Remarkable Conservation of Distinct Nonclassical MHC Class I Lineages in Divergent Amphibian Species

Ana Goyos, Jessica Sowa, Yuko Ohta and Jacques Robert

*J Immunol* 2011; 186:372-381; Prepublished online 29 November 2010; doi: 10.4049/jimmunol.1001467

[http://www.jimmunol.org/content/186/1/372](http://www.jimmunol.org/content/186/1/372)

**Supplementary Material**  [http://www.jimmunol.org/content/suppl/2010/11/29/jimmunol.1001467.DC1](http://www.jimmunol.org/content/suppl/2010/11/29/jimmunol.1001467.DC1)

**References**  This article cites 46 articles, 11 of which you can access for free at: [http://www.jimmunol.org/content/186/1/372.full#ref-list-1](http://www.jimmunol.org/content/186/1/372.full#ref-list-1)

**Why The JI?**  Submit online.
- **Rapid Reviews! 30 days** from submission to initial decision
- **No Triage!** Every submission reviewed by practicing scientists
- **Fast Publication!** 4 weeks from acceptance to publication

*average

**Subscription**  Information about subscribing to *The Journal of Immunology* is online at: [http://jimmunol.org/subscription](http://jimmunol.org/subscription)

**Permissions**  Submit copyright permission requests at: [http://www.aai.org/About/Publications/JI/copyright.html](http://www.aai.org/About/Publications/JI/copyright.html)

**Email Alerts**  Receive free email-alerts when new articles cite this article. Sign up at: [http://jimmunol.org/alerts](http://jimmunol.org/alerts)

*The Journal of Immunology* is published twice each month by The American Association of Immunologists, Inc., 1451 Rockville Pike, Suite 650, Rockville, MD 20852
All rights reserved.
Print ISSN: 0022-1767 Online ISSN: 1550-6606.
Remarkable Conservation of Distinct Nonclassical MHC Class I Lineages in Divergent Amphibian Species

Ana Goyos,* Jessica Sowa,* Yuko Ohta, † and Jacques Robert*

Nonclassical MHC class Ib (class Ib) genes are heterogeneous genes encoding molecules that are structurally similar to classical MHC class Ia molecules but with limited tissue distribution and polymorphism. Mammalian class Ib genes have diverse and often uncharacterized functions, and because of their rapid rate of evolution, class Ib phylogeny is difficult to establish. We have conducted an extensive genomic, molecular, and phylogenetic characterization of class Ib genes in two Xenopodinae amphibian species of different genera that diverged from a common ancestor as long ago as primates and rodents (~65 million years). In contrast with the unsteadiness of mammalian class Ib genes, our results reveal an unusual degree of conservation of most Xenopodinae class Ib gene lineages, including a novel monogenic lineage represented by the divergent Xenopus laevis XNC10 gene and its unequivocal Silurana (Xenopus) tropicalis orthologue, SNC10. The preferential expression of this gene line by thymocytes themselves from the onset of thymic organogenesis is consistent with a specialized role of class Ib in early T cell development and suggests such a function is conserved in all tetrapods.

The online version of this article contains supplemental material.

M ajor histocompatibility complex I and class II genes encode proteins that are involved in immune surveillance by presenting antigenic peptides to CD8 and CD4 T lymphocytes, respectively. MHC class I genes are further subdivided into classical MHC class Ia (class Ia) and nonclassical MHC class Ib (class Ib) based on structural and functional differences. Class Ia genes are central to the function of the immune system because they are involved in classical peptide Ag presentation. They are minimally polygenic, highly polymorphic, and codominantly expressed on cells in almost all tissues. Class Ib genes, in contrast, are structurally similar to class Ia genes and often function in immune responses by acting as indicators of intracellular stress and malignancy (1). This is partly due to the ability of different class Ib molecules to bind a diverse array of ligands including stimulatory or inhibitory receptors expressed on T, NK, and/or NKT lymphocytes (2). Although class Ib genes are defined as heterogeneous genes encoding molecules with a limited tissue distribution and low polymorphism, there are many exceptions to this definition. For example, some class Ib proteins, such as MIC-A and MIC-B in humans, are polymorphic (3), and yet others, such as HLA-E, are ubiquitously expressed (4).

