geisler, R. H. a Plasterk, C. Lee, M. Westerfield, P. J. de
811 Jong, L. I. Zon, J. H. Postlethwait, C. Nüsslein-Volhard, T. J. P. Hubbard, H. Roest Crollius, J.
812 Rogers, and D. L. Stemple. 2013. The zebrafish reference genome sequence and its
813 relationship to the human genome. *Nature* 496: 498–503.
814 29. Schoggins, J. W., S. J. Wilson, M. Panis, M. Y. Murphy, C. T. Jones, P. Bieniasz, and C. M.
815 Rice. 2011. A diverse range of gene products are effectors of the type I interferon antiviral
816 response. *Nature* 472: 481–5.
817 30. Rusinova, I., S. Forster, S. Yu, A. Kannan, M. Masse, H. Cumming, R. Chapman, and P. J.
818 Hertzog. 2013. INTERFEROME v2.0: An updated database of annotated interferon-regulated
819 genes. *Nucleic Acids Res.* 41: 1040–1046.
820 31. Levraud, J.-P., E. Colucci-Guyon, M. J. Redd, G. Lutfalla, and P. Herbomel. 2008. In vivo
821 analysis of zebrafish innate immunity. *Methods Mol. Biol.* 415: 337–363.
822 32. Biacchesi, S., M. LeBerre, A. Lamoureux, Y. Louise, E. Lauret, P. Boudinot, and M.
823 Brémont. 2009. Mitochondrial antiviral signaling protein plays a major role in induction of
824 the fish innate immune response against RNA and DNA viruses. *J. Virol.* 83: 7815–27.
825 33. Letunic, I., and P. Bork. 2018. 20 years of the SMART protein domain annotation
826 resource. *Nucleic Acids Res.* 46: D493-496.
827 34. Boudinot, P., L. M. van der Aa, L. Jouneau, L. Pasquier, P. Pontarotti, V. Briolat, A.

828 Benmansour, and J. P. Levraud. 2011. Origin and evolution of TRIM proteins: New insights
829 from the complete TRIM repertoire of zebrafish and pufferfish. *PLoS One* 6: e22022.
830 35. Howe, K., P. H. Schiffer, J. Zielinski, T. Wiehe, G. K. Laird, J. C. Marioni, O. Soylemez, F.
831 Kondrashov, and M. Leptin. 2016. Structure and evolutionary history of a large family of NLR
832 proteins in the zebrafish. *Open Biol.* 6: 160009.
833 36. Stein, C., M. Caccamo, G. Laird, and M. Leptin. 2007. Conservation and divergence of
834 gene families encoding components of innate immune response systems in zebrafish.
835 *Genome Biol.* 8: R251.
836 37. Laing, K. J., M. K. Purcell, J. R. Winton, and J. D. Hansen. 2008. A genomic view of the
837 NOD-like receptor family in teleost fish: Identification of a novel NLR subfamily in zebrafish.
838 *BMC Evol. Biol.* 8: 42.
839 38. Zhang, F., Y. Tao, Z. Zhang, X. Guo, P. An, Y. Shen, Q. Wu, Y. Yu, and F. Wang. 2012.
840 Metalloreductase steap3 coordinates the regulation of iron homeostasis and inflammatory
841 responses. *Haematologica* 97: 1826–1835.
842 39. Zou, J., C. Tafalla, J. Truckle, and C. J. Secombes. 2007. Identification of a second group of
843 type I IFNs in fish sheds light on IFN evolution in vertebrates. *J. Immunol.* 179: 3859–3871.
844 40. Clement, A. M., B. King, S. Giles, B. Choo, P. E. Ahlberg, G. C. Young, and J. A. Long. 2018.
845 Neurocranial anatomy of an enigmatic Early Devonian fish sheds light on early osteichthyan
846 evolution. *Elife* 7: e34349.
847 41. Venkatesh, B., A. P. Lee, V. Ravi, A. K. Maurya, M. M. Lian, J. B. Swann, Y. Ohta, M. F.
848 Flajnik, Y. Sutoh, M. Kasahara, S. Hoon, V. Gangu, S. W. Roy, M. Irimia, V. Korzh, I.
849 Kondrychyn, Z. W. Lim, B. H. Tay, S. Tohari, K. W. Kong, S. Ho, B. Lorente-Galdos, J. Quilez, T.
850 Marques-Bonet, B. J. Raney, P. W. Ingham, A. Tay, L. W. Hillier, P. Minx, T. Boehm, R. K.
851 Wilson, S. Brenner, and W. C. Warren. 2014. Elephant shark genome provides unique

852 insights into gnathostome evolution. *Nature* 505: 174–179.

853 42. tenOever, B. R. 2016. The Evolution of Antiviral Defense Systems. *Cell Host Microbe* 19:

854 142–149.

