Quantitative trait loci for resistance to Flavobacterium psychrophilum in rainbow trout: effect of the mode of infection and evidence of epistatic interactions

Pierre Boudinot, Steve Bird, Louis Du Pasquier, Bertrand Collet

Developmental and Comparative Immunology 116 (2021) 103929

Contents lists available at ScienceDirect

Developmental and Comparative Immunology

journal homepage: www.elsevier.com/locate/devcompimm

The repertoire of vertebrate STAT transcription factors: Origin and variations in fish

Pierre Boudinot [a,**], Steve Bird [b], Louis Du Pasquier [c], Bertrand Collet [a,*]

[a] Université Paris-Saclay, INRAE, UVSQ, VIM, 78350, Jouy-en-Josas, France
[b] Biomedical Unit, School of Science, University of Waikato, Hamilton, 3240, New Zealand
[c] Zoology and Evolutionary Biology, University of Basel, 4051, Basel, Switzerland

ARTICLE INFO

Keywords:
STAT
Interferon signalling
Gene duplication
Comparative immunology
Evolution
Vertebrates

ABSTRACT

The *stat* gene family diversified during early vertebrate evolution thanks to two rounds of whole genome duplication (WGD) to produce a typical repertoire composed of 6 STAT factors (named 1–6). In contrast, only one or two *stat* genes have been reported in *C. elegans* and in *D. melanogaster*. The main types of STAT found from bony fish to mammals are present in Agnathan genomes, but a typical STAT1-6 repertoire is only observed in jawed vertebrates. Comparative syntenies showed that STAT6 was the closest to the ancestor of the family. An extensive survey of *stat* genes across fish including polyploid species showed that whole genome duplications did not lead to a uniform expansion of *stat* genes. While 2 to 5 *stat*1 are present in salmonids, whose genome duplicated about 35My ago, only one copy of *stat*2 and *stat*6 is retained. In contrast, common carp, with a recent whole genome duplication (5–10My), possesses a doubled *stat* repertoire indicating that the elimination of *stat*2 and *stat*6 additional copies is not immediate. Altogether our data shed light on the multiplicity of evolutionary pathways followed by key components of the canonical cytokine receptor signalling pathway, and point to differential selective constraints exerted on these factors.

1. Introduction

Animals have evolved a number of efficient strategies to combat a large diversity of pathogens. In mammals, complex immune mechanisms are orchestrated and regulated by a network of cytokines acting through cognate ligand/receptor on multiple specialized immunocytes (Aaronson and Horvath, 2002). In mammals, a large number of these cytokines signal through the JAK/STAT signalling factors (Kiu and Nicholson, 2012) composed of a particular combination of four Janus kinases (JAK1-3, Tyrosine Kinase TYK2) and one of 7 Signal Transducer and Activator of Transcription (STAT1-4, 5A, 5B, 6) (Villarino et al., 2017). The pathway involves a cascade of phosphorylation reactions (Decker and Kovarik, 2000), multimeric complex formation and nuclear translocation (Reich, 2013) resulting in the induction of a particular set of genes responsible for a specific cellular response (Liongue et al., 2016). Gamma interferon activation site (GAS) is the core genomic motif targeted by STAT1 homodimers (Nast et al., 2019; Stark and Darnell, 2012; Decker et al., 1997). STATs heterodimers, associated with additional transcription factors, can bind variants of GAS motifs such as interferon-sensitive responsive element (ISRE) resulting in transcriptional regulation of large gene sets. Such gene sets leading to particular immune responses were associated to different STAT-dependant signalling. In addition, variations in the epigenetic status of genomic elements and in the type of immune cells involved explain, at least in part, the "specificity paradox" of the JAK/STAT signalling pathway, namely, how a 7-member protein family can ensure the specificity of response of dozens of cytokines (Lin and Leonard, 2019). The human STAT repertoire is composed of 7 transcription factors encoded by genes located on 3 chromosomes: *STAT1* and *STAT4* closely linked on chromosome 2, *STAT2* and *STAT6* on chromosome 12, and *STAT3*, *STAT5A* and *STAT5B* closely linked on chromosome 17.

All these proteins share four domains: a *N*-terminal Protein

Abbreviations: CCD, coiled coil domain; CBP, CREB-binding protein; GAS, Gamma interferon activation site; ISRE, interferon-sensitive responsive element; SH2, Src homology 2 domain; STAT, signal transducer and activator of transcription; TAD, C terminal transactivation domain; TAZ, Transcription Adaptor putative Zinc finger; WGD, whole genome duplication.

* Corresponding author.
** Corresponding author.
E-mail addresses: Pierre.Boudinot@inrae.fr (P. Boudinot), Bertrand.Collet@inrae.fr (B. Collet).

https://doi.org/10.1016/j.dci.2020.103929
Received 2 October 2020; Received in revised form 15 November 2020; Accepted 15 November 2020
Available online 1 December 2020

interaction domain ("STAT-i"), a coiled coil domain ("STAT a", CCD), a DNA-binding domain ("STAT b", DBD) and a Src homology 2 domain ("STAT-SH2"). Additionally, STAT1 and STAT2 comprise a C terminal transactivation (TAD) domain: the STAT1 transactivation domain (IPR022752) binds selectively to the Transcription Adaptor putative Zinc finger (TAZ)2 domain of C CREB-binding protein (CBP)/p300, while the STAT2 transactivation domain (IPR022756) binds to the TAZ1 domain of this protein (Wojciak et al., 2009; Bhattacharya et al., 1996; Zhang et al., 1996) (Fig. 1A). This domain confers to STAT1 and STAT2 an additional capacity to regulate gene expression since CBP and P300 are histone acetyltransferases that control acetylation of histones in nucleosomes, thus regulating chromatin remodelling and gene transcription.

The vertebrate *stat* repertoire emerged from an ancestral sequence present in the common ancestor of protostomes and deuterostomes with all STAT typical domains (Wang and Levy, 2012), through WGD, tandem duplication and dispersion (Copeland et al., 1995). Liongue et al. (2012), proposed that vertebrate *stat* genes originated from a set of two paralogs produced by local duplication, subsequently duplicated "en bloc" by the two rounds of WGD that occurred during early vertebrate evolution, leading to four copies of this cluster. Three of these copies (*STAT3-STAT5, STAT2-STAT6,* and *STAT1-STAT4*) have been retained in human and most vertebrates. In zebrafish, additional copies of *stat1* and *stat5* were found, likely due to the additional, teleost-specific WGD (Liongue et al., 2012).

In this work, we revisited the origins and the evolutive dynamic of the vertebrate *stat* gene repertoire. To find out whether duplicated *stat* copies were retained or lost, we focused on groups and species in which additional WGD occurred. We thus focused on ray finned fish because their genomes were subjected to several WGD events including a teleost-specific WGD event ("3R") that occurred at the root of this lineage about 350 million years ago (Myr) and more recent events for example in salmonids 50–60 Myr ago (Lien et al., 2016; Pasquier et al., 2016) and carps 5–10 Myr ago. In addition, salmonid fish such as Atlantic salmon and Rainbow trout are the most relevant species for the fish farming industry in Europe and worldwide, and their genomes are among the best characterized in teleost fish. We also characterized *stat* genes from Chondrichthyans, Agnathans and non-vertebrate deuterostomians to clarify how these transcription factors evolved during the emergence of vertebrates.

2. Results and discussion

2.1. The repertoire of stat genes is well conserved across tetrapods

A fundamental repertoire of six *stat* genes is well-conserved across all tetrapod classes and in the coelacanth, as illustrated in Fig. 1B (see also Table S1). One-to-one orthology relationships between tetrapod and