The distinction between class Ia and class Ib genes becomes even more blurred when their evolutionary relationships are analyzed. Genomic and phylogenetic analysis of class Ia and class Ib genes reveals that in general, different orders or families of mammals have different numbers of genes or genetic loci. For example, the human class Ia loci A, B, and C are shared only by hominoid species (e.g., human, gorilla, and orangutan), but the New World monkeys (e.g., tamarin) and nonprimate mammals do not have these human orthologues (5). In other words, different families or orders of mammals usually do not have orthologous genes. This pattern of evolution of MHC class I genes has been explained by the birth and death model of evolution (6) in which new genes are created by gene duplication, and some duplicated genes are maintained in the genome for a long time, whereas others are deleted or become nonfunctional through harmful mutations. Because of this rapid rate of evolution, it becomes difficult to translate directly to humans what is learned from studies of class Ib structure and function in nonhuman animal model systems. MHC class I genes (both class Ia and class Ib) have been isolated from all studied representatives of jawed vertebrates (7). Comparative analyses of MHC class I genes across distantly related nonmammalian species thus provides the opportunity to relate the evolution of structure with function and promotes understanding of the impact of convergent evolution on MHC class I molecules.

Species from the Anuran amphibian subfamily Xenopodinae exemplify one of the few documented cases in vertebrates of speciation by genome duplication via allopolyploidization (8); this taxon includes various species ranging in ploidy from diploid to dodecaploid. Although all of these species are grouped into one taxon, multiple lines of evidence indicate that they belong to two divergent genera named Xenopus and Silurana (9, 10). Xenopus laevis is a representative species of the Xenopus genus, whereas Xenopus tropicalis is a representative species of the Silurana genus and as such is now named Silurana tropicalis. These two species, X. laevis and S. tropicalis, whose divergence from a common ancestor is estimated to have occurred more than 65 million years ago (MYA), are serving as valuable nonmammalian animal models in the study of early vertebrate development and comparative immunology (9). To date, the tetraploid frog X. laevis has served as the most widely used nonmammalian comparative

*Department of Microbiology and Immunology, University of Rochester Medical Center, Rochester, NY 14642; and Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, MD 21201

Received for publication May 5, 2010. Accepted for publication October 29, 2010.

This work was supported by National Institutes of Health Grants F31-AI068610 (to A.G.), AI27877 (to Y.O.), R03-HD-06167-01, and R24-AI-059830.

Address correspondence and reprint requests to Dr. Jacques Robert, Department of Microbiology and Immunology, University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, NY 14642. E-mail address: Jacques_Robert@urmc.rochester.edu

The online version of this article contains supplemental material.

Abbreviations used in this paper: Ψ, pseudogene; AN, accession number; B, body; class Ia, classical MHC class Ia; class Ib, nonclassical MHC class Ib; EST, expressed sequence tag; H, head; B2m, β2-microglobulin; MYA, million years ago; pPBR, predicted peptide-binding residue; RACE-PCR, rapid amplification of cDNA ends; RACE-PCR, rapid amplification of cDNA ends-PCR; RT, reverse transcriptase; SNC, S. tropicalis nonclassical MHC class Ib; SNC, X. laevis nonclassical MHC class Ib.

Copyright © 2010 by The American Association of Immunologists, Inc. 0022-1767/10/$16.00

www.jimmunol.org/cgi/doi/10.4049/jimmunol.1001467
animal model for the study of immunity in general, including class Ib genes (11). The MHC of this organism has been extensively studied at the functional, biochemical, and molecular levels. And although it is not known precisely how many \textit{X. laevis} nonclassical MHC class Ib (XNC) genes there are, XNC genes are expressed and can be categorized into subfamilies based on sequence similarity; sequences within a subfamily are greater than 90% identical in amino acid sequence in their \textalpha1 domains (12). Eleven XNC subfamilies have been identified (12, 13), although the precise number of genes in each subfamily is not known because a genome sequence is not currently available for \textit{X. laevis}.