855 43. McConnell, S. C., K. M. Hernandez, D. J. Wcisel, R. N. Kettleborough, D. L. Stemple, J. A.

856 Yoder, J. Andrade, and J. L. O. de Jong. 2016. Alternative haplotypes of antigen processing

857 genes in zebrafish diverged early in vertebrate evolution. *Proc. Natl. Acad. Sci.* 113: E5014–

858 E5023.

859 44. Päri, M., A. Kuusksalu, A. Lopp, T. Reintamm, J. Justesen, and M. Kelve. 2007. Expression

860 and characterization of recombinant 2′,5′- oligoadenylate synthetase from the marine

861 sponge Geodia cydonium. *FEBS J.* 274: 3462–3474.

862 45. Kranzusch, P. J., S. C. Wilson, A. S. Y. Lee, J. M. Berger, J. A. Doudna, and R. E. Vance.

863 2015. Ancient Origin of cGAS-STING Reveals Mechanism of Universal 2′,3′ cGAMP Signaling.

864 *Mol. Cell* 59: 891–903.

865 46. Liongue, C., R. Sertori, and A. C. Ward. 2016. Evolution of Cytokine Receptor Signaling. *J.*

866 *Immunol.* 197: 11–18.

867 47. Langevin, C., J. P. Levraud, and P. Boudinot. 2019. Fish antiviral tripartite motif (TRIM)

868 proteins. *Fish Shellfish Immunol.* 86: 724–733.

869 48. Langevin, C., E. Aleksejeva, A. Houel, V. Briolat, C. Torhy, A. Lunazzi, J.-P. Levraud, and P.

870 Boudinot. 2017. FTR83, a member of the large fish-specific finTRIM family, triggers IFN

871 pathway and counters viral infection. *Front. Immunol.* 8: 617.

872 49. Chuong, E. B., N. C. Elde, and C. Feschotte. 2016. Regulatory evolution of innate

873 immunity through co-option of endogenous retroviruses. *Science (80-.).* 351: 1083–7.

874 50. Arnold, M. L., and K. Kunte. 2017. Adaptive Genetic Exchange: A Tangled History of

875 Admixture and Evolutionary Innovation. *Trends Ecol. Evol.* 32: 601–611.

876 51. Ashley, C. L., A. Abendroth, B. P. McSharry, and B. Slobedman. 2019. Interferon-
877 Independent Upregulation of Interferon-Stimulated Genes during Human Cytomegalovirus
878 Infection is Dependent on IRF3 Expression. *Viruses* 11: E246.
879 52. Kanemaru, H., F. Yamane, K. Fukushima, T. Matsuki, T. Kawasaki, I. Ebina, K. Kuniyoshi, H.
880 Tanaka, K. Maruyama, K. Maeda, T. Satoh, and S. Akira. 2017. Antitumor effect of Batf2
881 through IL-12 p40 up-regulation in tumor-associated macrophages. *Proc. Natl. Acad. Sci.*
882 114: E7331–E7340.
883 53. Fusco, D. N., H. Pratt, S. Kandilas, S. S. Y. Cheon, W. Lin, D. A. Cronkite, M. Basavappa, K.
884 L. Jeffrey, A. Anselmo, R. Sadreyev, C. Yapp, X. Shi, J. F. O'Sullivan, R. E. Gerszten, T. Tomaru,
885 S. Yoshino, T. Satoh, and R. T. Chung. 2017. HELZ2 is an IFN effector mediating suppression
886 of dengue virus. *Front. Microbiol.* 8: 240.
887 54. Caprara, G., E. Prosperini, V. Piccolo, G. Sigismondo, A. Melacarne, A. Cuomo, M.
888 Boothby, M. Rescigno, T. Bonaldi, and G. Natoli. 2018. PARP14 Controls the Nuclear
889 Accumulation of a Subset of Type I IFN–Inducible Proteins. *J. Immunol.* 200: 2439–2454.
890 55. Power, D., N. Santoso, M. Dieringer, J. Yu, H. Huang, S. Simpson, I. Seth, H. Miao, and J.
891 Zhu. 2015. IFI44 suppresses HIV-1 LTR promoter activity and facilitates its latency. *Virology*
892 481: 142–50.
893 56. Yan, N. 2017. Immune Diseases Associated with TREX1 and STING Dysfunction. *J. Interf.*
894 *Cytokine Res*. 37: 198–206.
895 57. Li, L., H. Zhao, P. Liu, C. Li, N. Quanquin, X. Ji, N. Sun, P. Du, C. F. Qin, N. Lu, and G. Cheng.
896 2018. PARP12 suppresses Zika virus infection through PARP-dependent degradation of NS1
897 and NS3 viral proteins. *Sci. Signal.* 11: eaas9332.
898 58. McPherson, R. L., R. Abraham, E. Sreekumar, S.-E. Ong, S.-J. Cheng, V. K. Baxter, H. A. V.
899 Kistemaker, D. V. Filippov, D. E. Griffin, and A. K. L. Leung. 2017. ADP-ribosylhydrolase