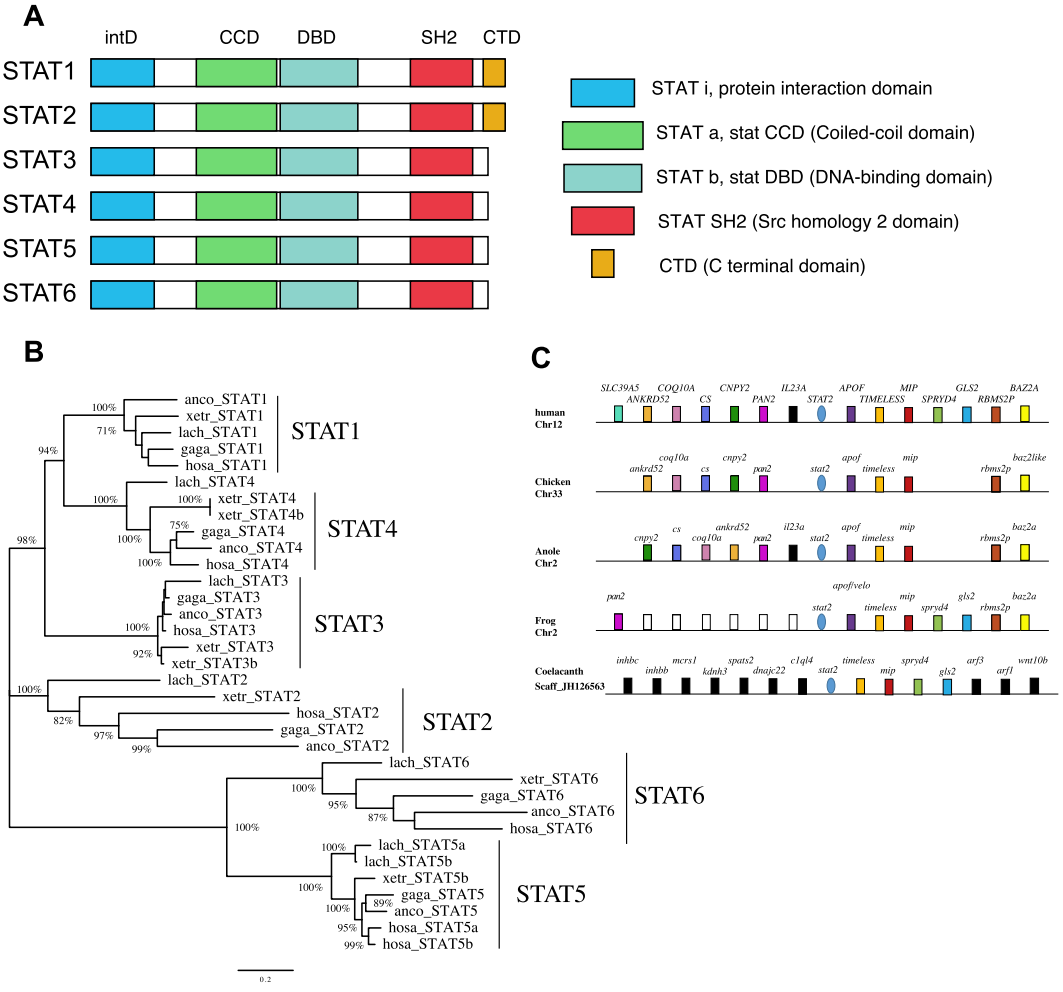


Fig. 1. Evolutionary history of STAT transcription factors across tetrapods. A. Domain structure of vertebrate STAT proteins. **B**. Maximum likelihood phylogenetic tree of STAT amino-acid sequences from human *Homo sapiens* (hosa), chicken *Gallus gallus* (gaga), *Anolis carolinensis* anole (anco), clawed frog *Xenopus tropicalis* (xetr) and *Latimeria chalumnae* coelacanth (lach). Bootstrap values (in %) of key nodes are indicated. Bootstrap values lower than 60% are not indicated. All sequences and sequence ID are provided in Table S1. **C**. Conservation of genomic neighborhood of *stat2* genes in the same tetrapod species (based on genome assemblies from Ensembl release 100).

coelacanth genes are also supported by conserved synteny groups comprising several markers flanking all *stat* gene clusters (as shown for *stat2* in Fig. 1C).

2.2. Loss and retention of stat genes after WGD during fish evolution reveal contrasted constraints on different stat subtypes

In ray-finned fishes, the *stat* repertoire comprises the same types as in the coelacanth and tetrapods, with *stat1, stat2, stat3, stat4, stat5* and *stat6* present in all species across teleosts. After a WGD occurred some 350 Myr during the early evolution of this group, two copies of each *stat* gene should have been generated (Van de Peer, 2004; Jaillon et al., 2004; Christoffels et al., 2006). This *stat* repertoire has been reshaped by further duplication and gene loss.

In fish groups that did not undergo additional WGD, such as herring (*Clupea harengus*), pike (*Esox lucius*), zebrafish (*Danio rerio*), stickleback (*Gasterosteus aculeatus*) and the marine species fugu (*Fugu rubripes*) and sea bream (*Sparus aurata)* (Fig. 2 and Table S1), *stat*2-6 could be found as single copy in contrast with two or more, *stat1* paralogs. One *stat1* paralog (named "b") is always linked to *stat4* as observed across tetrapods, while the other copy (named "a") is located on another chromosome. This was also the case of Atlantic cod (*Gadus morhua)*, a gadiform species with a particular immune system lacking CD4 and a functional MHC class II pathway. In some cases, an additional *stat1* can be found like in herring on a third chromosome (Fig. 2). In zebrafish, a *stat1* pseudogene has been described close to *stat1b* (Liongue et al., 2012), but is not present in the last genome assembly. Only one copy of *stat3, 4* and *5* was generally present in these species with a few exceptions as a double *stat5* in zebrafish, produced by a local duplication. In contrast, the retention of two functional *stat1* genes across multiple families of ray-finned fish suggests that different types of selection pressures may affect this gene, compared to other stat family members.

2.3. Multiple stat1 paralogs are also retained in tetraploid salmonids

To further test this hypothesis, we then focused on tetraploid species in which larger *stat* repertoires have been produced by an additional WGD, providing the opportunity to test their evolutionary fate.

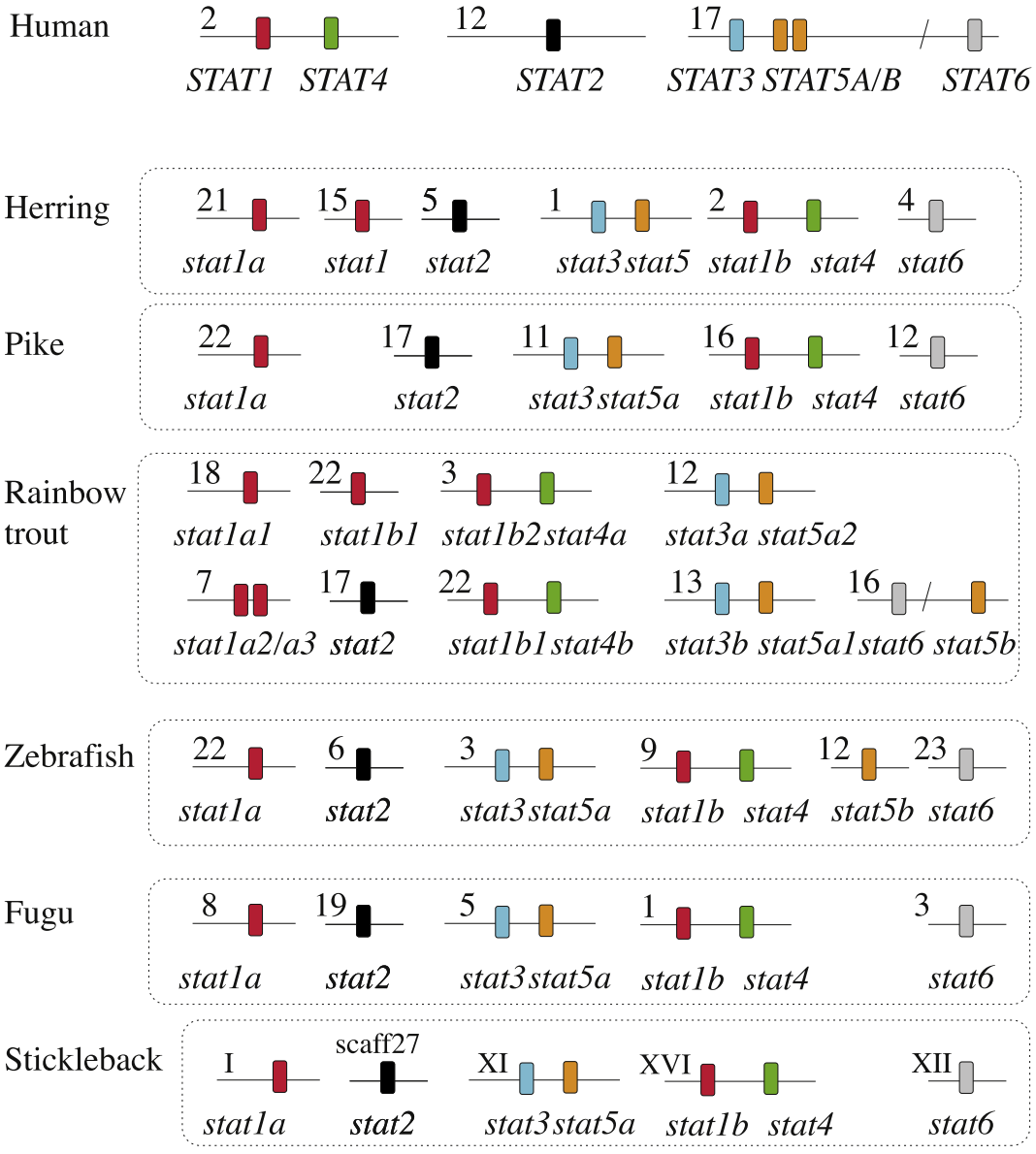


Fig. 2. Repertoires of fish STAT amino-acid sequences and *stat* genes chromosomic distribution. Data from genome assembly Omyk_1.0 (*Oncorhynchus mykiss*, rainbow trout, RefSeq GCF_002163495.1) and for other species from Ensembl release 100.

We first performed a comprehensive survey of *stat* genes in salmonids, a fish family tetraploidized by an additional WGD that occurred about 50–60 Myr ago. In these species, we typically found two blocks *stat3+5*, two blocks *stat4+1,* four or five copies of *stat1,* but only one *stat2* and one *stat6* gene (as for rainbow trout in Fig. 2). A comprehensive characterization of *stat* genes across salmonids is presented in Table 1. Among the two genera *Oncorhynchus* and *Salmo,* we analysed six species for which high quality genomes were available: Sockeye salmon *O. nerka*, rainbow trout *O. mykiss*, chinook salmon *O. tshawytscha*, Coho salmon *O. kisutch*, brown trout *S. trutta* and Atlantic salmon *S. salar*. A total of 16 *stat* loci were found in these six salmonid genomes (Table 1, Fig. 2). They were located on 9 chromosomes corresponding to 6 chromosomes in zebrafish (Fig. 3), a diploid cyprinid. Linkage analyses showed that *stat1a1-3, stat1b-4* and *stat3-5b* duplicated blocks generated by the salmonid-specific WGD were retained (Fig. 3), while there was no evidence of multiple copies of *stat2* and *stat6* (not even pseudogenes).

Phylogenetic and synteny block analyses across species provided consistent insights into the origin of these *stat* genes (Fig. 3 and Figure S1) and allowed unambiguous identification and annotation. For example, all *stat1a* were linked to *ccr4not* and *ftcd* – as zebrafish *stat1a* – while *stat1b* genes were associated to *stat4* and *slc40* genes. A number of sequences encoding ORFs with size lower than 50% of the average size of STATs proteins were additionally found in the rainbow trout, brown trout and Atlantic salmon (Table 2). These, which likely are assembly artefacts or pseudogenes, were not included in the phylogenetic analysis.

