HEMATOPOIETIC CELL LINEAGE SWITCHING MEDIATED BY ZEBRAFISH STAT1B

by

HAO SONG

A DISSERTATION

Presented to the Department of Biology
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2010
University of Oregon Graduate School

Confirmation of Approval and Acceptance of Dissertation prepared by:

Hao Song

Title:

"Hematopoietic Cell Lineage Switching Mediated by Zebrafish Stat1b"

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Biology by:

William Cresko, Chairperson, Biology
John Postlethwait, Advisor, Biology
Judith Eisen, Member, Biology
Jan Spitsbergen, Member, Not from U of O
J. Andrew Berglund, Outside Member, Chemistry

and Richard Linton, Vice President for Research and Graduate Studies/Dean of the Graduate School for the University of Oregon.

June 14, 2010

Original approval signatures are on file with the Graduate School and the University of Oregon Libraries.
An Abstract of the Dissertation of

Hao Song

for the degree of Doctor of Philosophy

in the Department of Biology to be taken June 2010

Title: HEMATOPOIETIC CELL LINEAGE SWITCHING MEDIATED BY ZEBRAFISH STAT1B

Approved: John H. Postlethwait

A critical question for developmental biology is the mechanism by which cells make fate decisions. In the hematopoietic system, stem cells differentiate into several different cell types, but the mechanisms that affect this process are incompletely known. Understanding these mechanisms is important because abnormal regulation of hematopoiesis can result in disease.

STAT1 protein plays crucial roles in mediating innate immunity by transducing interferon signals, but recent results have also related STAT1 to hematopoietic cell differentiation. Here we cloned a previously uncharacterized zebrafish co-ortholog of the human STAT1 gene we call stat1b and investigated the functions of two zebrafish Stat1 proteins in hematopoiesis. The advantage of the zebrafish model is that, due to a whole genome duplication (WGD), some human genes have two co-orthologs in zebrafish. During evolution, co-orthologs have retained or acquired similar, complimentary, or new
functions.

Both *stat1a* and *stat1b* encode all four characteristic domains of the human STAT1 protein. Phylogenetic and conserved synteny analyses showed that *stat1b* and *stat1a* arose as duplicates in the teleost genome duplication event, and these analyses clarified the historical origin of the entire vertebrate STAT gene family. RT-PCR demonstrated maternal expression of both *stat1a* and *stat1b*. Expression of *stat1b*, but not *stat1a*, was detected in hematopoietic domains of embryos by in situ hybridization. Morpholino knockdown of *stat1b*, but not *stat1a*, mRNA expression resulted in a decrease in expression of the myeloid cell marker genes *spi* and *mpx* and an increase in expression of the hematopoietic progenitor marker gene *scl* and the erythrocyte marker gene *gata1*. These results show that in zebrafish, Stat1b protein functions in the commitment of hematopoietic cells to a myeloid cell fate.
CURRICULUM VITAE

NAME OF AUTHOR: Hao Song

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon, USA
Beijing Medical University, Beijing, China

DEGREES AWARDED:

Doctor of Philosophy in developmental biology, University of Oregon, June 2010
Bachelor of Biomedical Science, Beijing Medical University, July 1996

AREAS OF SPECIAL INTEREST:

Developmental biology, molecular biology, cell biology
Java programming

PROFESSIONAL EXPERIENCE:

Associate scientist, GeneCore BioTechnologies Co., Ltd., Shanghai, China
09/2003–06/2004

Lecturer, China Medical University, Shenyang, China
08/1996–06/1999
GRANTS, AWARDS AND HONORS:

Cancer Federation Scholarship, 2007

PUBLICATIONS:


ACKNOWLEDGMENTS

I wish to express sincere appreciation to my advisor Dr. John Postlethwait for his five-year direction and support. The advice is always helpful. In addition, special thanks are due to professor Jan M. Spitsbergen, whose expertise in fish hematopoiesis and thoughtful discussion led me through several stages in this work. I also thank the members of the community councils for their valuable suggestions and encouragements. Finally, the lab members in Postlethwait lab are always ready to help in the wet work and discussion and deserve my highest appreciation.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1. Hematopoiesis in Zebrafish</td>
<td>1</td>
</tr>
<tr>
<td>1.2. STAT1 in Hematopoiesis</td>
<td>3</td>
</tr>
<tr>
<td>II. MATERIALS AND METHODS</td>
<td>7</td>
</tr>
<tr>
<td>2.1. Primers for Cloning statb cDNA</td>
<td>7</td>
</tr>
<tr>
<td>2.2. Phylogenetic Analysis</td>
<td>8</td>
</tr>
<tr>
<td>2.3. Zebrafish, Morpholinos, and In Situ Hybridization</td>
<td>9</td>
</tr>
<tr>
<td>2.4. RT-PCR and Quantitative RT-PCR</td>
<td>9</td>
</tr>
<tr>
<td>III. RESULTS</td>
<td>12</td>
</tr>
<tr>
<td>3.1. Isolation of a cDNA for a Zebrafish Co-Ortholog of STAT1, and Gene and Protein Structure</td>
<td>12</td>
</tr>
<tr>
<td>3.2. The Evolutionary Origin of statb</td>
<td>14</td>
</tr>
<tr>
<td>3.3. Both stat1a and stat1b Are Expressed at Early Stage of Zebrafish Embryo Development</td>
<td>19</td>
</tr>
<tr>
<td>3.4. The Expression of stat1b Is Specific in Hematopoiesis Region</td>
<td>20</td>
</tr>
<tr>
<td>3.5. Loss of stat1b Function Results in Hematopoietic Cell Lineage Switching</td>
<td>22</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Exon structures of human STAT1 alpha and beta, and zebrafish stat1a and stat1b on genome</td>
<td>14</td>
</tr>
<tr>
<td>2.</td>
<td>STAT specific domains appear in proteins encoded by the two STAT1 splicing forms of the human gene and from zebrafish Stat1a and Stat1b</td>
<td>14</td>
</tr>
<tr>
<td>3.</td>
<td>Phylogenetic tree constructed by maximum likelihood for STAT1, STAT2, and STAT3 clades, using Ciona intestinalis STATa as outgroup</td>
<td>16</td>
</tr>
<tr>
<td>4.</td>
<td>Conserved syntenies verify the history of the stat1 gene family.</td>
<td>18</td>
</tr>
<tr>
<td>5.</td>
<td>RT-PCR with stat1a and stat1b primers was run to detect their expression levels along zebrafish development</td>
<td>20</td>
</tr>
<tr>
<td>6.</td>
<td>Whole mount in situ hybridization shows specific expression pattern of stat1b in zebrafish hematopoietic regions</td>
<td>22</td>
</tr>
<tr>
<td>7.</td>
<td>In situ hybridization with blood gene markers after stat1a and stat1b MO injections.</td>
<td>24</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