900 activity of Chikungunya virus macrodomain is critical for virus replication and virulence. *Proc.*
901 *Natl. Acad. Sci.* 114: 16666–1671.
902 59. Langevin, C., L. M. van der Aa, a Houel, C. Torhy, V. Briolat, a Lunazzi, a Harmache, M.
903 Bremont, J.-P. Levraud, and P. Boudinot. 2013. Zebrafish ISG15 exerts a strong antiviral
904 activity against RNA and DNA viruses and regulates the interferon response. *J. Virol.* 87:
905 10025–36.
906 60. Gizzi, A. S., T. L. Grove, J. J. Arnold, J. Jose, R. K. Jangra, S. J. Garforth, Q. Du, S. M. Cahill,
907 N. G. Dulyaninova, J. D. Love, K. Chandran, A. R. Bresnick, C. E. Cameron, and S. C. Almo.
908 2018. A naturally occurring antiviral ribonucleotide encoded by the human genome. *Nature*
909 558: 610–614.
910 61. Samuel, C. E. 2019. Adenosine deaminase acting on RNA (ADAR1), a suppressor of
911 double-stranded RNA-triggered innate immune responses. *J. Biol. Chem.* 294: 1710–1720.
912 62. Kasher, P. R., E. M. Jenkinson, V. Briolat, D. Gent, C. Morrissey, L. A. H. Zeef, G. I. Rice, J.-
913 P. Levraud, and Y. J. Crow. 2015. Characterization of samhd1 morphant zebrafish
914 recapitulates features of the human type i interferonopathy Aicardi-Goutières syndrome. *J.*
915 *Immunol.* 194.
916 63. Guerra, S., L. A. López-Fernández, M. A. García, A. Zaballos, and M. Esteban. 2006.
917 Human gene profiling in response to the active protein kinase, interferon-induced
918 serine/threonine protein kinase (PKR), in infected cells: Involvement of the transcription
919 factor ATF-3 in PKR-induced apoptosis. *J. Biol. Chem.* 281: 18734–18745.
920 64. Meng, X., D. Yang, R. Yu, and H. Zhu. 2015. EPSTI1 is involved in IL-28A-mediated
921 inhibition of HCV Infection. *Mediators Inflamm.* 2015: 716315.
922 65. Rothenburg, S., N. Deigendesch, K. Dittmar, F. Koch-Nolte, F. Haag, K. Lowenhaupt, and
923 A. Rich. 2005. A PKR-like eukaryotic initiation factor 2 kinase from zebrafish contains Z-DNA

924 binding domains instead of dsRNA binding domains. *Proc. Natl. Acad. Sci.* 102: 1602–1607.
925 66. Sepahi, A., L. Tacchi, E. Casadei, F. Takizawa, S. E. LaPatra, and I. Salinas. 2017. CK12a, a
926 CCL19-like Chemokine That Orchestrates both Nasal and Systemic Antiviral Immune
927 Responses in Rainbow Trout. *J. Immunol.* 199: 3900–3913.
928 67. Nomiyama, H., N. Osada, and O. Yoshie. 2013. Systematic classification of vertebrate
929 chemokines based on conserved synteny and evolutionary history. *Genes to Cells* 18: 1–16.
930 68. Chen, J., Q. Xu, T. Wang, B. Collet, Y. Corripio-Miyar, S. Bird, P. Xie, P. Nie, C. J. Secombes,
931 and J. Zou. 2013. Phylogenetic analysis of vertebrate CXC chemokines reveals novel lineage
932 specific groups in teleost fish. *Dev. Comp. Immunol.* 41: 137–152.
933 69. Torraca, V., C. Cui, R. Boland, J.-P. Bebelman, A. M. van der Sar, M. J. Smit, M. Siderius, H.
934 P. Spaink, and A. H. Meijer. 2015. The CXCR3-CXCL11 signaling axis mediates macrophage
935 recruitment and dissemination of mycobacterial infection. *Dis. Model. Mech.* 7: 253–269.
936 70. Machala, E. A., S. Avdic, L. Stern, D. M. Zajonc, C. A. Benedict, E. Blyth, D. J. Gottlieb, A.
937 Abendroth, B. P. McSharry, and B. Slobedman. 2018. Restriction of Human Cytomegalovirus
938 Infection by Galectin-9. *J. Virol.* 93: e01746-18.
939 71. Abdel-Mohsen, M., L. Chavez, R. Tandon, G. M. Chew, X. Deng, A. Danesh, S. Keating, M.
940 Lanteri, M. L. Samuels, R. Hoh, J. B. Sacha, P. J. Norris, T. Niki, C. M. Shikuma, M. Hirashima,
941 S. G. Deeks, L. C. Ndhlovu, and S. K. Pillai. 2016. Human Galectin-9 Is a Potent Mediator of
942 HIV Transcription and Reactivation. *PLoS Pathog.* 12: e1005677.
943 72. Zhu, C., A. C. Anderson, A. Schubart, H. Xiong, J. Imitola, S. J. Khoury, X. X. Zheng, T. B.
944 Strom, and V. K. Kuchroo. 2005. The Tim-3 ligand galectin-9 negatively regulates T helper
945 type 1 immunity. *Nat. Immunol.* 6: 1245–1252.
946 73. Brosseau, C., L. Colas, A. Magnan, and S. Brouard. 2018. CD9 tetraspanin: A new pathway
947 for the regulation of inflammation? *Front. Immunol.* 9: 2316.