In salmonids, an additional (fifth) *stat1* gene that we named *stat1a3* was found immediately downstream of *stat1a2,* suggesting it was generated by local duplication. Interestingly, in the rainbow trout, brown trout and Coho salmon, the STAT1A3 protein is twice the size of the normal size of the STAT1. These long STAT1 proteins contain twice the typical set of domains in tandem [STATi- STATa- STATb- STATSH2-CTD- STATi- STATa- STATb- STATSH2-CTD] and seem to be due to a local duplication-fusion of two *stat1* ORFs. The double *stat1a3* was confirmed in the rainbow trout by the EST CA361350 covering the junction area between the end of the putative first *stat1* and the beginning of the second, which excludes that *stat1a3* has been produced by an assembly error. Further functional studies are required to determine the function of the encoded protein, its potential intramolecular dimerization and GAS elements binding abilities.

While the ancient WGD that in early teleost fish has left two *stat1* but only one of the other *stat* paralogs in diploid species, the more recent salmonid-specific WGD resulted in five *stat1* being retained. In contrast, only one copy of *stat2* and *stat6* were kept, either due to an early complete loss post-WGD or because of consistent selection pressures in favour of a single copy.

2.3. Up-regulation of salmonid stat genes during antiviral responses

Fig. 4 shows the expression profile of all chinook salmon *stat* genes from an RNAseq experiment carried out on the EC cell line (Dehler et al., 2019). We checked whether the salmonid *stat1* and *stat2* genes were induced by type I IFN in a manner consistent with zebrafish where, in zebrafish larva, recombinant IFNφ1 induces a robust up-regulation of *stat1b* and *stat2,* but not of *stat1a* (Levraud et al., 2019). In the chinook salmon cell line EC (Dehler et al., 2019) *stat1b1* and *stat2* were induced with a FC > 1.5 following stimulation by salmonid recombinant type I IFN. *Stat1a1* was also induced to some extent (Fig. 4). However, *stat1b2* was not up-regulated. At steady state, *stat1a* paralogs were more expressed than *stat1b,* as in zebrafish, a pattern consistent with a functional constitutive expression of *stat1a* genes (Dehler et al., 2019).

Thus, there is no strict conservation of the *stat1/2* genes inducibility between salmonids and zebrafish, although the most upregulated genes are *stat1b* and *stat2* in both species. Overall, the paralogs of a given genes may be expressed at low levels in healthy cells, but can reach much higher levels after stimulation, offering opportunities for complex regulations. Whether this profile is different in other cells or under different stimulation conditions remains to be clarified. Similar variations of steady state expression levels were also observed for *stat5: stat5.1* and *stat5.2* were detected at low levels, while *stat5.3* transcripts were at least 10 times more abundant (Fig. 4).

Duplicated genes in polyploid species are expected to be eliminated by deletion/accumulation of mutations, if they do not acquire new functions (neo/sub-functionalization) or are not kept by selection for gene dosage (Levasseur and Pontarotti, 2011). Our data about zebrafish and salmonid multiple *stat1* paralogs strongly suggest that they were

Table 1
List of potentially functional stat genes in five species of salmonids fish. The nomenclature is extended from *Danio rerio*.

	Oncorhynchus	*Salmo*				
	nerka	*mykiss*	*tshawytscha*	*Kisutch*[f]	*Truttae*	*salar*
stat 1a1	115129673 (5/+)	100136755 (18/-)	112266551 (14/-)[a]	109906681 (16/-)	115195984 (6/+)	100136558 (16/+)
stat 1a2	115133971 (2/-)	100137016 (7/-)	112253897 (7/-)[b]	109890649 (5/-)	115179269 (39/-)[##]	
stat1a3	115133986 (2/+)	110527523 (7/+)*[,d,g]	112253898 (7/+)[c]	109890433 (5/+)[d]	115179399 (39/+)[d]	106575214[#] (17/-)
stat 1b1	115138513 (1/-)	110501544 (22/+)	112244575 (Un/+)	109871062 (26/+)	115156118 (20/+)	100196256 (21/-)[e]
stat 1b2	115108041 (3/-)	110520020 (3/-)	112235369 (3/-)	109865728 (2/+)	115160666 (24/-)	106586142 (25/-)
stat 2	115143174 (15/+)	110494323 (17/+)	112217577 (2/+)	109895769 (1/+)	115207827 (14/-)	100270812 (12/+)
stat3a	115110281 (26/-)	110538194 (12/-)	112225989 (27/-)***	116374454 (6/-)	115171116 (32/-)	106601025 (3/-)
stat3b	115104128 (21/-)	100136756 (13/-)	112258660 (9/-)	109898422 (10/-)	115197994 (1/-)	106607297 (6/+)
stat4a	115108086 (3/-)	110520023 (3/-)	112235400 (3/-)	109865765 (2/-)	115160669 (24/-)	106586145(25/-)
stat4b	115138591 (1/-)	110501546 (22/+)	112225151 (26/-)***	109870786 (26/+)	115156112 (20/+)	100380385 (21/-)
stat5a1	115103495 (21/-)	100135887 (13/-)**	112258659 (9/-)[##]	109897291 (10/-)	115198021 (1/-)	106607295 (6/+)**
stat5a2	115110283 (26/-)	110538192 (12/-)	112225727 (27/-)[##]	109893324 (6/-)	115171115 (32/-)	100380532 (3/-)
stat5b	115106643 (23/+)	110491683 (16/+)	112223777 (24/+)	109865530 (20/+)	115161213 (2/+)	106579144 (19/+)
stat6	115102666 (20/+)	110491929 (16/+)	112221662 (22/+)	109869625 (24/+)	115167799 (30/+)	106567004 (13/+)

*annotated as "uncharaterised protein", **RefSeq status indicated as provisional, ***annotated as "low qbity protein", #annotated as pseudogene in NCBI but as coding for ENSSSAT00000091945 in Ensembl, ##annotated as pseudogene.

[a] Duplicated due to genome assembly errors, identical to 112253955.

[b] Duplicated due to genome assembly errors, identical to 112253778.

[c] Duplicated due to genome assembly errors, identical to 112253779.

[d] Double size.

[e] Smallsize (not incl[e]ded in phylogenetic analysis).

[f] The chinook genome still contains some assembly errors resulting in artificially duplicated regions in particular between the chromosomes 7 and 14 and within the chromosome 7. In this species, only one stat5 was annotated as functional in contrast to 3 stat5 genes found in the five other species analysed, named stat5.1-3.

[g] stat1a3 is doubled in NCBI but separate in Ensembl as ENSOMYG00000034815 and ENSOMYG00000035706, both annotated as stat1a.

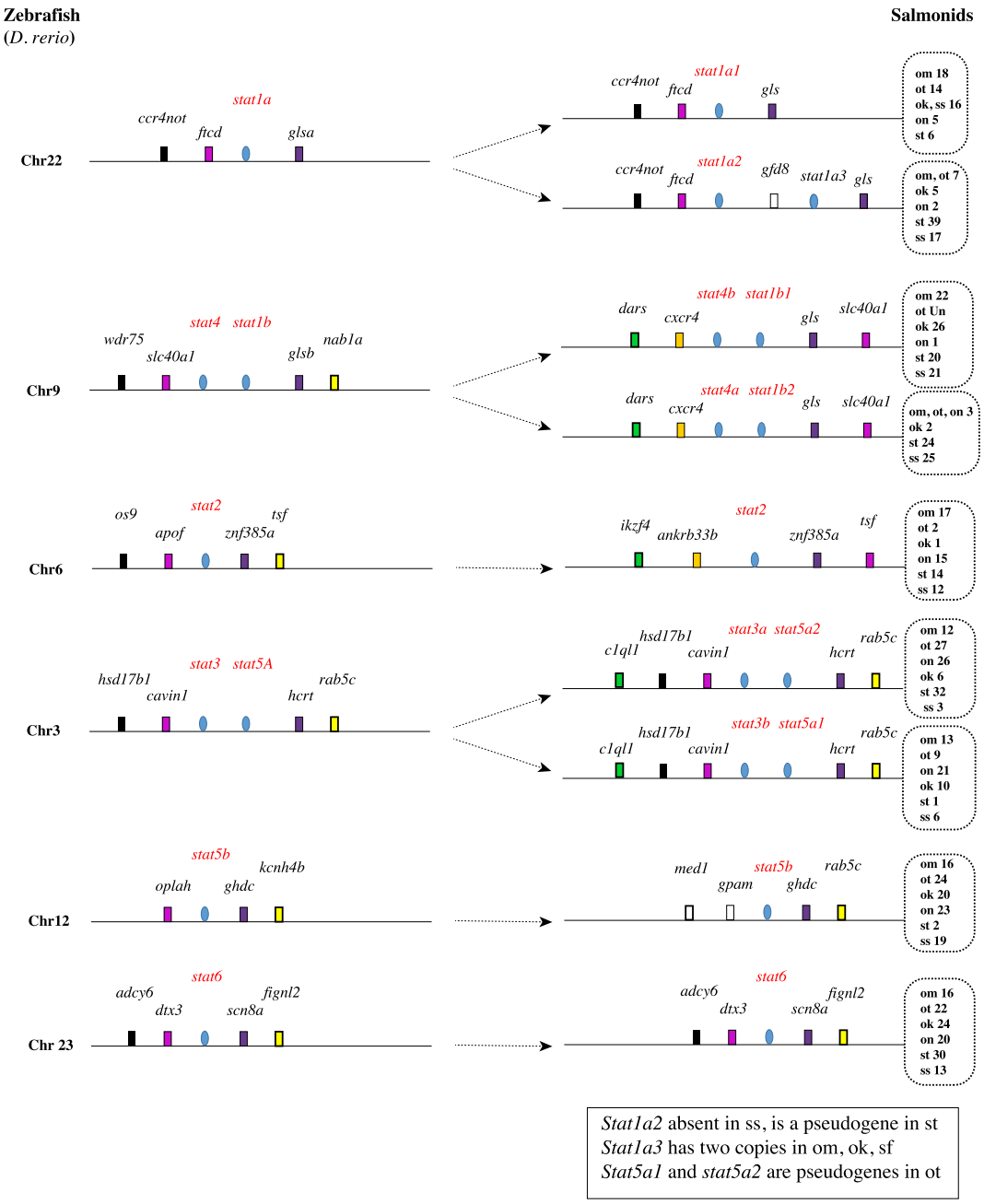


Fig. 3. *stat* genes from genomes of Salmonidae (4 *Oncorhynchus* species and 2 *Salmo species*): synteny block conservation analysis. Data based on NCBI genome assemblies: Okis_V2, *Oncorhynchus kisutch* (coho salmon): GCF_002021735.2; Oner_1.0, *Oncorhynchus nerka* (sockeye salmon): GCF_006149115.1; Otsh_v1.0, *Oncorhynchus tshawytscha* (Chinook salmon): GCF_002872995.1; Oket_V1, *Oncorhynchus keta* (chum salmon): GCF_012931545.1; Omyk_1.0, *Oncorhynchus mykiss* (rainbow trout): GCF_002163495.1; fSalTru1.1, *Salmo trutta* (river trout): GCF_901001165.1; ICSASG_v2, *Salmo salar* (Atlantic salmon): GCF_000233375.1.

indeed subjected to neofunctionalization. More functional work will be necessary to establish if this is also true for salmonid *stat3, 4* and *5* paralogs.