1.1. Hematopoiesis in Zebrafish

The process of hematopoiesis consists of cell fate decision and lineage differentiation and maturation of all blood cells. SCL/TAL1 is a helix-loop-helix transcription factor that determines the commitment of hemangioblast, the common ancestor of endothelial cells and blood cells, to hematopoiesis stem cells (HSC) (Forrai and Robb, 2003; Xiong, 2008). Absence of SCL/TAL1 was embryonic lethal in knockout mice due to complete and early block in embryonic erythropoiesis (Shivdasani et al., 1995). For each of the three main blood cell lineages, erythroid, myeloid, and lymphoid, several important transcription factors exist for regulating the development and are used as markers for specific cell lineages. Among them, GATA1 is a zinc finger protein essential for erythroid cell differentiation. GATA1 binds to its specific DNA sites and recruits co-factors to promote transcription of downstream genes, including hbb1 for hemoglobin (Fujiwara et al., 1996; Crispino, 2005). PU.1/SP11 is an Ets-family transcription factor and promotes expression of various myeloid specific genes including the M-CSF receptor (Oikawa et al., 1999; Kastner et
IKAROS, a hemopoietic-specific zinc finger transcription factors, is necessary for both embryonic lymphoid cell fate decision and adult T-/B- cell maturation, cooperating with many factors, especially AIOLOS (Georgopoulos, 1997; Schmitt, 2002). Nevertheless, the detailed mechanisms of their functions and the intricate networks of co-factors remain to be revealed. Besides, MPO/MPX (myeloperoxidase) is a heme-containing enzyme mainly synthesized during neutrophils maturation and executes the microbicidal activity of neutrophils by generating highly active free radical species (Hampton et al., 1998; Klebanoff, 2005) and thus is used as a maker for differentiated neutrophils.

Zebrafish is a good model in hematopoiesis research. Vertebrates have two waves of hematopoiesis, primitive (embryonic) hematopoiesis and definitive (adult) hematopoiesis. The transcriptional mechanisms of hematopoietic stem cell (HSC) differentiation are evolutionarily conserved among vertebrates in both hematopoietic waves, although the location varies among species (de Jong and Zon, 2005). In zebrafish, primitive hematopoiesis begins around 10 hours post fertilization (hpf) when embryonic HSC differentiate from the ventral lateral plate mesoderm (Davidson et al., 2003). Primitive HSC reside at two sites in zebrafish: the intermediate cell mass (ICM) and the rostral blood island (RBI). The ICM is located in the trunk ventral to the notochord and contains stem cells that differentiate mainly into pro-erythroblasts. The RBI develops from the cephalic mesoderm and produces mainly macrophages (Long et
al., 1997). Primitive hematopoiesis generates mainly erythrocytes. At about 24hpf, about 300 erythrocytes enter the blood stream from the ICM (Patterson et al., 2005). Primitive hematopoiesis lasts until about 32hpf, and about an hour earlier, definitive HSC begin to appear in the ventral wall of the dorsal aorta in a region known as the aorta-gonad-mesonephros (AGM) (Burns et al., 2005). Hematopoiesis continues in the AGM from 30hpf to 40hpf, when the kidney marrow takes over as the site of lifelong definitive hematopoiesis, analogous to the bone marrow in tetrapods. Cytospin preparations of kidney marrow reveal that all circulating hematopoietic blood cell types are present from 40hpf onwards (the embryo hatches on day 3) (Traver, 2004).