The amphibian \textit{S. tropicalis}, in contrast, has been selected as a model organism for a whole genome sequencing project because of its important phylogenetic position and because it is the only diploid species of the \textit{Xenopodinae} subfamily (http://genome.jgi-psf.org/Xentr4/Xentr4.home.html). Analysis of the \textit{S. tropicalis} genome reveals an extensive degree of conserved gene synteny with human and chicken genomes and also a high degree of conservation of genes associated with human disease (14). Scaffolds containing the MHC locus have been annotated and characterized in detail (15), and these studies indicate that compared with the mammalian MHCs, the amphibian MHC is evolutionarily highly conserved. These studies also demonstrate that, compared with \textit{S. tropicalis} MHC, the vertebrate MHC experienced a vigorous rearrangement in the bony fish and with some modification in bird lineages and a translocation and expansion of the MHC class I genes in the mammalian lineage. For this reason, the amphibian is considered to be a valuable model to study the evolution of the MHC. We have therefore identified and characterized a large family of \textit{S. tropicalis} nonclassical MHC class Ib (SNC) genes and compared them with XNC subfamilies to evaluate the degree of evolutionary conservation of class Ib genes within the \textit{Xenopodinae} subfamily.

\section*{Materials and Methods}

\subsection*{Animals}

Outbred \textit{X. laevis} and \textit{S. tropicalis} adults and larvae were obtained from our breeding colony (http://www.urmc.rochester.edu/smd/nhb/xenopus/index.html). Animals were sacrificed by immersion in tricaine methane sulfonate (5 mg/ml for adults and 1 mg/ml for tadpoles). Sublethal gamma-irradiation (10 Gy) was delivered to metamorphic larvae stage 56–58 with a cobalt source. All animals were handled under strict laboratory and university committee on animal resources regulations (approval no. 100577/2003-151 and 2004-199), minimizing discomfort at all times.

\subsection*{RNA extraction, rapid amplification of cDNA ends-PCR, and RT-PCR}

Total RNA was isolated from tissues of pooled adults (10 to 20 individuals) and larvae (30 individuals). All RNA extractions were carried out using 1 ml TRIzol reagent. For rapid amplification of cDNA ends (RACE), 5' and 3' rapid amplification of cDNA ends-PCR (RACE-PCR)-ready cDNAs were prepared using the SuperSMART PCR cDNA synthesis kit from Clontech (Mountain View, CA). RACE-PCR was carried out using the Advantage 2 PCR Enzyme System from Clontech. RACE-PCR primers used were \textbf{XNCT}rap-5'RACE-1 (5'–CCCT TCC TCT GGT GTT ACC TCC AC-3') and \textbf{XNCT}rap-3'RACE-2 (5’–GCC ACT CTC TGA CTC TGA GCT GG-3'). cDNA synthesis was performed using the iScript cDNA synthesis kit from Bio-Rad (Hercules, CA) and diluted 2:1. RT-PCR primers specific for the SNC genes 2.2, 4, 6.1, 6.2, 7.1, 12, and 13.3 as well as for EF-1\textalpha were designed, and the annealing temperatures were determined using gradient PCR. All RT-PCRs included water and reverse transcriptase (RT)-minus controls (omission of RT during cDNA synthesis).

\subsection*{Southern blotting}

Genomic DNA from \textit{Xenopus} erythrocytes was isolated as described (16) and digested to completion with restriction endonuclease. The digested DNA was separated on a 1% agarose gel and transferred onto nylon membranes by the capillary blotting technique in 10X SSC. Increasing amounts of DNA were loaded in higher-ploidy animals according to the ploidy level.

\section*{Bioinformatics tools}

Nucleotide and amino acid sequences were analyzed using utilities at the National Center for Biotechnology Information Web site (http://www.ncbi.nlm.nih.gov). Nucleotide and amino acid sequences were aligned using ClustalX (http://www.clustal.org/), and alignments were edited and shaded with the GeneDoc program (http://www.nrblrc.bscs.qub.ac.uk/genedoc/). The nucleotide and amino acid sequences of known genes were retrieved from GenBank using ENTREZ at http://www.ncbi.nlm.nih.gov. Genomic sequences were retrieved from the \textit{S. tropicalis} JGI Web site (http://genome.jgi-psf.org/Xentr4/Xentr4.home.html). Homology searches were performed using TBLASTN and BLASTN programs. Phylogenetic analysis was performed using Molecular Evolutionary Genetics Analysis (MEGA, version 4.0; http://www.megasoftware.net/). Phylogenetic trees were generated by the neighbor-joining method of Saitou and Nei (17). Genetic distances were calculated by estimating the number of amino acid substitutions using the p-distance method with pairwise deletion of gaps. Numbers on nodes represent percentages of 1000 bootstrap replicates supporting each partition.