2.5. Classification and nomenclature of stat genes in tetraploid species based on the example of salmonids

The survey of the *stat* gene cluster in salmonid fish highlighted a nomenclature issue for *stat* genes in polyploid species. The current annotation of such complex duplicated genomes is often misleading because of assembly errors. Some annotations inherited the nomenclature used at the time of the first and often single gene discovery by homology cloning and lack consistency with annotation in other fish species. Regarding salmonid *stat* genes, the rainbow trout *stat1a1* and *stat1a2* were annotated *stat1-1* and *stat1-2* with no reference to the *stat1a* group defined previously in non-salmonid teleost such as zebrafish. The *stat1a3* was left annotated as "uncharacterized protein" whereas phylogeny and blast against the mammalian protein database allocated it to the *stat1* group. Based on our results from phylogeny and synteny conservation, we therefore established a coherent nomenclature (Fig. 3, Table S1). A similar approach may be followed in other groups of tetraploid vertebrates for example in Amphibians.

2.6. Other tetraploid genomes tell more about stat evolutionary dynamics

We also studied the *stat* genes from the common carp (*Cyprinus*

Table 2
Additional genes encoding short ORF with significant hit to signal transducer and activator of transcription. None could be identified in *O. tshawytscha* nor in *O. kisutch.*

Species	GeneID (chromosome/orientation)	Size of the largest isoform (aa)
O. mykiss	110503197 (24/+)	243
S. truttae	115188323 (Un/+)	184
	115189982 (Un/+)	249
	115190002 (Un/-)	136
	115156120 (20/+)	249
	115156122 (20/+)	186
	115156123 (20/+)	179
	115156124 (20/+)	147
	115156138 (20/+)	186
	115156139 (20/+)	184
S. salar	106583229 (22/-)	314

carpio, Ensembl 100: German_Mirror_carp_1.0), an allotetraploid teleost due to a recent WGD that occurred relatively recently 5–10 Myr ago. In this species, all duplicated loci were retained, with exactly twice as many genes as in the diploid cyprinid zebrafish with 4 *stat1*, 2 *stat2*, 2 *stat3*, 2 *stat4*, 4 *stat5* and 2 *stat6* (Table S1). The phylogenetic tree and the distribution of these genes in contigs indicate that they correspond to a duplication of the blocks typically found in zebrafish and other diploid teleosts (Table S1, Fig. 4A).

Polyploid species also originate by allopolyploidization, *i.e.* by genome association due to hybridization among different species. The availability of the genome for the frog *Xenopus laevis* (2n = 36) offers an opportunity to estimate the effect of evolution of the two subgenomes of an allotetraploid species that were combined about 17–18 Myr ago, on the diversification of the *stat* gene family (Session et al., 2016). In parallel, we analysed the *stat* repertoire from the genome of *Xenopus tropicalis* (2n = 20), which is not made of obvious pairs of homoeologous chromosomes (Uno et al., 2013). Thirteen (13) and seven (7) *stat* genes were identified in the genome of *X. laevis* and *X. tropicalis*, respectively. *X. laevis* shows an almost perfect duplication, with the exception of the loss of *stat4. S* (Table S1; Fig. 4B), while 8.3% and 31.5% of *X. laevis* genes with clear 1:1 or 2:1 orthologs in *X. tropicalis* were lost, respectively, from L and S subgenomes (Session et al., 2016).

Interestingly, these observations in the common carp *Cyprinus carpio* and the African clawed frog *Xenopus laevis* show that additional *stat* genes in polyploid species are not rapidly eliminated, maybe because different copies can get specialized functions easily and quickly. The presence of two functional *stat2* and *sta6* genes is tolerated in both cases, and the loss of one copy is not necessarily immediate after duplication. Furthermore, the pattern of evolution of duplicated genomes in salmonids suggest some selection pressures possibly associated to viral subversions strategies (Nan et al., 2017).

2.7. Agnathan-specific stat genes shed light on the origin of vertebrate STAT transcription factors

The repertoire of *stat* genes is generally more diverse in Vertebrates than in other Metazoans (Liongue et al., 2012), likely due to the two cycles of WGD that occurred in the early evolution of this lineage. To get insight into the early steps of *stat* evolution in vertebrates, we analysed genomes from cartilaginous fish (*i.e.*, Chondrichthyans) and Agnathans.

Orthologs of all vertebrate *stat* were found in cartilaginous fish (Table 3, Fig. 5A). *Stat1, 2, 3, 5* and *6* have been annotated in most species of sharks and rays for which a genome is available (Table S1). A typical *stat4* genomic sequence was not detected in shark genomes except in the whale shark *Rhincodon typus* (Genbank ID XP_020376005). An EST was also found in the dogfish shark *Squalus acanthias* (Genbank ID EE627912). Phylogenetic analysis confirmed that *stat* genes from Chondrichthyans have human orthologs (Fig. 5A).

In contrast, the list of *stat* genes was different in Agnathans: in two species, the sea lamprey *Petromyzon marinus* and the hagfish *Eptatretus burgeri,* phylogenetic analysis identified orthologs of human *STAT3* and *STAT5* (Fig. 5A). Two other *stat* sequences clustered with group1-4 (later referred as "*stat1-4*") and group5-6 (later referred as "*stat5-6*") respectively but could not be assigned to a particular set, suggesting that the *stat* repertoire of "modern" vertebrates was consolidated and standardized in Gnathostomes. Additionally, the genomic neighborhood of

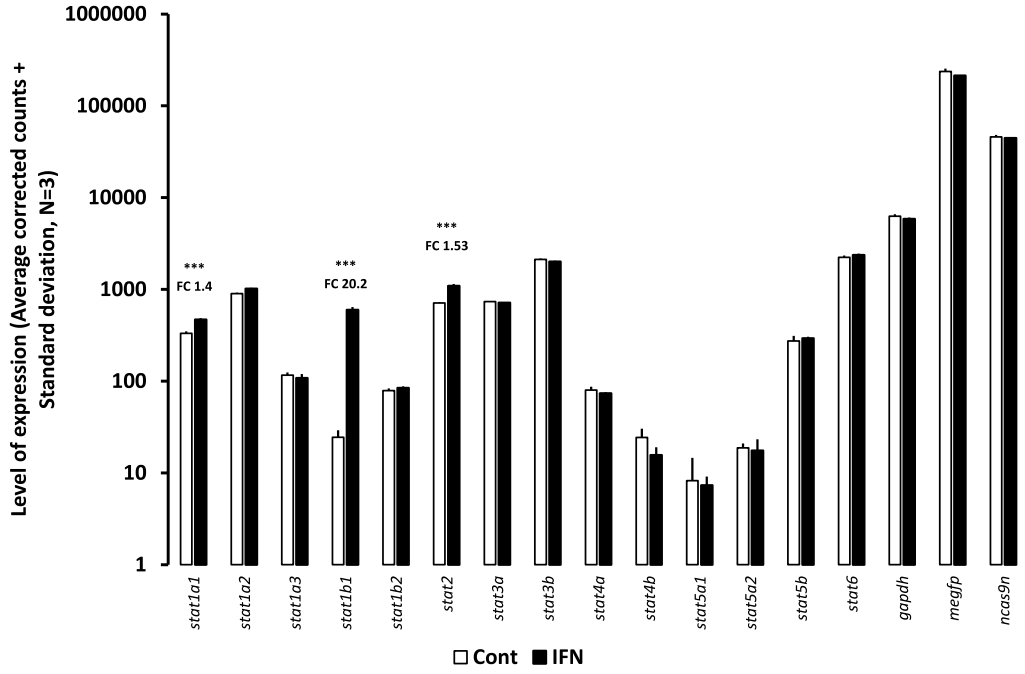


Fig. 4. Expression levels of *stat* genes (basal and induced by recombinant type I interferon) determined by RNAseq in CHSE-EC cell *O. tshawytscha* (Dehler et al., 2019)**.** Data are on a log scale and represent the average + Standard deviation (N = 3). When the induction is statistically significant (***$p < 0.001$), the Fold Change is indicated.

Table 3
Presence and number of *stat* genes in genomes of Chondrichthyans and Agnathans.