1.2. STAT1 in Hematopoiesis

Hematopoiesis may also involve STAT1 (signal transducer and activator of transcription-1). The lack of $FANCC$ activity leads to Fanconi Anemia, a condition of hematopoietic bone marrow failure and leukemia (Joenje and Patel, 2001; Wang, 2007), and failure to activate optimally STAT1 in response to cytokines or growth factors (Pang et al., 2000; Pang et al., 2001). STAT1 mediates signaling of both type I and type II interferons (IFNs) in the innate immune response against viral infection through an IFN-JAK-STAT relay (Schindler et al., 1992). IFNs bind to IFN receptors, thereby activating JAKs, which then phosphorylate STAT1. Phosphorylated STAT1 translocates from the cytosol into the nucleus, and in minutes promotes changes in gene
expression (Stark et al., 1998). Now we know STAT1 can transduce signals in response to a variety of cytokines and growth factors (Najjar et al., 2010a) and regulate different cellular functions, mainly in immunity system. In addition to by type I and type II IFNs to fight against opportunistic virus infections (Meraz et al., 1996; Durbin et al., 1996), growth factors including EGF (Chin et al., 1996) and PDGF (Vignais et al., 1996) have been shown to activate STAT1. TNF has anti-viral function and can activate STAT1, although TNF is mainly released in response to bacterial lipopolysaccharide (LPS) rather than viruses (Singh K et al., 2010). Moreover, STAT1 is activated by IL-4 for cell growth inhibition and T cell maturation (Chang et al., 2000; Acacia et al., 2007); by IL-6 to inhibit cartilage growth (Legendre et al., 2003); by IL-7 to initiate T cell activity (Rose et al., 2010); by IL-9 for cell growth arrest and differentiation (Demoulin et al., 1999; Demoulin et al., 2001); by IL-10 to activate monocytic cells (Rahimi et al., 2005); by IL-11 to protect endothelial cells and trigger gastric tumorigenesis (Mahboubi et al., 2000; Ernst et al., 2008); by IL-13 to induce degranulation of mast cells leading to asthma (Chiba et al., 2009); by IL-21 to promote cell apoptosis (de Totero et al., 2008); by IL-22 to increase cultured cell line proliferation (Lejeune et al., 2002; Brand et al., 2006); and by IL-27 to regulation T cell differentiation and mediate inflammation through monocytes activation (Kalliolas and Ivashkiv, 2008; Ouaked et al., 2009). STAT1 also plays a role in the humoral immunity, promoting the expression of the B-cell receptor in the lymphoid lineage (Najjar et al., 2010b) and helping neutrophil
granulocytes from the myeloid lineage fight against mycobacterial infection (Al-Muhsen et al., 2008; Bussmeyer et al., 2010). In addition, STAT1 is involved in the activation of dendritic cells and macrophages in antigen presentation (Cantrell et al., 2009). STAT1 might even promote late-stage melanoma progression (Schultz et al., 2010). Many of these various functions of STAT1 could be related to its effect on cell growth retardation and apoptosis. For example, activation of STAT1 is involved in renal cells damage in hyperglycemia (Ortiz-Muñoz et al., 2010), pathophysiological change in the CNS (Hofer et al., 2010) and osteoblast differentiation inhibition in bone (Tajima et al., 2010). But several showed the opposite, especially in carcinogenesis and tumor metastasis.

Although a substantial amount of work on STAT1 has occurred in mature tissue culture cells, comparatively little study has been directed towards the roles of STAT1 in embryonic development. The zebrafish *Danio rerio* (Hamilton, 1822) has several advantages for an investigation of the roles of STAT1 in embryonic development. Due to a whole genome duplication event at the base of the teleost radiation, some single copy genes in human have two orthologs in zebrafish (Amores et al., 1998; Amores et al., 2004; Postlethwait et al., 2004). Such duplicates often share between them ancestral gene functions that are conserved with the mammalian genes (Force et al., 1999). Thus, the knock down of just one of the duplicates by morpholino (MO) antisense oligonucleotide while keep the other functional can reveal gene functions that might be
difficult to resolve in mammals due to pleiotropy. Our work on the two zebrafish STAT1 co-orthologs provides new insights into the roles of STAT1 in hematopoietic development. The stat1a gene of zebrafish has been shown to rescue interferon-signaling mediated cell growth inhibition in a STAT1-deficient human cell line (Oates et al., 1999). Here we report the cloning of stat1b, an additional stat1 co-ortholog in zebrafish. We show that stat1b arose in the teleost genome duplication event, that it is expressed in the embryonic hematopoietic domain, and that it promotes blood cell lineage switching in primitive hematopoiesis development.
CHAPTER II
MATERIALS AND METHODS

2.1. Primers for Cloning stat1b cDNA

Primers covering overlapping fragments of XR_029517 were designed as follows for RT-PCR, with two pairs of primers for each fragment. PCR products were sequenced with the TOPO TA cloning kit into the pCR4-TOPO vector (Invitrogen, Carlsbad, CA, USA) and sequences were assembled to obtain full-length cDNA sequence.

stat1bF10  CGTTGAAAGATGACGCTCTG
stat1bR10  ACGGATCTGCTGGAGACACT
stat1bF11  GCTTATCCCGAGATACACTCC
stat1bR11  TCTTCAGCTGCTGACGGATCT
stat1bF20  GCTGAAGGTGTGCAAATGAA
stat1bR20  AGGCACTGGGTAACTGGTTG
stat1bF21  GCTGAAGGTGTGCAAATGAA
stat1bR21  TTCACCGCTGAGCATGTTGTA
stat1bF30  GACGCCTCTCATCGTTACAGA
2.2. Phylogenetic Analysis

Related STAT protein sequences were downloaded from NCBI and Ensembl and aligned by Muscle (MUltiple Sequence Comparison by Log-Expectation, http://www.ebi.ac.uk/Tools/muscle/index.html) using the JTT model with ProtTest (Abascal et al., 2005). Finally, all information was submitted to PhyML (Guindon and Gascuel, 2003; http://www.atgc-montpellier.fr/phylml/) for construction of the phylogenetic tree.
2.3. Zebrafish, Morpholinos, and In Situ Hybridization