948 74. Richardson, R. B., M. B. Ohlson, J. L. Eitson, A. Kumar, M. B. McDougal, I. N. Boys, K. B.
949 Mar, P. C. De La Cruz-Rivera, C. Douglas, G. Konopka, C. Xing, and J. W. Schoggins. 2018. A
950 CRISPR screen identifies IFI6 as an ER-resident interferon effector that blocks flavivirus
951 replication. *Nat. Microbiol.* 3: 1214–1223.
952 75. Cheriyath, V., D. W. Leaman, and E. C. Borden. 2011. Emerging Roles of FAM14 Family
953 Members (G1P3/ISG 6–16 and ISG12/IFI27) in Innate Immunity and Cancer. *J. Interf.*
954 *Cytokine Res*. 31: 173–181.
955 76. Xue, B., D. Yang, J. Wang, Y. Xu, X. Wang, Y. Qin, R. Tian, S. Chen, Q. Xie, N. Liu, and H.
956 Zhu. 2016. ISG12a Restricts Hepatitis C Virus Infection through the Ubiquitination-
957 Dependent Degradation Pathway. *J. Virol.* 90: 6832–6845.
958 77. Taylor, H. E., A. K. Khatua, and W. Popik. 2014. The Innate Immune Factor Apolipoprotein
959 L1 Restricts HIV-1 Infection. *J. Virol.* 88: 592–603.
960 78. Mekhedov, S. L., K. S. Makarova, and E. V. Koonin. 2017. The complex domain
961 architecture of SAMD9 family proteins, predicted STAND-like NTPases, suggests new links to
962 inflammation and apoptosis. *Biol. Direct* 12: 13.
963 79. Meng, X., F. Zhang, B. Yan, C. Si, H. Honda, A. Nagamachi, L. Z. Sun, and Y. Xiang. 2018. A
964 paralogous pair of mammalian host restriction factors form a critical host barrier against
965 poxvirus infection. *PLoS Pathog.* 14: e1006884.
966
967