Phylum	Species	Gene number	Gene name	Conserved domains
Chondrichthyans	Elephant shark *Callorhincus milli*	≥4	ENSCMIG00000003696 (3)	STATi-STATa-STATb-SH2
			ENSCMIG00000003757 (5b)	STATi-STATa-STATb-SH2
			ENSCMIG00000010732 (1)	STATi-STATa-STATb-SH2
			ENSCMIG00000015418 (1)	STATi-STATa-STATb-SH2**
	Whale shark *Rhycodon typus*		XP_020376005	STATi-STATa-STATb-SH2
Agnathans	Lamprey *Petromyzon marinus*	4	ENSPMAG00000000244	STATi-STATa-STATb-SH2
			ENSPMAG00000002770	STATi-STATa-STATb-SH2
			ENSPMAG00000006622	… STATa-STATb …
			ENSPMAG00000009008	STATi-STATa-STATb-SH2
	Hagfish *Eptatretus burgeri*	4/5	ENSEBUG00000002998& ENSEBUG00000003439	… STATb-SH2
			&ENSEBUG00000006280	…-SH2 (separated by gls)
			ENSEBUG00000007506	STATi-STATa-STATb-SH2
			ENSEBUG00000013827	STATi-STATa-STATb

**there are 2 STAT gene in this entry !!

agnathan *stat* did not fit the well-conserved synteny blocks observed in jawed vertebrates (Fig. 5B). These regions contain markers located close to *stat* genes in vertebrates, such as in *ab1, gls, myo1b, cavin1, tmeff2, slc39A10, dnah7*. However, these markers do not seem to be associated consistently with *stat* sequences in agnathans and jawed vertebrates, suggesting that these regions were produced by several duplications of an ancestral segment followed by extensive gene loss, making the reconstitution of the history of this region difficult. Markers have been best conserved in the regions encoding tetrapod *stat1, stat2, stat3 and stat4* and lamprey *stat1-4* (Fig. 5B). While two *stat* genes closely linked on lamprey scaffold 5 are most similar to *stat3* and *stat5/6* respectively, the markers found at close proximity do not match with genes located close to human *stat3* and *5*: in human, *nab1*, *gls* and *myo1* homologs are located on chromosome 2 close to *stat2* and *stat4*. Moreover, the lamprey *stat5* is linked to *cavin*, a marker associated to human *stat*3/5, but also to *timeless* which is found close to human *stat2* on chromosome 12.

Thus, all vertebrates seem to possess genes from both *stat1-4* and *stat5-6* groups, encoded in genomic blocks inherited from an ancestral region containing *nab1, gls, myo1b, cavin1, tmeff2, slc39A10,* and *dnah7* genes. However, the standardized *stat* repertoire found in human was apparently established later in early gnathostomes. Further assemblies of agnathan genomes will help to better understand the evolution of this region.

2.8. Conserved linkages indicate that stat6 is a genomic environment closest to the ancestral stat gene

We then analysed genomes from other deuterostomians. In these species, the repertoire of *stat* genes was significantly smaller compared to vertebrates (Table 4, Figure S2): three *stat* sequences were found in the cephalochordate lancelet *Branchiostoma floridae* and in the tunicate *Ciona intestinalis,* and one in the appendicularia *Oikopleura dioica*, in the hemichordate *Saccoglossus kowalevskii* and in the sea urchin *Strongylocentrotus purpuratus.* Most sequences clustered in phylogenetic trees with STAT5 and STAT6 (data not shown), as reported previously for non-vertebrate STAT sequences (Liongue et al., 2012). Only one sequence from the lancelet was more similar to the STAT1-4 group (Figure S2). STAT5 and STAT6 have pleiotropic roles and are involved as transcription factors in the biology of different cell types including epithelial and haematopoietic as well as immune cells. Such critical functions in a wide range of contexts are consistent with a primordial status of these genes within the family. Overall, these results confirmed that a complete STAT1-6 repertoire could not be found in these species.

As published previously (Wang and Levy, 2012), protostomians genomes also contain typical *stat* genes, sometimes with multiple copies such as in the annelids *Helobdella* and *Capitella* (Table S1). These sequences were most similar to vertebrate *stat5* and *stat6* as previously reported (Liongue et al., 2016). However, we were not able to find any *stat* synteny blocks shared between these species and vertebrates. In contrast, linkage groups with *stat* genes from the placozoan *Trichoplax adhaerens* stand out as an intriguing exception (Fig. 6), which reminds of our previous report about MHC (Suurväli et al., 2014). In this species, *stat* genes were mainly located close to each other on scaffold 2. Seven genes flanking this cluster were homologous to 7 markers located on human chromosome 12, most of them in the close neighborhood of *STAT6* and *STAT2*, to 3 markers on human chromosome 2 close to *STAT1*, and to one marker on human chromosome 17 close to *STAT5* and *STAT3*. Interestingly, the best conserved set of linkages involved the region of *stat6*, which appears to be most closely related to the ancestral *stat* in phylogenetic analyses. These observations are consistent with the idea that both vertebrate and *Trichoplax* genomes evolved relatively slowly while those of Protostomes were subjected to extensive rearrangements. It also establishes a link between vertebrate *stat* genes and basal bilaterians.

3. Conclusions

The canonical signalling pathway “cytokine receptor - JAK/STAT“ contributes to many functions in invertebrates, as illustrated by in depth studies in Drosophila. In this species, this axis in involved in embryonic segmentation, in stem cell proliferation, in growth as well as in immunity (Perrimon and Mahowald, 1986; Zeidler et al., 1999; Gregory et al., 2008; Morin-Poulard et al., 2013; Rajan and Perrimon, 2012; Vanha-aho et al., 2016; Stepkowski et al., 2008a; Stepkowski et al., 2008b). This repertoire of types of STAT transcription factors was remarkably stable during tetrapod evolution. We found the same types of STAT in Chondrichthyans but not in Agnathans, showing that this repertoire was likely standardized with the emergence of Gnathostomes. In ray finned fish, successive WGD offered multiple opportunities of further functional diversification and specialization. Our work shows that only *stat1* paralogs were retained after the R3 WGD, with one being constitutive and the other strongly induced by IFN. Focusing on Salmonids, we found several *stat-1*, *-3*, *-4* and *-5* due to the most recent WGD, while one copy of the *stat2/6* block has been retained. With 5 paralogs and a remarkable long version with additional domains, *stat1* stands out as the only member of the family prone to expansion and diversification. We have already reported that chinook salmon cells in which *stat1a1* and *stat1a2* (with constitutive expression) have been disrupted, completely lost type I IFN responsiveness (Dehler et al., 2019). Further work is needed to dissect the specialized functions of these multiple *stat1* in various cell

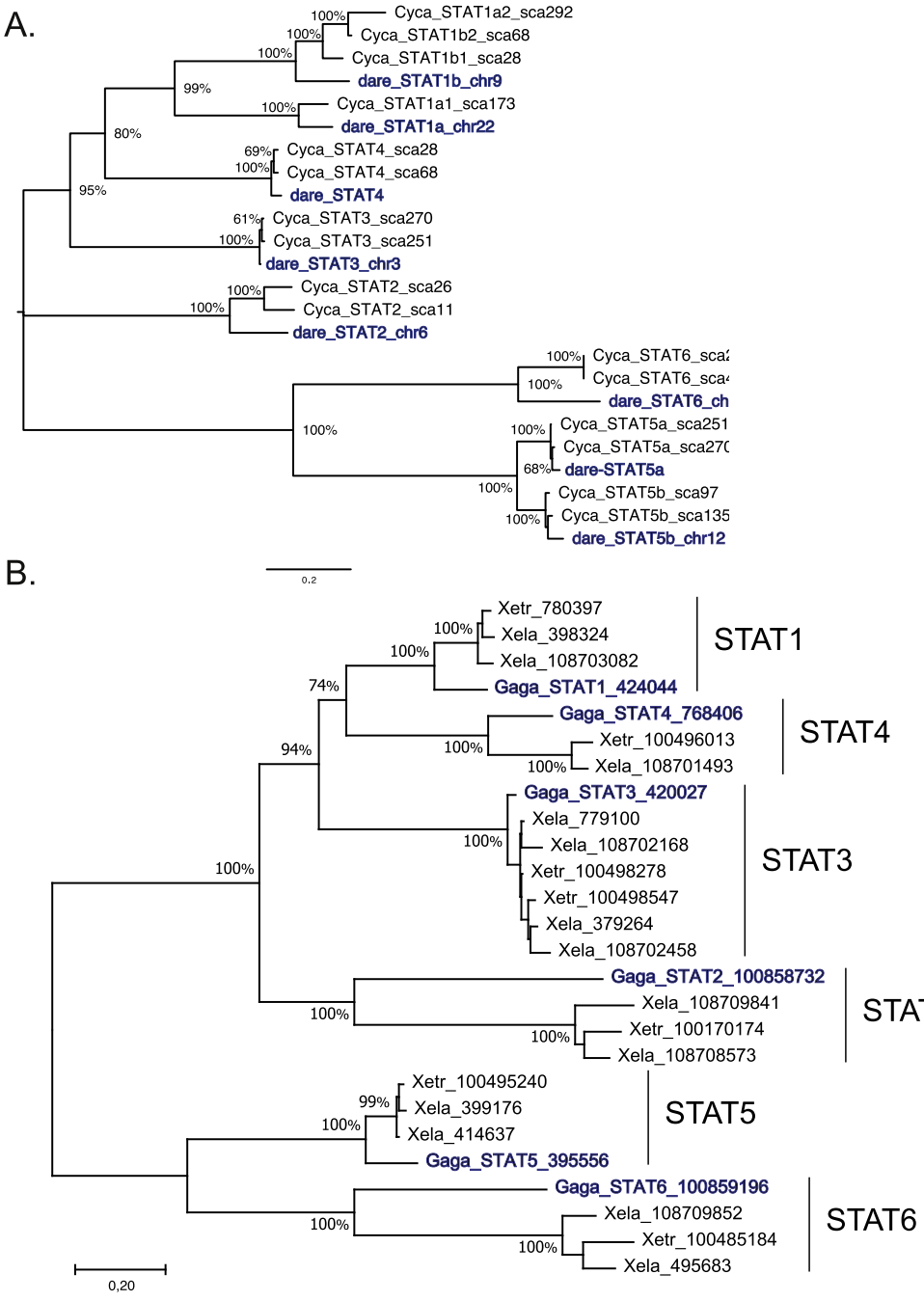


Fig. 5. STAT repertoires in other polyploid species. **A.** Phylogenetic tree of STAT proteins from common carp *Cyprinus carpio* (Cyca) and zebrafish *Danio rerio* (Dare). The evolutionary history was inferred using the Maximum likelihood method (number of bootstrap tests:1000 replicates). Bootstrap values (in %) of key nodes are indicated. Bootstrap values lower than 60% are not shown. All ambiguous positions were removed for each sequence pair (pairwise deletion option). The chromosome (for zebrafish) or the scaffold (for carp) are indicated. **B.** Phylogenetic tree of STAT proteins from *Xenopus tropicalis* (Xetr), *Xenopus laevis* (Xela) and *Gallus gallus* (gaga). The gene ID are indicated and refer to Table S1. The evolutionary history was inferred using the Maximum likelihood method as for A.