Zebrafish (ABC X TU strain) were maintained as described (Kimmel et al., 1995). The sequence of \textit{stat1b} morpholino (MO) was AAAATGTAGCGGATGTTACTTCGAC (Gene Tools, LLC, Philomath, OR, USA), which targets the splice donor site of exon5-intron5, leading to the inclusion of intron 5 and a premature translation stop codon. The sequence of \textit{stat1a} MO is TCATGTGGTCAACAGGCACCTGCAA, which targets the splice donor site of exon3-intron3, leading to the inclusion of intron 3 and a premature translation stop. The sequence of the control MO was CCTCTTACCTCAGTTACAATTTATA, whose sequence has no significant similarity found in the zebrafish reference genome sequence database. Embryos were collected at 16hpf, 24hpf, 32hpf, and 48hpf for in situ hybridization with zebrafish \textit{stat1a} probe, \textit{stat1b} probe, and a negative control (cuttlefish, GenBank accession number GU388435). In situ was carried out following methods previously described (Yan et al., 2002) except that probe incubation was done at 50 °C. MO (5 mg/ml for \textit{stat1b} MO and 10 mg/ml for \textit{stat1a} MO) was injected into one-cell stage embryos. In situ hybridization with blood markers was carried out at 16hpf and 32hpf following methods previously described (Yan et al., 2002).

2.4. RT-PCR and Quantitative RT-PCR

Fertilized eggs were collected at indicated time points. Total RNA was extracted
with TRI reagent (Molecular Research Center, Cincinnati, OH, USA) and genomic DNA was removed with DNA-free™ DNase Treatment and Removal Reagents (ABI, Foster City, CA, USA). Further total RNA sample was quantified with spectrometry at 260nm and loading amount to RT-PCR was calibrated to be the same. Product from each pair of primers was sequenced and verified to be the target mRNA fragment before the actual RT-PCR. RT-PCR was run with OneStep RT-PCR kit (QIAGEN, Valencia, CA, USA). Primers used are listed below:

actinF GCCAACACAGTGCTGTCTGGAGG
actinR GGTCATGGACGCCCATTGTGAGG
stat1aMEF AGTCGCAGCAATGACTCAGTG
stat1aMER C7GCTGATGATCATCGCCATTG
stat1bMEF AGGTGACTCCATGCAGGGGAAC
stat1bMER GCGGGTCGTTGCTGTAGGTGA

cDNAs were prepared with the High Capacity RNA-to-cDNA kit (ABI, Foster City, CA, USA). Quantitative PCR was carried out using Power SYBR® Green PCR Master Mix (ABI, Foster City, CA, USA). Q-PCR requires biological repeats and sampling repeats. To make the level of RNA comparable among different samples, the expression level of a house-keeping gene zgc:136952 (NM_001040043), which encodes a ribosomal protein, was run as an internal control. Primers used are listed below:
gata1qF GAATGCAGCTTCAGAGGTTTATCC
gata1qR TGGGTTCAGAGAATACGCTCCTA
mpxqF CCAAACCTCAGGGATGTTCTTG
mpxqR CCCAAACTACGAGTCCCTATGC
spiqF GGGTAGCCATCACATCCCTCTTAG
spiqR TGGACGTTGTGAGGGTAACACA
sclqF GACTAATTTCGCGGCTGACAACCTA
sclqR GGGCGTAAACAGAAAGTCTTACGTA
RPL32qF CCCTCACCAACCTAAGATCGT
RPL32qR CTCCAGTTTCGCCCTGATCTTG
CHAPTER III
RESULTS

3.1. Isolation of a cDNA for a Zebrafish Co-Ortholog of STAT1, and Gene and Protein Structure

To search for a possible duplicate of the described zebrafish stat1a gene, we used the stat1a gene sequence (RefSeq NM_131480) in a BLAST search against the zebrafish EST database and identified the gene model sequence XR_029517 as the most similar sequence. We designed primers for overlapping fragments of the EST sequence and used them to amplify cDNA from a total RNA sample of mixed 24hpf and 48hpf embryos from an ABC X TU cross, and then combined sequences of the cDNA fragments bioinformatically. We identified a 3,569bp nucleotide sequence that we called statlb and submitted it to GenBank (accession number FJ986224). Comparing the structure of this cDNA with the genomic DNA at Ensembl shows 23 exons. The translated exons of human STAT1 and zebrafish stat1a and statlb are orthologs. The untranslated exon 1 in human STAT1 was lost in both zebrafish stat1 genes and exon 2 was partially kept. At 3' end, exon 24 to 26 in stat1a encodes the STAT domain (see below and Fig. 2) just like what exon 24 in human STAT1 does (Fig. 1). The statlb gene
should be translated into a 725-amino-acid peptide (accession number ACR83062).