968 **Figure legends**

969

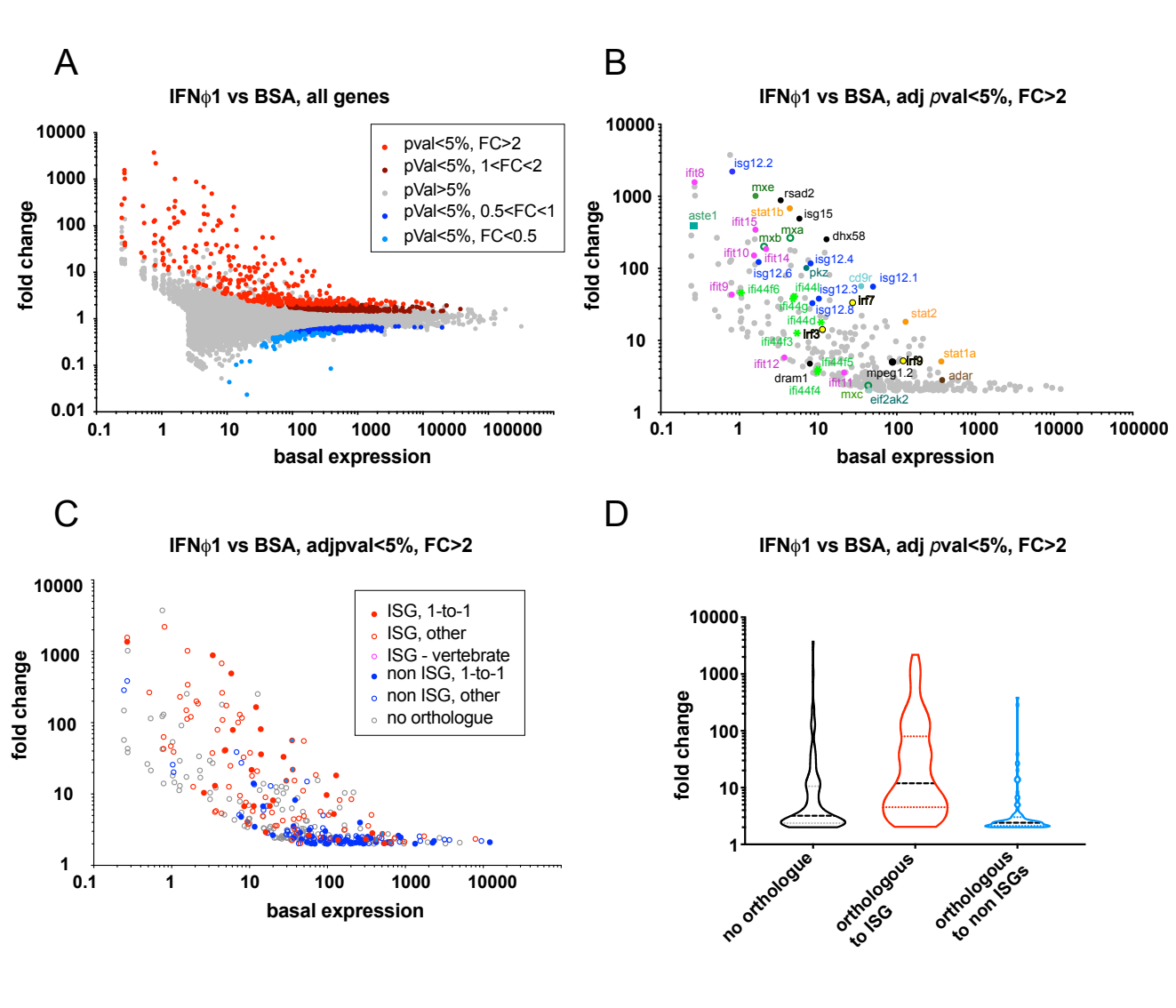
970 **Figure 1. Zebrafish larva transcriptional response to type I IFN** (A) Fold change (FC)/basal
971 expression representation with all genes detected in the analysis. Basal expression was the
972 average of read numbers mapped onto a given gene in control (BSA-injected) larvae. FC is
973 the ratio of reads numbers in IFNφ1-injected larvae divided by basal expression. pVal
974 corresponds to adjusted p-Value. (B) FC/basal expression representation, limited to ISGs
975 identified in IFNφ1-injected larvae (e.g. red dots in panel A); key genes commented in the
976 text are indicated with gene families identified by different colors/symbols. (C) FC/basal
977 expression representation for zebrafish ISGs represented according to the type of orthology
978 with human genes, and if human orthologues include at least one human ISG. 1-to-1: single
979 zebrafish gene orthologous to single human gene. Other: other orthology relationships
980 between zebrafish and human genes (many-to-many, 1-to-many, many-to-1). Vertebrate:
981 zebrafish gene sharing a common ancestor with a human ISG at the basal or jawed
982 vertebrate level, but not orthologous. No orthologue: no orthology relationship with a
983 human gene, and no known common ancestor with human ISG at vertebrate level. Non-ISG:
984 orthologous human gene(s) do not include any ISG. (D) Fold change distribution for zebrafish
985 ISG with different types of orthology relationship to human genes, using the same color code
986 as in C but pooling genes with single and multiple human orthologues. Zebrafish ISGs with a
987 human ISG within a vertebrate-level orthology groupes were not included in this analysis.

988

989 **Figure 2. Main families of fish-specific ISG and their domain organisation.** Each panel show
990 the typical domain organization of a family of fish-specific ISGs (e.g., with no human
991 orthologue), as determined by SMART analysis (http://smart.embl-heidelberg.de/). The