Table 4
Presence and number of *stat* genes in genomes of other deuterostomians.

Phylum	Species	Gene number	Gene name	Conserved domains
Non vertebrate Deuterostomians	Sea urchin *Strongylocentrotus purpuratus*	1	SP-STAT	STATi-STATa-STATb-SH2
	Saccoglossus kowalevskii (Hemichordates)	1	XP_006814941	STATi-STATa-STATb-SH2
	Branchiostoma floridae (Cephalochordates)	2	XP_019630041/BL18533	STATi-STATa-STATb-SH2
			XP_002594129BL09530	STATi-STATa-STATb-SH2
	Oikopleura dioica (Appendicularia)	1	AAS21327	STATi-STATa-STATb-SH2
	Ciona intestinalis (Tunicates)	3	ENSCING00000004044	STATi-STATa-STATb-SH2
			ENSCING00000010308	…STATa-STATb-SH2
			ENSCING00000024295	STATi-STATa

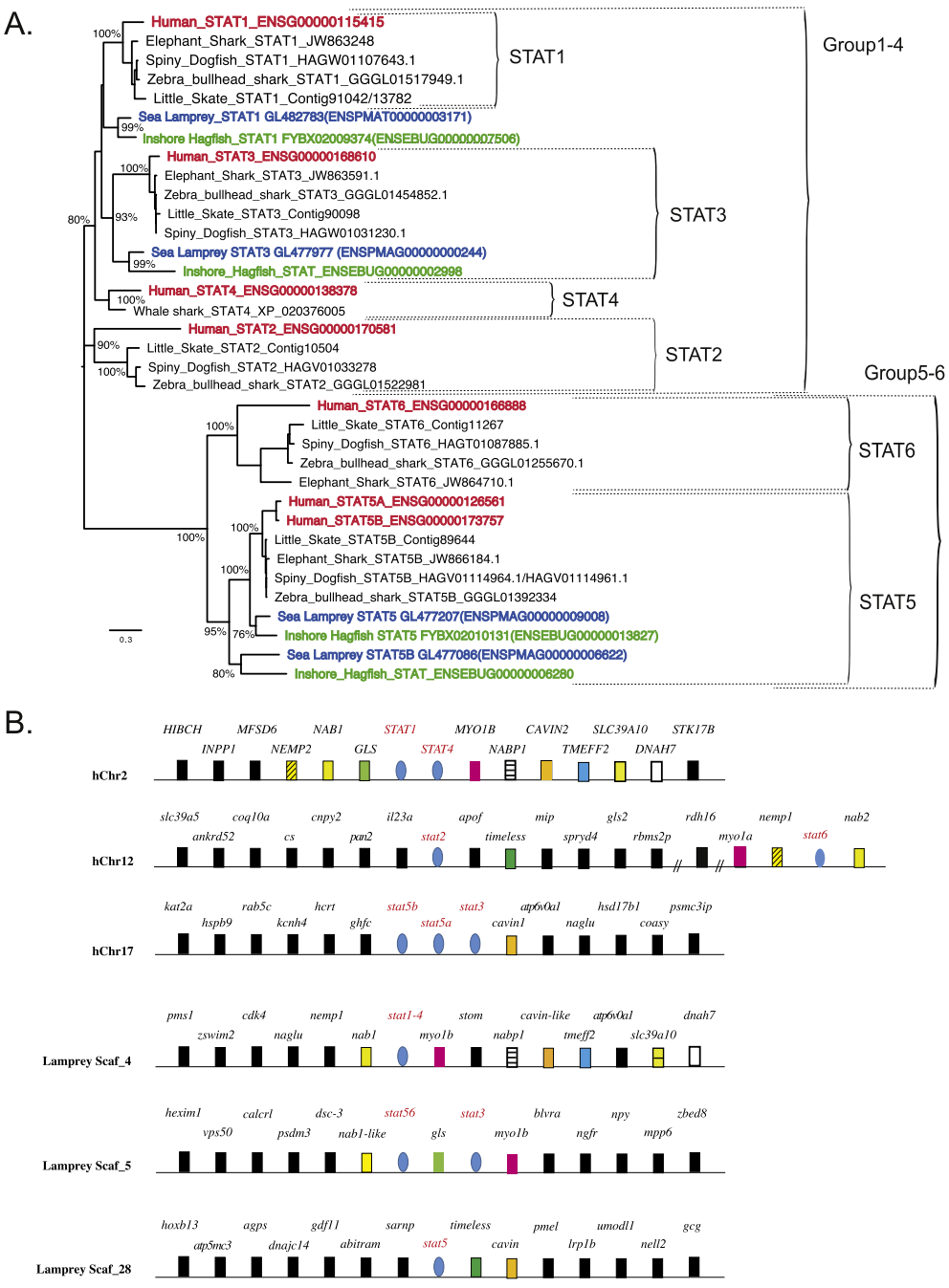


Fig. 6. STAT amino-acid sequences from Chondrichthyans and Agnathans. A. Maximum likelihood phylogenetic tree of STAT amino-acid sequences from human, elephant shark (a chimera), spiny dogfish and zebra bullhead sharks, the little skate (a ray) and sea lamprey and inshore hagfish (Agnathans). Bootstrap values (in %) of key nodes are indicated. Bootstrap values lower than 60% are not indicated. All sequences and sequence ID are provided in Table S1. B. Genomic context of *stat* genes in human and sea lamprey based on data from Ensembl release 100 (Human GRCh38. p13 and Sea Lamprey Pmarinus_7.0).

types and infectious contexts. This evolutionary trend seems to be supported by the high number of *stat1* genes in cyprinids which have been subjected to an independent WGD. Overall, our work shows that the kinetics of *stat* loss is consistently variable across the members of the family (Fig. 7). Hence, the *stat* gene family is particularly suited to study the fate of recently duplicated genes and in particular, loss-of function (or pseudogeneization), dosage effect and neofunctionalization aspects (Lan and Pritchard, 2016). Contrasted inducibility of *stat* paralogs, which is a key mechanism of *stat* mediated immune responses, provide a fast and efficient pathway towards neo/sub-functionalization for these critical factors (Fig. 8).

4. Material and methods

4.1. Identification of stat sequences

Genomes analyses were carried out using the Ensembl (Release 100) and NCBI web interfaces. tBlastn and delta blast searches on the Refseq genomes, and genome annotations searches were combined to pull out all the members of the *stat* gene family. The NCBI genomes released version are as follows: Omyk_1.0 for *Oncorhynchus mykiss*, Okis_V2 for *O. kisucht*, Otsh_v1.0 for *O. tshawytscha*, Oner_1.0 for *O. nerka*, fSalTru1.1 for *Salmo trutta*, ICSASG_v2 for *S. salar*, fSpaAur1.1 for *Sparus aurata*, gadMor3.0 for *Gadus morhua*, GRCz11 for *Danio rerio*, UCB_Xtro_10.0 for *Xenopus tropicalis*, Xenopus_laevis_v2 for *X. laevis*, GRCg6a for *Gallus gallus* and GRCh38. p13 for *Homo sapiens*. The domain structure of the proteins encoded by *stat* genes was checked using SMART and