As predicted by NCBI BLASTP, the translated peptide Stat1b has four (Int, Alpha, Bind, and SH2 superfamily) of the five specific domains that are found in human STAT1 (Fig. 2). The two domains at the N-end, STAT-int and STAT-alpha, are responsible for the interaction between STAT proteins after they bind DNA sites, usually enhance their transcription activation effects; the STAT-bind domain recognizes and binds specific DNA fragments, determining the specificity of STAT proteins; SH2 domain, or Src Homology 2, near to the C-end, bind a phosphorylated tryosine residue in IFN receptors or other STAT proteins, thus recruiting STATs to IFN receptors or forming dimers with other STAT proteins, respectively. The human STAT1 gene has two splice forms, STAT1-α and STAT1-β. The difference between the two splicing isoforms is that STAT1-α has an extra domain at the C-end, which, interestingly, also appears in zebrafish Stat1a C-end. This domain is approximately 20 amino acids in length and binds selectively to the TAZ2 domain of CRB (CREB-binding protein) to act as a transcriptional activator. Taken together, we conclude the newly cloned gene is indeed a member of STAT family.
Fig. 1. Exon structures of human STAT1 alpha and beta, and zebrafish stat1a and stat1b. Exons are represented as labeled unfilled boxes (5' and 3' un-translated regions (UTRs)) or filled boxes (translated regions). The width of boxes is proportional to the actual length of each exon. Introns are represented with thin lines connecting exons and their length is indicated in kilobases but is not proportional to length.

Fig. 2. STAT specific domains appear in proteins encoded by the two STAT1 splicing forms of the human gene and from zebrafish Stat1a and Stat1b. Five domains are represented as colored boxes. The length of each domain is indicated with a scale of numbers of amino acids on the top. Linker regions of domains in the proteins are represented as white boxes.

3.2. The Evolutionary Origin of stat1b

To help understand the origin of stat1b, we conducted a phylogenetic analysis of full-length STAT proteins using maximum likelihood (Guindon and Gascuel, 2003) rooted on the STATa protein of the urochordate Ciona intestinalis (Fig. 3). The STAT1, STAT2, and STAT3 clades are well supported with high bootstrap values (470/500 for
STAT1, 500/500 for STAT3, and 490/500 for STAT2). The newly cloned statlb sequence clearly belongs to the STAT1 clade. Note that the teleost Stat1b and Stat1a clades do not branch as sisters basal to the tetrapod STAT1 clade. Because the bootstrap value for the zebrafish Stat1a branch is lower than 60%, the node was collapsed. The evidence from this tree topology neither strongly supports nor rules out the hypothesis that statla and statlb arose in the teleost genome duplication event. We conducted BLAST searches with statlb sequence against genome databases and found statlb sequences for only two species, zebrafish and goldfish, both of which are Cypriniform fishes. In contrast, statla sequences were found not only in Cypriniforms, but also in Percomorph fish, including stickleback, pufferfish, and medaka. It appears that the statlb gene was lost in the Percomorph lineage.

The phylogenetic analysis showed that the duplication event that produced STAT1, STAT2, and STAT3 clades occurred before the divergence of tetrapod and teleost genomes. To determine the historical relationships of these genes, we used the Synteny Database (Catchen et al., 2009) to identify human paralogs of genes in a 20Mb interval surrounding STAT1 on human chromosome 2 (Hsa2). The algorithm marches down Hsa2 gene by gene, identifies paralogs, and directly above or below the Hsa2 gene, plots a plus
Fig. 3. Phylogenetic tree constructed by maximum likelihood for STAT1, STAT2, and STAT3 clades, using Ciona intestinalis STATTa as outgroup. Bootstrap values are shown as the percentage of a data set of 500 trials. The newly cloned sequence clearly belongs to the STAT1 clade. Protein sequences used in tree construction:

- Chicken.STAT1: NP_001012932
- Chicken.STAT3: NP_001026102
- Ciona.STATa: BAE06716
- Goldfish.Stat1: AA088245
- Human.STAT1: NP_009330
- Human.STAT2: NP_005410
- Human.STAT3: NP_644805
- Medaka.Stat3: AAT64912
- Mouse.STAT1: NP_033309
- Mouse.STAT2: NP_064347
- Mouse.STAT3: NP_035616
- Pufferfish.Stat1: AAL09414
- Pufferfish.Stat3: AAL09415
- Stickleback.Stat1: GENSCAN00000019309
- Stickleback.Stat3: GENSCAN00000017667
- Fugu.Stat1: ENSTRUP0000000223
- Zebrafish.Statla: NP_571555
- Zebrafish.Statlb: ACR83062
- Zebrafish.Stat3: AAH68320

sign on the human chromosome that has the paralog. Figure 4A shows that Hsa2 has numerous paralogs on Hsa3, 7, 12, and 17. The most parsimonious explanation is that these four chromosome segments (Hsa2, 3+7, 12, 17) arose in two rounds of whole genome duplication at the base of the vertebrate radiation (Dehal and Boore, 2005).
These events would have produced a fourth STAT1-related paralog that is now missing from all studied genomes. The Ensembl database shows that STAT1 lies adjacent to STAT4 and both genes are transcribed in the same orientation (Fig. 4B), suggesting that they arose by a tandem duplication event. STAT3 lies adjacent to STAT5A, which is next to STAT5B, with STAT3 and STAT5B transcribed in the same orientation. Finally, STAT3 is very near STAT6, only about 750kb distant, on Hsa12. We conclude that the STAT gene family arose by a tandem duplication to give a STAT123 gene and a STAT456 gene in pre-vertebrate chordates. In the R1 and R2 rounds of genome duplication that preceded the vertebrate radiation, four copies of this tandem gene pair were produced, STAT1 STAT4, STAT2 STAT6, STAT3 STAT5, and a fourth pair that is now missing. A subsequent tandem duplication event produced STAT5A and STAT5B in the human lineage after it diverged from the chicken lineage.