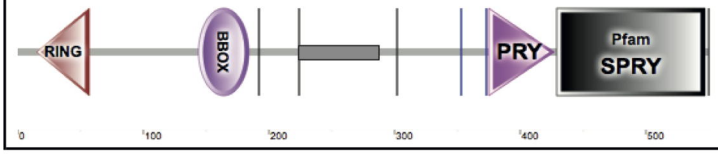
992 accession numbers of IFNφ1-induced genes (FC>2, pVal<5%) and the fraction they represent
993 within the family are indicated. Vertical lines represent exon boudaries. Besides named
994 domains, boxes represent coiled-coil regions (grey), low complexity regsions (striped),
995 transmembrane domains (black), and leader peptides (white).
996
997 **Figure 3. Comparison of ISG repertoires identified by IFNφ1 induction and by CHIKV**
998 **infection.** (A) Venn analysis of genes up- or down-regulated in zebrafish larva by IFNφ1
999 treatment (IFNφ1), CHIKV infection (ChikV), and CHIKV infection in the context of type I IFN
1000 receptor knock-down ("MoCrfb"). ISG are identified either by their responsiveness to IFNφ1
1001 or by a CHIKV induction abolished in crfb1+2 morphants (venn diagram at the top). Genes
1002 induced by CHIKV infection in an IFNφ1 independent way are analysed in the Venn diagram
1003 at the bottom of the panel. (B) FC/FC representation of transcriptome response to IFNφ1
1004 and CHIKV. Color code identifies genes significantly induced by IFNφ1 and/or CHIKV, and
1005 genes which are significantly induced in one condition (FC>2, adj p val< 0.05) while almost
1006 significantly induced in the other ("quasi Sig"; thresholds FC>1.5 and/or adj p value < 0.2).
1007 (C) Venn analysis of genes significantly induced by IFNφ1 and/or CHIKV. Gene subsets
1008 corresponding to genes induced quasi-significantly is represented within dotted lines.
1009
1010 **Figure 4. Graphic overview of the ancestral ISG repertoire**, organized by functional
1011 modules.
1012
1013 **Figure 5. Evolution of the ISG repertoire** since the LCATT (last common ancestor of
1014 Tetrapods and Teleosts). Genes names are given as examples, with no attempt to be

exhaustive.


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
ifi44
isg12.3
irf7
ifi44g
isg12.8
stat2
ifi44d
irf3
ifi44f3
stat1a
ifit12
ifi44f5
irf9
mpeg1.2
adar
dram1
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

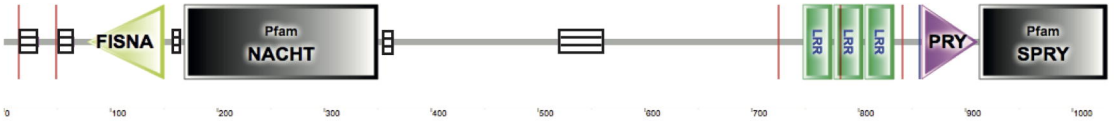
finTRIMs 21 ISGs out of ~80 paralogues

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



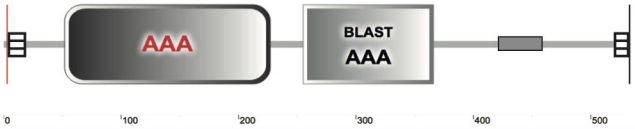
fish-specific NLRs 6 ISGs out of ~350 paralogues (several subfamilies)

ENSDARG00000102440 ENSDARG00000077488 ENSDARG00000077673 ENSDARG00000090033 ENSDARG00000078912 ENSDARG00000087597



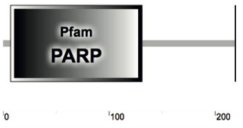
septin-related 2 ISGs out of 15 paralogues

ENSDARG00000053761 ENSDARG00000078093



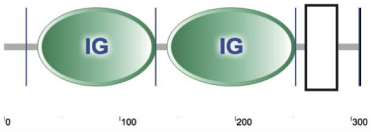
Gig2 family 2 ISGs of 11 paralogues

ENSDARG00000088260 ENSDARG00000086903



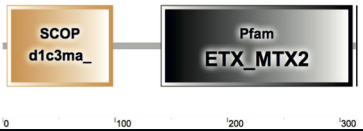
Fish two-Ig domain 3 ISGs out of ~140 paralogues

ENSDARG00000099577 ENSDARG00000095643 ENSDARG00000071657



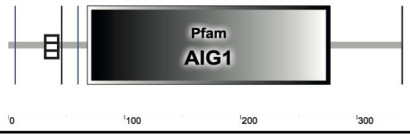
lectin 2 ISGs out of 14 paralogues

ENSDARG00000087318 ENSDARG00000039099



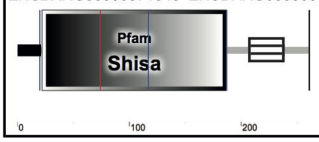
AIG1 domain 2 ISGs out of 4 paralogues

ENSDARG00000018587 ENSDARG00000078182



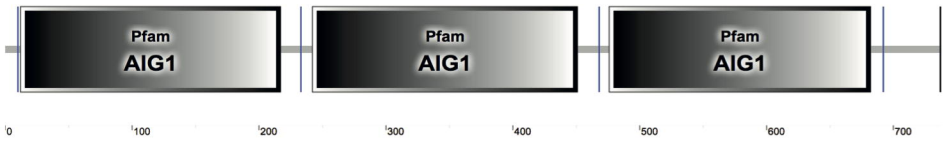
Shisa-related 2 ISGs out of 2 paralogues

ENSDARG00000071543 ENSDARG00000013681



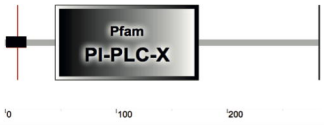
multi AIG1 domains 2 ISGs out of 10 paralogues