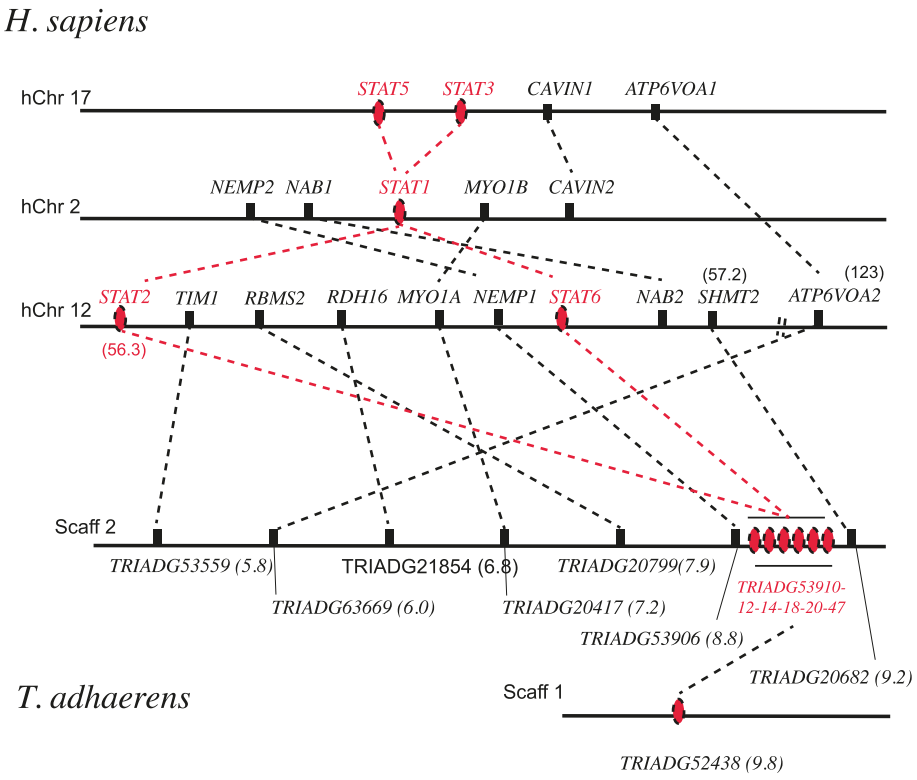
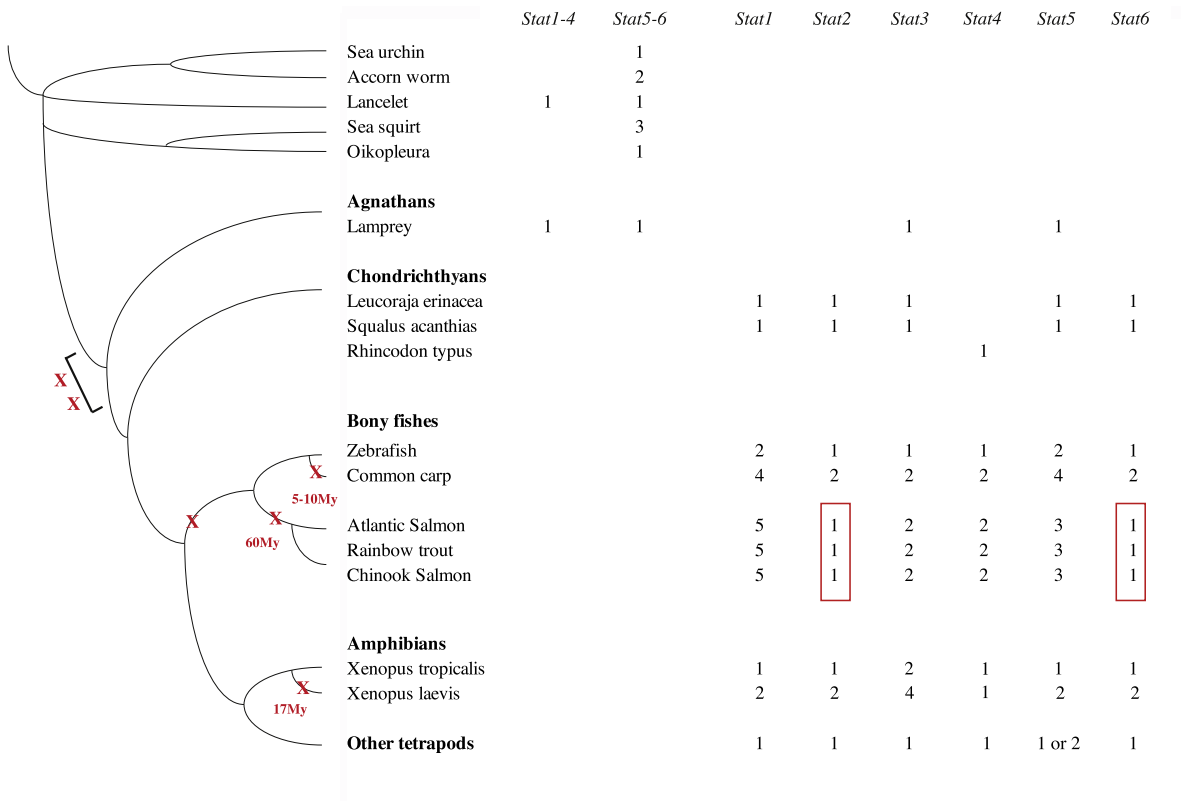


Fig. 7. Conserved genomic neighborhood between human STAT genes located on chromosomes 2, 12 and 17, and *stat* genes found in *Trichoplax adhaerens*. The location of markers is indicated besides gene names between brackets when relevant. Data from genome assemblies in Ensembl release 100 (Human GRCh38. p13 and *Trichoplax adhaerens* ASM15027v1).



	Stat1-4	*Stat5-6*	*Stat1*	*Stat2*	*Stat3*	*Stat4*	*Stat5*	*Stat6*
Sea urchin		1						
Accorn worm		2						
Lancelet	1	1						
Sea squirt		3						
Oikopleura		1						
Agnathans								
Lamprey	1	1			1		1	
Chondrichthyans								
Leucoraja erinacea			1	1	1		1	1
Squalus acanthias			1	1	1		1	1
Rhincodon typus						1		
Bony fishes								
Zebrafish			2	1	1	1	2	1
Common carp			4	2	2	2	4	2
Atlantic Salmon			5	1	2	2	3	1
Rainbow trout			5	1	2	2	3	1
Chinook Salmon			5	1	2	2	3	1
Amphibians								
Xenopus tropicalis			1	1	2	1	1	1
Xenopus laevis			2	2	4	1	2	2
Other tetrapods			1	1	1	1	1 or 2	1

Fig. 8. Evolutionary pathways of stat genes in Deuterostomians. WGD are indicated by red "X", and the date indicated for the most recent events. Salmonid *stat* genes for which paralogs have not been retained are boxed in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

pfam to look for assembly problems and fragmentary sequences. MegaX software was used to carry out phylogenetic analyses and confirm the homology relationships between sequences. The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model. The bootstrap consensus tree inferred from 1000 replicates was taken to represent the evolutionary history of the taxa analysed. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Joining and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value.

4.2. Microsynteny analysis

Synteny were retrieved from Genomicus (version 100) and the orthology/paralogy relationships available in Ensembl, and complemented by visual examination of the graphical interface in both Ensembl and NCBI. A linkage was considered a conserved microsynteny only when three or more such genes were linked in such a way in two species.

4.3. Expression analysis

RNAseq transcriptome analysis on the chinook salmon STAT2-KO GS2 and CHSE-EC cell lines was described by (Dehler et al., 2019). Briefly, STAT2 KO or control CHSE-EC cells were stimulated (or not) in EMEM medium supplemented with 250 ng/ml of recombinant *O. mykiss* IFNA2. Three biological replicates (Flask 1–3) were used for library construction for each group, and RNA-Seq libraries were prepared using TruSeq Stranded mRNA Sample Preparation Kit (Illumina) according to the manufacturer's instructions. Libraries were validated for quality on Agilent DNA1000 Kit, pooled in equimolar amounts and sequenced in pair-ends 2x75 bp on Illumina NextSeq 500/550. For each library, a depth 20 M reads were generated. Reads were then spliced-aligned to 47,898 genes (47,022 Gnomon, 876 RefSeq, GCF_002163495.1_O-myk_1.0_genomic.gff from the NCBI).

Declaration of competing interest

The authors declare no commercial or financial conflict of interest.

Acknowledgements

We thank Luc Jouneau for help with transcriptome analysis and annotation. This work was funded in part by the VetBioNet project (EU Grant Agreement INFRA-2016-1 731014) and by institutional grant from INRAE. BC, PB, SB and LDP carried out *in silico* analyses. BC, PB and LDP wrote the manuscript. All authors read and approved the final manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dci.2020.103929.

Data availability

All data generated or analysed during this study are included in this published article and its supplementary information files.

References

Aaronson, D.S., Horvath, C.M., 2002. A road map for those who don't know JAK-STAT. Science 80. https://doi.org/10.1126/science.1071545.

Bhattacharya, S., Eckner, R., Grossman, S., Oldread, E., Arany, Z., D'Andrea, A., Livingston, D.M., 1996. Cooperation of Stat2 and p300/CBP in signalling induced by interferon- α. Nature 383, 344–347. https://doi.org/10.1038/383344a0.

Christoffels, A., Brenner, S., Venkatesh, B., 2006. Tetraodon genome analysis provides further evidence for whole-genome duplication in the ray-finned fish lineage. Comp. Biochem. Physiol. Genom. Proteonomics 1, 13–19. https://doi.org/10.1016/j.cbd.2005.06.001.

Copeland, N.G., Gilbert, D.J., Schindler, C., Zhong, Z., Wen, Z., Darnell, J.E., Mui, A.L.-F., Miyajima, A., Quelle, F.W., Ihle, J.N., Jenkins, N.A., 1995. Distribution of the mammalian stat gene family in mouse chromosomes. Genomics 29, 225–228. https://doi.org/10.1006/geno.1995.1235.

Decker, T., Kovarik, P., 2000. Serine phosphorylation of STATs. Oncogene 19, 2628–2637. https://doi.org/10.1038/sj.onc.1203481.

Decker, T., Kovarik, P., Meinke, A., 1997. GAS elements: a few nucleotides with a major impact on cytokine-induced gene expression. J. Interferon Cytokine Res. 17, 121–134. https://doi.org/10.1089/jir.1997.17.121.

Dehler, C.E., Lester, K., Della Pelle, G., Jouneau, L., Houel, A., Collins, C., Dovgan, T., Machat, R., Zou, J., Boudinot, P., Martin, S.A.M., Collet, B., 2019. Viral resistance and IFN signaling in STAT2 knockout fish cells. J. Immunol. 203, 465–475. https://doi.org/10.4049/jimmunol.1801376.

Gregory, L., Came, P.J., Brown, S., 2008. Stem cell regulation by JAK/STAT signaling in Drosophila. Semin. Cell Dev. Biol. https://doi.org/10.1016/j.semcdb.2008.06.003.