Evidence from the phylogenetic analysis was not conclusive regarding the origin of stat1a and stat1b with respect to the teleost genome duplication event. To clarify the origin of the stat1a and stat1b gene pair, we used the Synteny Database (Catchen et al., 2009) to search for regions in the zebrafish genome with conserved synteny to the STAT1 region on Hsa2. Results showed that stat1b lies next to stat4 on Dre9 along with about 300 other genes with conserved synteny between the two chromosomes, only a few of which are shown in Fig. 3B. Except for two inversions, genes in human and zebrafish are in the same order in these regions, including stat1b and stat4 and near
neighbors. Likewise, several orthologous pairs of genes with conserved synteny are shared between the human \( STAT1 \) region and the region of Dre22 that contains \( stat1a \) (Fig. 4). An inversion breakpoint is in the location expected for a duplicate copy of \( stat4 \), suggesting that the inversion event may have destroyed the \( stat4 \) copy on Dre22, which would have had no phenotypic penalty if its functions were fully redundant with the \( stat4 \) copy on Dre9. Because the orientation of transcription is \( stat4 > stat1b > \), an inversion event between the genes could have broken \( cis \)-acting regulatory elements located in the intergenic region, which is 5' to \( stat1b \). We conclude that \( stat1a \) and \( stat1b \) arose in the teleost genome duplication event.

3.3. Both \( stat1a \) and \( stat1b \) Are Expressed at Early Stage of Zebrafish Embryo Development

A gene’s expression pattern provides important hints about its function. We followed the time course of \( stat1a \) and \( stat1b \) expression by RT-PCR using total RNA samples from zebrafish embryos of various ages. To calibrate this semi-quantitative technique, we made dilutions of the template and used just 21 cycles in PCRs (Fig. 5a). Results showed that the intensity of the amplified band depended on template concentration. When we used the same conditions on embryo RNAs, we found strong and steady expression of \( stat1a \) from 0.5hpf until 96hpf, about 1 day after hatching, as well as in unfertilized eggs. These results suggest a maternal expression pattern followed by
continued zygotic expression (Fig. 5b). In contrast, *stat1b* was not maternally expressed and started low expression around 0.5hpf but rose until about 9hpf and continued strong thereafter (Fig. 5b).

Fig. 4. Conserved syntenies verify the history of the *stat1* gene family. A. Paralogy analysis of 20Mb in the human genome surrounding *STAT1*. Human paralogs of genes residing in the region of Hsa2 between 180 and 200Mb were plotted on the rest of the human chromosomes with gene order according to the location on human chromosome 2 (Hsa2) using the Dotplot feature of the Synteny database (Catchen et al., 2009). Results showed that Hsa2 genes showed extensive paralogy with Hsa7, 12, and 17. *STAT1* and its paralogs *STAT2* and *STAT3* are circled. These results would be expected if these chromosome segments arose by two rounds of whole genome duplication. B. Conserved syntenies around human *STAT1* were searched in the zebrafish genome using the Synteny Database (Catchen et al., 2009). Extensive conserved syntenies were found around *stat1b* on zebrafish chromosome 9 (Dre9) and *stat1a* on Dre22. Lines connect orthologs. The positions shown in black rectangles on the chromosomes in the upper right are blown up in the main figure. Arrows indicate the positions of *stat* genes. Each box represents a gene. Genes: HUMAN, dre9gene, dre22gene, STAT4, stat4; STAT1, stat1b, stat1a; GLS, gls, gls; NAB1, NP_001116745.1, st:dkeyp-84a8.1; TMEM194B, B0UYT4; MFSD6, zgc:92925, st:dkey-188p4.2; HIBCH, hibch; MSTN, mstn, gdf8; PMS1, pms1; ORMDL1, ormdl1; OSGEPL1, osgepl1; ANKAR, ankar; slc40a1, slc40a1; WDR75, wdr75. The following genes are HUMAN, dre22gene: RQCD1, rqcd1; PLCD4, plcd4a; PTH2R, st:dkeyp-4h4.1; FAM119A, zgc:110528; FAM119A; AOX1, aox1; BZW1, bzw1a.
Fig. 5. A). Control experiment with a gradient of RNA concentrations and 21 amplification cycles. Results verified that these RT-PCR conditions can show the relative amount of target mRNA semi-quantitatively. B). RT-PCR with stat1a and stat1b primers were run to detect their expression levels along zebrafish development. The gene of beta actin was used as loading control. RNA samples from zebrafish embryos were extracted at different time points as indicated in the lanes. –, negative control; u, unfertilized eggs; 0.5, 2, et al., time points at hours post fertilization. While the expression level of stat1a mRNA is steady, that of stat1b varies. Especially, stat1a is clearly maternal expressed, while stat1b is turned on after fertilization.