\section*{Results}

\subsection*{Genomic characterization of SNC genes}

XNC gene transcripts can be grouped into at least 11 subfamilies, numbered XNC1 through XNC11, based on amino acid sequence similarity of their \textalpha1 domains (12, 13). Although \textit{X. laevis} is estimated to have roughly 20 XNC genes per haplotype by Southern blot analysis (12), in absence of genomic sequences their exact total number and the respective number of XNC genes within each subfamily is unknown. In contrast, the genome of the only diploid frog of the \textit{Xenopodinae} subfamily, \textit{S. tropicalis}, has recently been fully sequenced and annotated. Because class Ia and class Ib genes of mammalian species studied to date are generally not conserved between animals as divergent as \textit{X. laevis} and \textit{S. tropicalis}, we sought to identify and characterize SNC genes in \textit{S. tropicalis} and compare them with XNC gene subfamilies of \textit{X. laevis}. We used in silico analysis and searched the \textit{S. tropicalis} genomic database version 4.1 (http://genome.jgi-psf.org/Xentr4/Xentr4.home.html) using the TBLASTN algorithm with deduced amino acid sequences of XNC transcripts. The BLASTN algorithm was also used to search the genome with nucleotide sequences of XNC \textalpha1, \textalpha2, and \textalpha3 domain exons.

Two scaffolds, 557 and 427, encompassing 1.76 Mbp, contain a total of 29 automatically predicted SNC genes, including pseudogenes and gene fragments (Fig. 1, Table I). Some of these genes were missing exons containing either Leader or Transmembrane and Cytoplasmic domains. Manual searches were performed around these SNC genes, and if the exons were present, they were manually annotated. Several class I-like sequences were also found in scaffolds 2166, 145, and 3581. However, their sequences were mainly incomplete and not supported by expressed sequence tag (EST) databases. The characterization of these class I-like sequences was therefore not considered further in this study. In summary, a total of 29 SNC genes, including some pseudogenes and gene fragments, was retained for further analysis.

Because our objective was to assess orthologous relationships between XNC gene subfamilies in \textit{X. laevis} and SNC genes in \textit{S. tropicalis}, we chose a nomenclature and the methodology used in \textit{X. laevis} that is based on the most divergent \textalpha1 domain to discriminate subfamilies (12). To assign SNC genes to particular subfamilies that correspond with XNC \textalpha1 domains among SNC genes. This resulted in 14 separate subfamilies defined by separate clades of SNC genes with bootstrap support <70. To determine orthologies to XNC subfamilies,
an amino acid alignment of SNC and XNC α1 domains was made. This alignment was used to generate a neighbor-joining phylogenetic tree. Orthologous XNC and SNC gene subfamilies were defined as any clade having a bootstrap value of 75 or higher. As such, the amino acid sequence similarity of individual α1 domain sequences within each clade is greater than 60%. Each S. tropicalis gene model was named with a three-letter code, SNC, followed by a distinctive number. We matched the numbers between XNC subfamilies and SNC orthologues [or co-orthologues, defined as two or more S. tropicalis SNC genes that are collectively orthologous to one X. laevis XNC subfamily, likely due to recent gene duplication (18)] based on sequence similarity determined by nucleotide (data not shown) and amino acid alignments of the α1 (Supplemental Fig. 1) and α2 (Supplemental Fig. 2) domains, as well as by phylogenetic analysis of the α1 domains (Fig. 2A). SNC genes without an orthologous XNC subfamily were given a number higher than 11 (i.e., because of the 11 XNC subfamilies in X. laevis characterized to date). Pseudogenes were labeled with the character “C.” In a few cases, particular genes could have been grouped into one of two subfamilies. For example, SNC genes SNC2/3.3, SNC2/3.4, and SNC2/3.5 were similar to both XNC2 and XNC3. We decided to name the genes with an “SNC2/3” prefix simply because this group of genes was not different enough from either XNC2 (α1 domain) or XNC3 (α3 domain) to be grouped into its own unique subfamily and because regardless of whether these particular SNC genes would have been grouped into the XNC2 or XNC3 subfamily, SNC lineages 1, 2, 3, 4, and 16 all evolved from a recent common ancestor. Our analysis indicates that S. tropicalis SNC genes group into 14 subfamilies based on sequence similarity in their α1 domains (Fig. 2A). Most SNC subfamilies are monogenic, and a few others have experienced lineage-specific expansions. Surprisingly, most SNC genes are orthologous (or co-orthologous) to XNC subfamilies. Of the 14 SNC subfamilies, 9 are conserved with X. laevis. Although most SNC genes are conserved with XNC subfamilies, there are some SNC genes that do not have a corresponding orthologue in X. laevis, and vice versa, and therefore represent species-specific class Ib genes.