ENSDARG00000011581 ENSDARG00000101671



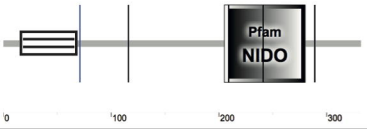
phospholipase 2 ISGs out of 5 paralogues

ENSDARG00000075622 ENSDARG00000099715



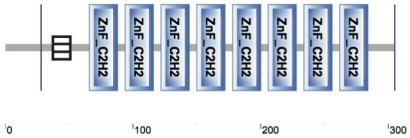
NIDO domain 2 ISGs of ouf 6 paralogues

ENSDARG00000103973 ENSDARG00000101930



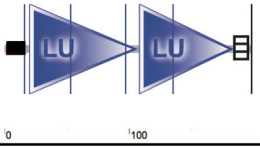
tandem ZNFs 2 ISGs out of ~260 genes

ENSDARG00000091243 ENSDARG00000099872



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

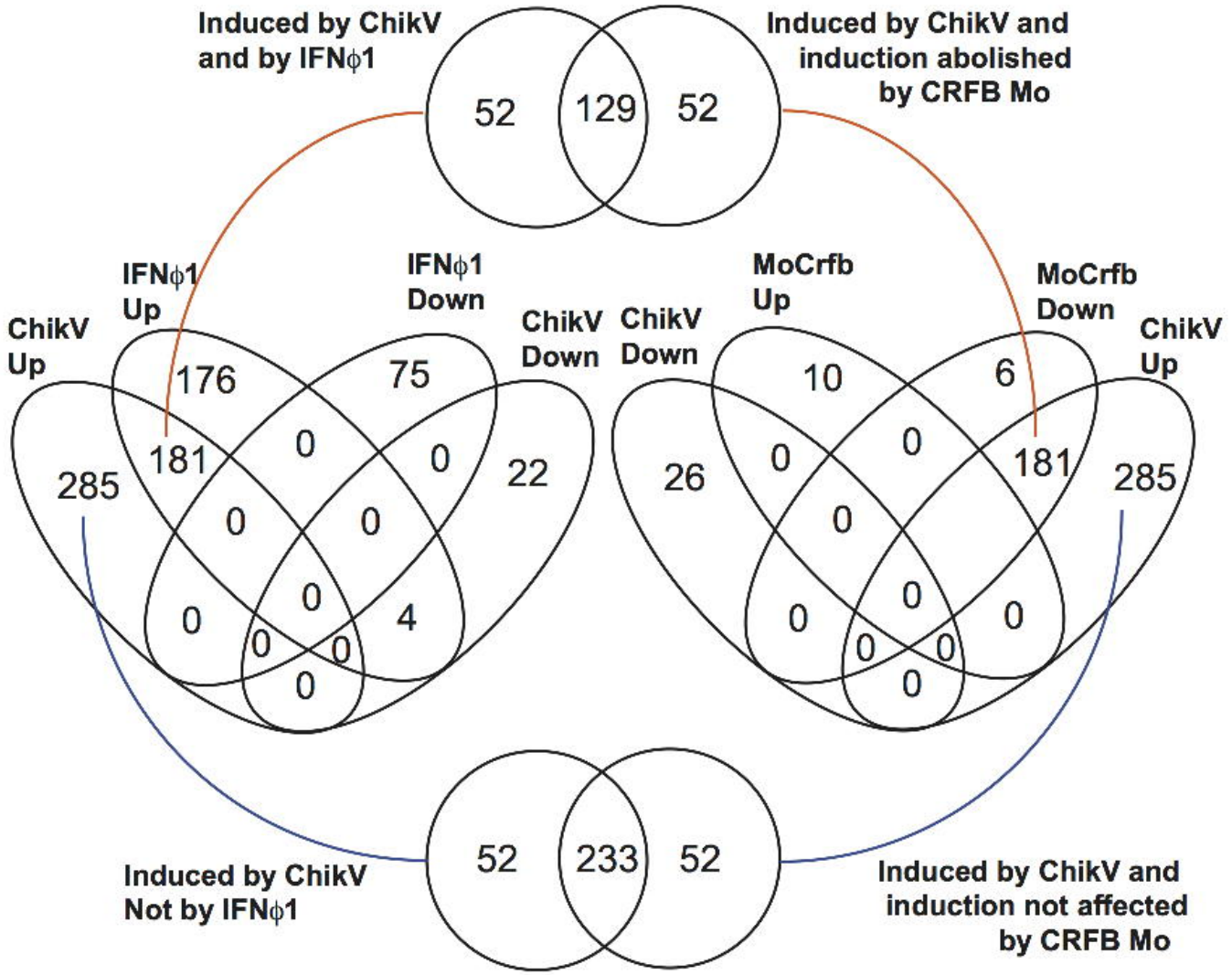
ENSDARG00000104775 ENSDARG00000089123

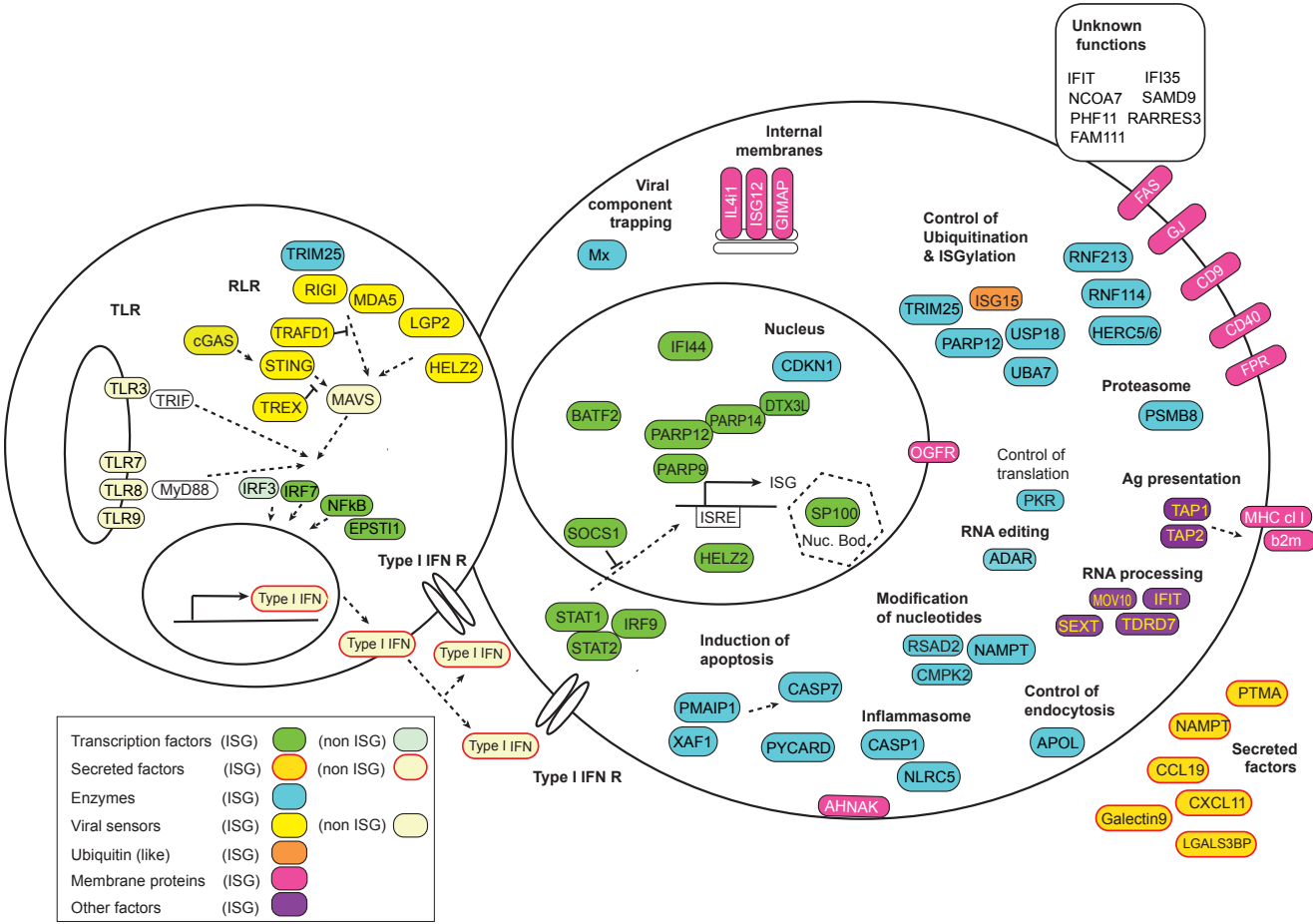


large, URGCP-related 3 ISGs out of 15 paralogues

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
176
181
75
22
0
0
0
0
0
0
0
0
4
0
0
0
0
0
0
0
0


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

Ancestral ISG maintained in teleosts as

single copy:

rsad2, trim25, isg15, apol

or multiple paralogues:

ifi44, mx, isg2,herc5/6, sp100

Conserved ISG

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 ISG in their common ancestor. Orthologous genes are produced by speciation (orthogenesis), by opposition to paralogous genes produced by duplication. Thus, an osteichtyan-level orthology group includes human and zebrafish genes with a direct ancestor in the LCATT. Vertebrate level orthologues share this ancestral gene at the basal or jawed vertebrate level. 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.