3.4. The Expression of stat1b Is Specific in Hematopoiesis Region

To identify the spatial localization of stat1 transcripts, we used whole mount in situ hybridization. Confirming a previous report (Oates et al., 1999), we did not find a
specific expression pattern for stat1a (data not shown), despite detecting it readily in RT-PCR studies. In contrast, stat1b RNA probes showed strong and restricted expression in the hematopoietic domain. Expression of stat1b appeared in the characteristic two stripes that flank the paraxial mesoderm of the posterior embryo starting at the 2-somite stage (about 11hpf) (Fig. 6CD) and meet anteriorly at the 18-somite stage (Fig. 6EF). Most cells in that region at 16hpf are erythroid progenitors (de Jong and Zon, 2005). At 24hpf, the expression pattern of stat1b is in the ICM (Fig. 6G), where most erythrocytes reside at this time. Later, at 32hpf and 48hpf, stat1b expression advanced to the heart and the ducts of Cuvier, which form after 25hpf and contain both erythrocytes and myeloid cells (Fig. 6HI). The expression patterns of stat1b are consistent with hematopoietic marker genes, including gata1, gata2 and hbbe (Detrich et al., 1995).
Fig. 6. Whole mount in situ hybridization shows specific expression pattern of stat1b in zebrafish hematopoietic regions. A. and B. No expression was detected at 3hpf and 9hpf. C. and D. First expression pattern of stat1b was detected at 2 somites and the expression was constrained to stripes that flank the paraxial mesoderm, shown with side and dorsal, respectively. E. and F. stat1b was expressed in internal cell mass at 18 somites, where hematopoietic progenitor cells reside, shown with side and dorsal, respectively. G. At 24hpf embryos, the expression region of stat1b is merging into one stripe. H. At 32hpf, in addition to internal cell mass, stat1b expression was found in the heart and ducts of Cuvier, a blood vessel network on yolk. I. At 48hpf, the expression of stat1b keeps on in the hematopoietic regions. Abbreviations: dc, ducts of Cuvier; h, heart; icm, intermediate cell mass; s, stripes flanking the paraxial mesoderm.

3.5. Loss of stat1b Function Results in Hematopoietic Cell Lineage Switching

The expression domain of stat1b in zebrafish ICM suggested the hypothesis that stat1b plays a role in primitive hematopoietic development. To knockdown activity of stat1a and stat1b in developing zebrafish embryos, we injected one-cell fertilized eggs with
either stat1a or stat1b MO, and then queried whether various lineages of blood cells developed differently by following the expression of lineage-specific gene markers. The *scl/tall* gene is a marker for primitive hematopoietic progenitor cells. Results showed that the *scl/tall* expression domain in 16hpf embryos injected with stat1b MO was broader than normal with more cells expressing this gene. In contrast, animals injected with the *stat1a* MO were normal (Fig. 7A-C). Quantitative PCR analysis confirmed the increased expression level of *scl/tall* in stat1b knockdown animals (Fig. 7G). We conclude that *stat1b*, but not *stat1a*, normally narrows the stripe of *scl/tall*-expressing hematopoietic progenitor cells. The *gata1* gene is a marker for cells in the primitive erythroid lineage. Knockdown of *stat1b* caused an elevation of *gata1* expression levels in both 16 and 32hpf embryos as detected by in situ hybridization experiments, and this result was confirmed by Q-PCR (Fig. 7H-N). The *pu.1/spi1* gene provides a marker for myeloid cells in general, and *mpo/mpx* is a marker of heterophil granulocytes. Both markers were depressed both in the in situ hybridization assays and in the Q-PCR assays (Fig. 7M-Y). We conclude that Stat1b normally promotes myeloid differentiation but depresses that of erythroid cells and those two cell lineages compete with each other for the same pluripotential progenitor cell group. Analysis of expression by in situ hybridization for *c-myb* (a marker of definitive HSC), *l-plastin* (a marker of macrophages), and *rag1* (a marker of lymphoid cells) were also performed, but no significant difference of those markers was found in *stat1b* or *stat1a* MO embryos and
controls (data not shown).

Fig. 7. In situ hybridization and quantitative PCR with blood gene markers after stat1a and stat1b MO injections. 16hpf embryos injected with control sequence MO (A), stat1a MO (B), and stat1b MO (C). The expression of scl/tall in 16hpf embryos was elevated by stat1b MO and confirmed with quantitative PCR (G). Elevation of scl/tall expression in 32hpf embryos by stat1b MO was found by in situ (D-F) but not consistent with q-PCR results (G). In situ with gata1 probe showed that the expression was elevated by stat1b MO at both 16hpf (J) and 32hpf (M) and the results were confirmed by q-PCR (N). Whereas the stat1a MO had no effect (I and L) as those of control MO (H and K). On the contrary, the depression of pu.1/spil expression by stat1b MO was found in both 16hpf (O) and 32hpf (T) and confirmed with q-PCR (U). Again, the effects of stat1b MO (P and S) were similar to those of control MO (O and R). Finally mpo/mpx expression in stat1b knockdown embryos was suppressed dramatically at 32hpf (X) and the q-PCR suggested the same effect (Y). But the effect of stat1b MO (W) was similar to those of control MO (V). Taken together, those results suggested that stat1b is involved in promoting a lineage switch from erythroid cells to myeloid cells.
We cloned and investigated the functions of stat1b, a previously unidentified zebrafish co-ortholog of the human STAT1 gene. Phylogenetic analysis supports the conclusion that the newly cloned gene belongs to the STAT1 family with high bootstrap value (Fig. 3) and conserved syntenic analysis shows that the human and zebrafish genomes have been conserved in the regions of both STAT1 duplicates (Fig. 4). Only one other stat1b gene was identified in the NCBI database, a gene in gold fish, like zebrafish, a cyprinid fish. An explanation for this gene distribution could be a missing ohnolog (Postlethwait, 2007). Because most gene duplicates become pseudogenes, only about 20% genes in human genome have two co-orthologs in zebrafish (Postlethwait et al., 2000).