**Conservation of Xenopodinae class Ib gene lineages**

Analysis of class Ib genes of S. tropicalis and X. laevis reveal an unprecedented degree of conservation of multiple class Ib gene lineages between two divergent vertebrate species. These results suggest that compared with mammalian taxon, class Ib genes in Xenopodinae underwent different selective pressures, favoring maintenance of individual class Ib lineages. To address this hypothesis further, we evaluated the degree of similarity of the different S. tropicalis SNC genes with X. laevis XNC subfamilies and class Ia through nucleotide and amino acid alignments and phylogenetic analyses. To strengthen this analysis, we included MHC class I sequences of two other amphibian species available in GenBank: the northern leopard frog, *Rana pipiens*, and the axolotl, *Ambystoma mexicanum*, organisms whose most recent common ancestors with species from the Xenopodinae subfamily date back 220 and 270 MYA, respectively, as determined by both molecular and paleontological evidence (19, 20). Neighbor-joining consensus trees of amino acid alignments (i.e., nucleotide data gave similar results; data not shown) of individual α1, α2, and α3 extracellular domains were constructed with 1000 bootstrap replicates (Fig. 2A–C). Rather than forming species-specific clusters, the majority of class Ib loci of *S. tropicalis* and *X. laevis* cluster in a locus-specific manner (Fig. 2A, 2B). These gene lineages have
thus maintained their genetic identities at least since the S. tropicalis and X. laevis speciation event, more than 65 MYA (9, 10). Notably, both the α1 and α2 domain trees (Fig. 2A, 2B) demonstrate that Xenopodinae class Ia sequences are more similar to class Ia sequences from the Anuran amphibian R. pipiens than they are to class Ib gene sequences from their own species. Assuming similar rates of evolution, this result suggests that Anuran class Ia and class Ib genes diverged more than 200 MYA. This conclusion is, however, subjected to the limitation that the rates of evolution of these two gene families have probably not been constant and have varied independently. Almost certainly, particular regions in the α1 and α2 domains have been selected for divergence (e.g., interaction with peptides), whereas other regions have been selected to stay the same (e.g., interactions with β2m, TCRs, NK cell receptors, etc.). Moreover, the class Ib lineage should have been subjected to fast evolutionary changes when it specialized into new functions, but afterward changes may have been selected against. Importantly, and irrespective of the variation in the rate of evolution, this phylogeny also implies that the SNC10 gene along with its X. laevis XNC10 orthologue were the first class Ib genes to have diverged from the XNC/SNC common ancestor.

In contrast, a different evolutionary scenario emerges from the analysis of α3 domain sequences. Fig. 2C reveals an overall sequence similarity shared by all XNC and SNC α3 domains suggesting intraspecies homogenization and presumably subsequent selection to maintain important functions of α3 domain sequences of SNC and XNC subfamilies. This is similar to what is seen for class Ia α3 domains (21). Homogenization of α3 domains of MHC class I molecules might occur to maintain important functions, such as association with β2m. It is possible that the difference between α3 domains of class Ia and class Ib may come from their association with different accessory proteins (β2m, CD8). For example, the putative Xenopodinae β2m- and CD8-binding residues of the α3 domain show some distinctive features between class Ia and class Ib. For both β2m- and CD8-binding residues, there are some positions that are conserved between class Ia and class Ib and other positions that have a clear dichotomy (Supplemental Fig. 3). Therefore, for both SNC and XNC genes, α1 and α2 domain sequences, collectively forming the peptide-binding domain, cluster according to loci, whereas α3 Ig domain sequences cluster according to species.