Jaillon, O., Aury, J.M., Brunet, F., Petit, J.L., Stange-Thomann, N., Maucell, E., Bouneau, L., Fischer, C., Ozouf-Costaz, C., Bernot, A., Nicaud, S., Jaffe, D., Fisher, S., Lutfalla, G., Dossat, C., Segurens, B., Dasilva, C., Salanoubat, M., Levy, M., Houdet, N., Castellano, S., Anthouard, V., Jubin, C., Castelli, V., Katinka, M., Vacherie, B., Blémont, C., Skalli, Z., Cattolico, L., Poulain, J., De Berardinis, V., Cruaud, C., Dupart, S., Brottler, P., Coutanceau, J.P., Gouzy, J., Parra, G., Lardier, G., Chapple, C., McKernan, K.J., McEwan, P., Bosak, S., Kellis, M., Volff, J.N., Gulgó, R., Zody, M.C., Mesirov, J., Lindblad-Toh, K., Birren, B., Nusbaum, C., Kahn, D., Robinson-Rechavi, M., Laudet, V., Schachter, V., Quétler, F., Saurin, W., Scarpeill, C., Wincker, P., Lander, E.S., Weissenbach, J., Roest Crollius, H., 2004. Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. Nature 431, 946–957. https://doi.org/10.1038/nature03025.

Kiu, H., Nicholson, S.E., 2012. Biology and significance of the JAK/STAT signalling pathways. Growth Factors 30, 88–106. https://doi.org/10.3109/08977194.2012.660936.

Lan, X., Pritchard, J.K., 2016. Coregulation of tandem duplicate genes slows evolution of subfunctionalization in mammals. Science 352 (80), 1009–1013. https://doi.org/10.1126/science.aad8411.

Levasseur, A., Pontarotti, P., 2011. The role of duplications in the evolution of genomes highlights the need for evolutionary-based approaches in comparative genomics. Biol. Direct. https://doi.org/10.1186/1745-6150-6-11.

Levraud, J.-P., Jouneau, L., Briolat, V., Laghi, V., Boudinot, P., 2019. IFN-stimulated genes in zebrafish and humans define an ancient arsenal of antiviral immunity. J. Immunol. 203, 3361–3373. https://doi.org/10.4049/jimmunol.1900804.

Lien, S., Koop, B.F., Sandve, S.R., Miller, J.R., Kent, M.P., Nome, T., Hvidsten, T.R., Leong, J.S., Minkley, D.R., Zimin, A., Grammes, F., Grove, H., Gjuvsland, A., Walenz, B., Hermansen, R.A., Von Schalburg, K., Rondeau, E.B., Di Genova, A., Samy, J.K.A., Olav Vik, J., Vigeland, M.D., Caler, L., Grimholt, U., Jentoft, S., Inge Våge, D., De Jong, P., Moen, T., Baranski, M., Palti, Y., Smith, D.R., Yorke, J.A., Nederbragt, A.J., Tooming-Klunderud, A., Jakobsen, K.S., Jiang, X., Fan, D., Hu, Y., Liberles, D.A., Vidal, R., Iturra, P., Jones, S.J.M., Jonassen, I., Maass, A., Omholt, S. W., Davidson, W.S., 2016. The Atlantic salmon genome provides insights into rediploidization. Nature 533, 200–205. https://doi.org/10.1038/nature17164.

Lin, J.-X., Leonard, W.J., 2019. Fine-tuning cytokine signals. Annu. Rev. Immunol. 37, 295–324. https://doi.org/10.1146/annurev-immunol-042718-041447.

Liongue, C., O'Sullivan, L.A., Trengove, M.C., Ward, A.C., 2012. Evolution of JAK-STAT pathway components: mechanisms and role in immune system development. PLoS One 7, e32777. https://doi.org/10.1371/journal.pone.0032777.

Liongue, C., Sertori, R., Ward, A.C., 2016. Evolution of cytokine receptor signaling. J. Immunol. 197, 11–18. https://doi.org/10.4049/jimmunol.1600372.

Morin-Poulard, I., Vincent, A., Crozatier, M., 2013. The Drosophila JAK-STAT pathway in blood cell formation and immunity. JAK-STAT 2, e25700. https://doi.org/10.4161/jkst.25700.

Nan, Y., Wu, C., Zhang, Y.J., 2017. Interplay between Janus kinase/signal transducer and activator of transcription signaling activated by type I interferons and viral antagonism. Front. Immunol. https://doi.org/10.3389/fimmu.2017.01758.

Nast, R., Staab, J., Meyer, T., 2019. Gene activation by the cytokine-driven transcription factor STAT1. In: Gene Regulation. IntechOpen. https://doi.org/10.5772/intechopen.82699.

Pasquier, J., Cabau, C., Nguyen, T., Jouanno, E., Severac, D., Braasch, I., Journot, L., Pontarotti, P., Klopp, C., Postlethwait, J.H., Guiguen, Y., Bobe, J., 2016. Gene evolution and gene expression after whole genome duplication in fish: the PhyloFish database. BMC Genom. 17, 368. https://doi.org/10.1186/s12864-016-2709-z.

Perrimon, N., Mahowald, A.P., 1986. l(1)hopscotch, a larval-pupal zygotic lethal with a specific maternal effect on segmentation in Drosophila. Dev. Biol. 118, 28–41. https://doi.org/10.1016/0012-1606(86)90070-9.

Rajan, A., Perrimon, N., 2012. Drosophila cytokine unpaired 2 regulates physiological homeostasis by remotely controlling insulin secretion. Cell 151, 123–137. https://doi.org/10.1016/j.cell.2012.08.019.

Reich, N.C., 2013. STATs get their move on. JAK-STAT 2, e27080. https://doi.org/10.4161/jkst.27080.

Session, A.M., Uno, Y., Kwon, T., Chapman, J.A., Toyoda, A., Takahashi, S., Fukui, A., Hikosaka, A., Suzuki, A., Kondo, M., Van Heeringen, S.J., Quigley, I., Heinz, S., Ogino, H., Ochi, H., Hellsten, U., Lyons, J.B., Simakov, O., Putnam, N., Stites, J., Kuroki, Y., Tanaka, T., Michiue, T., Watanabe, M., Bogdanovic, O., Lister, R., Georgiou, G., Paranjpe, S.S., Van Kruijsbergen, I., Shu, S., Carlson, J., Kinoshita, T.,

Ohta, Y., Mawaribuchi, S., Jenkins, J., Grimwood, J., Schmutz, J., Mitros, T., Mozaffari, S.V., Suzuki, Y., Haramoto, Y., Yamamoto, T.S., Takagi, C., Heald, R., Miller, K., Haudenschild, C., Kitzman, J., Nakayama, T., Izutsu, Y., Robert, J., Fortriede, J., Burns, K., Lotay, V., Karimi, K., Yasuoka, Y., Dichmann, D.S., Flajnik, M.F., Houston, D.W., Shendure, J., Dupasquier, L., Vize, P.D., Zorn, A.M., Ito, M., Marcotte, E.M., Wallingford, J.B., Ito, Y., Asashima, M., Ueno, N., Matsuda, Y., Veenstra, G.J.C., Fujiyama, A., Harland, R.M., Taira, M., Rokhsar, D.S., 2016. Genome evolution in the allotetraploid frog Xenopus laevis. Nature 538, 336–343. https://doi.org/10.1038/nature19840.

Stark, G.R., Darnell, J.E., 2012. The JAK-STAT pathway at twenty. Immunity. https://doi.org/10.1016/j.immuni.2012.03.013.

Stepkowski, S.M., Chen, W., Ross, J.A., Nagy, Z.S., Kirken, R.A., 2008a. STAT3: an important regulator of multiple cytokine functions. Transplantation. https://doi.org/10.1097/TP.0b013e3181739d25.

Stepkowski, S.M., Chen, W., Ross, J.A., Nagy, Z.S., Kirken, R.A., 2008b. STAT3: an important regulator of multiple cytokine functions. Transplantation. https://doi.org/10.1097/TP.0b013e3181739d25.

Suurväli, J., Jouneau, L., Thépot, D., Grusea, S., Pontarotti, P., Du Pasquier, L., Rüütel Boudinot, S., Boudinot, P., 2014. The proto-MHC of placozoans, a region specialized in cellular stress and ubiquitination/proteasome pathways. J. Immunol. 193, 2891–2901. https://doi.org/10.4049/jimmunol.1401177.

Uno, Y., Nishida, C., Takagi, C., Ueno, N., Matsuda, Y., 2013. Homoeologous chromosomes of Xenopus laevis are highly conserved after whole-genome duplication. Heredity 111, 430–436. https://doi.org/10.1038/hdy.2013.65.

Van de Peer, Y., 2004. Tetraodon genome confirms Takifugu findings: most fish are ancient polyploids. Genome Biol. https://doi.org/10.1186/gb-2004-5-12-250.

Vanha-aho, L.M., Valanne, S., Rämet, M., 2016. Cytokines in Drosophila immunity. Immunol. Lett. https://doi.org/10.1016/j.imlet.2015.12.005.

Villarino, A.V., Kanno, Y., O'Shea, J.J., 2017. Mechanisms and consequences of Jak–STAT signaling in the immune system. Nat. Immunol. 18, 374–384. https://doi.org/10.1038/ni.3691.

Wang, Y., Levy, D.E., 2012. Comparative evolutionary genomics of the STAT family of transcription factors. JAK-STAT 1, 23–36. https://doi.org/10.4161/jkst.19418.

Wojciak, J.M., Martinez-Yamout, M.A., Dyson, H.J., Wright, P.E., 2009. Structural basis for recruitment of CBP/p300 coactivators by STAT1 and STAT2 transactivation domains. EMBO J. 28, 948–958. https://doi.org/10.1038/emboj.2009.30.

Zeidler, M.P., Perrimon, N., Strutt, D.I., 1999. Polarity determination in the Drosophila eye: a novel role for unpaired and JAK/STAT signaling. Genes Dev. 13, 1342–1353. https://doi.org/10.1101/gad.13.10.1342.

Zhang, J.J., Vinkemeier, U., Gu, W., Chakravarti, D., Horvath, C.M., Darnell, J.E., 1996. Two contact regions between Stat1 and CBP/p300 in interferon γ signaling. Proc. Natl. Acad. Sci. U.S.A. 93, 15092–15096. https://doi.org/10.1073/pnas.93.26.15092.