The alignment of translational exons of human STAT1 and zebrafish stat1a and stat1b showed highly conserved exons structures (Fig. 1). The first non-coding exon in human STAT1 was lost in both zebrafish stat1, which might be due to no selection pressure in that part. Four STAT specific domains appear in both Stat1a and Stat1b, with an extra domain at the C-end of Stat1a. The same scenario happens in two human
STAT1 splicing forms, of which STAT1-alpha has the extra domain (Fig. 2). It is interesting to postulate that some evolution pressure selected the splicing variants in human STAT1 while zebrafish may have evolved two different genes for a similar purpose after the WGD. While it is difficult to investigate the functional difference between the two human STAT1 splicing variants in vivo, zebrafish Stat1a and Stat1b provide us with an alternative way to investigate gene functions.

The stat1a and stat1b genes differed in several key respects. First, the in situ expression pattern of stat1b in the zebrafish hematopoietic region is a sharp contrast to that of stat1a: we and others (Oates et al., 1999) could not detect a specific expression pattern for stat1a by whole mount in situ hybridization. Because RT-PCR revealed substantial transcript for both stat1a and stat1b (Fig. 5), it may be that stat1a is expressed at low levels ubiquitously and that we detect stat1b expression by in situ hybridization due to its concentration in the hematopoietic system. That difference supports subfunction partitioning of two co-orthologs due to WGD. By conventional model of anti viral infection process, interferons were secreted by viral infected cells and then affected any surrounding cell through IFN-JAK-STAT1 pathway to invoke anti viral reaction. Stat1a in zebrafish might be the one against viral infection like revealed by Oates et al. in 1999 and Stat1b is involved in zebrafish hematopoiesis.

Second, our knockdown studies with anti-sense Morpholino (MO) did not suggest a role of Stat1a in blood cell lineage switching, even though RT-PCR experiments
showed high levels of maternal stat1a expression. Either of two possibilities might explain this result. First, stat1a may not play a role in blood development, or second, the splice-targeting MO for stat1a, which will not act on maternal message RNA, may not block the effect of maternal stat1a. A translation start site targeting MO for stat1a might resolve this question, but without a suitable antibody for zebrafish Stat1a, we could not verify the stat1a efficacy of the translation MO.

The specific expression pattern of stat1b in hematopoietic regions suggests that it plays a role in zebrafish hematopoiesis. Loss-of-function of stat1b by MO injection resulted in the increased expression of scl and gata1 in primitive hematopoiesis but decreased expression of spi and mpx in both experiments of in situ and quantitative PCR. Because in situ expression analyses showed broadened expression regions, but not condensed expression in cells, we conclude that the increased expression of scl and gata1 is due to increased cell numbers of corresponding blood lineage. The scl gene marks primitive hematopoietic stem cells (HSC) and gata1 marks erythroid cells, whereas spi and mpx mark myeloid cell lineages. Along zebrafish development, erythrocytes are predominant initially, but later several myeloid cell lineages appear. Thus, our results suggested that 1) both RBC and neutrophils compete for the same pluripotential progenitor cell group; 2) although a marker for progenitor cells, scl biased to RBC differentiation; several other markers exist for primitive HSCs and they might be responsible for other cell lineage differentiation; and 3) stat1b promotes lineage
switching from erythroid (gata1-expressing) cells to myeloid (spi- and 
mpx-expressing) cells. Note that the result from knocking down stat1b is 
complementary to the result from knocking down gata1, which leads to an increase of 
spi and mpx expression in zebrafish but did not affect other cell lineages (Galloway et 
al., 2005) (Rhodes et al., 2005).

In addition, that result is consistent with recent experiments on mice that IFN-γ 
produced by Th1 cells stimulates a significant expansion of LSK, a specific HSC 
lineage, which tends to differentiate into myeloid lineages, through a STAT1 dependent 
pathway (Zhao et al., 2010). They showed that IFN-γ induced depletion of total 
progenitor cells, but promoted LSK cells proliferation capacity into myeloid cells, 
which compensate for the loss of total progenitors. IFN-γ has been known for long to be 
capable for regulating hematopoietic cell proliferation (Raefsky et al., 1985) but its 
effects are complicated. Depending on the stages and lineages of hematopoiesis, IFN-γ 
either inhibits colony formation through inducing over apoptosis (Young et al., 1997) 
(Shimozato et al., 2002) or promotes certain type of stem cell expansion and 
differentiation into myeloid cells (Caux et al., 1992) (Brugger et al., 1993). In all of 
those experiments, STAT1 is the important mediator for IFN-γ function. In our 
experiments, Stat1b could act in any of a number of ways. It might(1) promote the 
apoptosis of erythroid cells; 2) inhibit the differentiation of progenitor cells into 
erthrocytes; or 3) promote the differentiation of progenitor cells into myeloid cells.
Taking into account the results mentioned above and the function of STAT1 in inducing apoptosis, we prefer the first hypothesis but cannot rule out the other two possibilities. Understanding the different regulatory effects of IFN and STAT1 in different hematopoietic cell lineages and stages would help to shed light on the mechanism of those two interesting proteins and to better manipulate them in clinical practice for the welfare of human patients. Meanwhile zebrafish represents a new model to dissect the intricate functions of STAT1 with function divergence between zebrafish Stat1a and Stat1b.
REFERENCES