It has been proposed that a few class Ib loci diverged from class Ia loci a long time ago and now have different functions (22). Because our phylogenetic analysis indicates that Xenopodinae class Ia and class Ib genes likely diverged more than 200 MYA, we wanted to compare SNC gene sequences with other jawed vertebrate class Ia and class Ib sequences. A phylogenetic analysis of expressed SNC gene sequences (see the section SNC genes are expressed for more details) and representative jawed vertebrate class Ia and class Ib sequences (Fig. 3) was conducted. We postulated that selective pressures should primarily affect Ag-binding functions of class Ib genes and therefore focused our analysis on α1 and α2 domain sequences. The tree presented in Fig. 3 shows that SNC and XNC genes form a monophyletic clade with respect to other vertebrate class Ia and class Ib genes, including amphibian class Ia. This provides support for a common evolutionary origin of class Ib genes from the Xenopodinae subfamily.

In conclusion, phylogenetic analysis indicates that SNC and XNC genes from S. tropicalis and X. laevis evolved from a common ancestor, with the XNC/SNC10 gene lineage being the first to diverge. In addition, although there has been expansion and contraction of individual class Ib genes, there is a remarkable preservation of several class Ib gene lineages over a relatively long period of time in two divergent amphibian species.

SNC genes are expressed

Class Ib molecules identified to date tend to have a more restricted tissue expression pattern compared with that of class Ia (4). To determine if the different SNC genes also have a limited tissue-specific distribution, we conducted EST database searches as well as RACE-PCR and RT-PCR analyses. We first searched the National Center for Biotechnology Information GeneBank S. tropicalis EST database to determine which of the 29 SNC genes have EST support. BLAST analyses were performed with the TBLASTN algorithm using SNC consensus α1 and α2 domain amino acid sequences from each of the 14 SNC subfamilies (see alignment in Supplemental Figs. 1 and 2). A total of 47 SNC transcripts were obtained, but only 35 contained sufficient sequence in the α1, α2, or Cytoplasmic domains to assign them to a particular gene. In total, nine genes were represented by ESTs (Table II). In addition, most EST transcripts were from adult thymus and spleen, whereas a few were from adult brain, fat bodies, intestine, liver, lung, skin, skeletal muscle, oviduct, and testes.

To substantiate our EST data, we performed two distinct 5′ RACE-PCR reactions: one with cDNA synthesized from a pool of 20 adult outbred thymuses, and another one with cDNA from a pool of 20 adult spleens. RACE-PCR products were separated by gel electrophoresis, and bands were cloned and sequenced. A total of 43 SNC sequences analyzed from adult thymus and spleen supported 7 SNC genes (Table II), 5 of which were common to those identified by ESTs. Finally, RT-PCR of cDNA from different organs from a pool of 20 adult tissues was performed using SNC

<table>
<thead>
<tr>
<th>Scaffold No.</th>
<th>Protein ID</th>
<th>Position</th>
<th>Gene Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>427</td>
<td>178628</td>
<td>963091–973347</td>
<td>SNC6.1</td>
</tr>
<tr>
<td></td>
<td>178629</td>
<td>977506–984314</td>
<td>SNC5.2</td>
</tr>
<tr>
<td></td>
<td>374185</td>
<td>990632–998282</td>
<td>SNC10</td>
</tr>
<tr>
<td></td>
<td>992046–1004069</td>
<td>SNC9P</td>
<td></td>
</tr>
<tr>
<td>557</td>
<td>386267</td>
<td>253–6190</td>
<td>SNC12</td>
</tr>
<tr>
<td></td>
<td>386272</td>
<td>16483–37957</td>
<td>SNC2.1P</td>
</tr>
<tr>
<td></td>
<td>386185</td>
<td>57583–76814</td>
<td>SNC2.2</td>
</tr>
<tr>
<td></td>
<td>386177</td>
<td>85023–95237</td>
<td>SNC1.1P</td>
</tr>
<tr>
<td></td>
<td>386268</td>
<td>110253–116573</td>
<td>SNC4</td>
</tr>
<tr>
<td></td>
<td>458818</td>
<td>126280–134843</td>
<td>SNC13.1</td>
</tr>
<tr>
<td></td>
<td>182055</td>
<td>153164–176381</td>
<td>SNC13.2</td>
</tr>
<tr>
<td></td>
<td>386203</td>
<td>180370–187959</td>
<td>SNC13.3</td>
</tr>
<tr>
<td></td>
<td>386310</td>
<td>196609–203924</td>
<td>SNC13.4</td>
</tr>
<tr>
<td></td>
<td>182058</td>
<td>232324–249472</td>
<td>SNC13.5</td>
</tr>
<tr>
<td></td>
<td>182059</td>
<td>256738–273249</td>
<td>SNC16.1</td>
</tr>
<tr>
<td></td>
<td>386304</td>
<td>279463–291543</td>
<td>SNC16.2</td>
</tr>
<tr>
<td></td>
<td>386297</td>
<td>297124–306904</td>
<td>SNC16.3</td>
</tr>
<tr>
<td></td>
<td>386192</td>
<td>331557–343917</td>
<td>SNC17</td>
</tr>
<tr>
<td></td>
<td>386184</td>
<td>354429–364805</td>
<td>SNC12.1</td>
</tr>
<tr>
<td></td>
<td>386302</td>
<td>377723–386183</td>
<td>SNC16.4</td>
</tr>
<tr>
<td></td>
<td>386241</td>
<td>410224–415542</td>
<td>SNC14</td>
</tr>
<tr>
<td></td>
<td>386246</td>
<td>425413–427874</td>
<td>SNC5</td>
</tr>
<tr>
<td></td>
<td>386189</td>
<td>433532–441878</td>
<td>SNC7.1</td>
</tr>
<tr>
<td></td>
<td>386240</td>
<td>487368–502788</td>
<td>SNC7.2</td>
</tr>
<tr>
<td></td>
<td>386261</td>
<td>511135–521252</td>
<td>SNC7.3</td>
</tr>
<tr>
<td></td>
<td>386244</td>
<td>529093–535588</td>
<td>SNC7.4</td>
</tr>
<tr>
<td></td>
<td>386249</td>
<td>548537–557354</td>
<td>SNC7.5</td>
</tr>
<tr>
<td></td>
<td>386214</td>
<td>583078–595673</td>
<td>SNC15</td>
</tr>
<tr>
<td></td>
<td>386300</td>
<td>610756–614712</td>
<td>SNC7.6</td>
</tr>
</tbody>
</table>

*Transmembrane/cyttoplasmic domain manually annotated for SNCs 4.5, 1.2, 2.1.

*SNC10 manually annotated and corrected.

Ψ, pseudogene.
gene-specific primers. In addition to the 11 SNC genes already found to be expressed, RT-PCR resulted in detection of a 12th expressed SNC gene; SNC10 transcripts were amplified from adult thymus (Table II).

In summary, evidence for expression for 12, of 29 total, SNC genes was obtained in this study (Fig. 1A, asterisks). As expected, many of these genes, including SNC10, displayed a restricted tissue expression pattern that mostly involved lymphoid organs. In addition, the levels of expression of several SNC genes were higher than those of others, as is the case for X. laevis as well as mammalian class Ib genes (23, 24).

Genes from the divergent Xenopodinae class Ib lineage, SNC10 and XNC10, are orthologous

It has been proposed that class Ib genes that are evolutionarily old might have evolved unique functions that became fixed early during vertebrate evolution (25). The orthologous relationship of the oldest Xenopodinae class Ib genes, SNC10 and XNC10, was investigated with the purpose of gaining functional insight. It is important to note that during our initial inspection of the 29 automatically predicted gene models, we did not identify an obvious XNC10 orthologue. However, it was observed that JGI gene model 374185 had an α1 domain exon similar to XNC9 and an α2 domain exon similar to XNC10. Therefore, this open reading frame was either annotated incorrectly by the gene prediction models or it was a hybrid SNC gene. To address these possibilities, the genomic region surrounding and including this predicted open reading frame was further inspected and manually annotated. Two unannotated exons coding for an XNC10-like Leader and α1 domain were identified in between the automatically predicted XNC9-like Leader-α1 domain-containing exon and the XNC10-like α2 domain exon for gene model 374185 (Fig. 1B). To know whether the S. tropicalis genome contains an XNC10 orthologue, we verified the gene architecture by PCR amplification of the genomic region containing the putative SNC10 gene followed by analytical endonuclease digestion (data not shown). These results confirmed that S. tropicalis contains an SNC gene that is orthologous to X. laevis XNC10.

One notable feature of the XNC10 protein was the uniqueness of its predicted peptide-binding residues (pPBRs) (13). To explore further the orthologous relationship between X. laevis XNC10 and S. tropicalis SNC10 and obtain indirect evidence of a critical, conserved immune function, we analyzed molecular features of the pPBRs of SNC10. As previously done for X. laevis class Ia and XNC genes (11, 12), we based our analysis on the amino acid alignment of these genes with HLA-A (26). Specifically, we com-
pared SNC10 amino acid residues that align with invariant amino acids of the peptide binding domain of human HLA-A (27) to analyze any conservation of residues potentially involved in peptide binding. In mammalian class Ia, side chains of these invariant residues point into the MHC class I cleft and are essential for peptide anchoring through hydrogen or ionic bonding to backbone atoms of peptide N and C termini (Table III). The putative peptide N terminus of XNC10 contains two XNC10-specific residues, Asp/Asn59 and His171 (13), which are not conserved with SNC10. The putative peptide C terminus docking residues of SNC10, in contrast, are conserved with those of XNC10 (Table III). For example, the peptide C terminus docking residues of XNC10 are very different compared with residues in the same position of other XNC molecules. In XNC10, Lys146 is conserved with class Ia and Tyr/Arg84 is conservatively substituted for Ser, a residue that is still capable of forming hydrogen bonds. Both of these C terminus docking residues are conserved with SNC10. The XNC10-specific residue Ala143 is also conserved in SNC10. Of interest, the only peptide C terminus residue that is not conserved between XNC10 and SNC10 is Trp147. Notably, this residue is identical to that found in the same position in class Ia. Therefore, although N terminus residues of SNC10 pPBRs are not very similar to those of XNC10 (they are actually more conserved with XNC1–9 pPBRs), putative peptide C terminus docking residues of SNC10 are remarkably similar to those of XNC10 (Table III). These results are consistent with the interesting possibility that the XNC/SNC10 amphibian class Ib subfamily binds a common or conserved antigenic motif with its putative peptide C terminus docking residues while simultaneously building flexibility into the system at the peptide N terminus binding residues. Whether XNC/SNC10 does indeed bind an Ag and the nature of this Ag (e.g., peptide, glycolipid, or other), however, remains to be determined.

One hallmark of class Ib genes, which contrasts with class Ia, is their limited polymorphism (29). This is thought to reflect dif-

Table II. Summary of adult expression pattern for SNC genes

<table>
<thead>
<tr>
<th>SNC Gene</th>
<th>No. EST</th>
<th>No. of Adult RACE-PCR</th>
<th>RT-PCR</th>
<th>SNC Tissue Expression Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNC2.2</td>
<td>1</td>
<td>8</td>
<td>Yes</td>
<td>Thymus, spleen, colon, intestine, liver, stomach</td>
</tr>
<tr>
<td>SNC4</td>
<td>0</td>
<td>10</td>
<td>Yes</td>
<td>Thymus, spleen, intestine, skin</td>
</tr>
<tr>
<td>SNC6.1</td>
<td>1</td>
<td>7</td>
<td>Yes</td>
<td>Thymus, spleen</td>
</tr>
<tr>
<td>SNC6.2</td>
<td>0</td>
<td>1</td>
<td>Spleen</td>
<td></td>
</tr>
<tr>
<td>SNC7.1</td>
<td>12</td>
<td>3</td>
<td>Yes</td>
<td>Ubiquitous</td>
</tr>
<tr>
<td>SNC7.2</td>
<td>1</td>
<td>0</td>
<td>ND</td>
<td>Thymus, spleen</td>
</tr>
<tr>
<td>SNC10</td>
<td>0</td>
<td>0</td>
<td>Yes</td>
<td>Thymus</td>
</tr>
<tr>
<td>SNC12</td>
<td>2</td>
<td>5</td>
<td>Yes</td>
<td>Ubiquitous</td>
</tr>
<tr>
<td>SNC13.1</td>
<td>2</td>
<td>0</td>
<td>ND</td>
<td>Skeletal muscle, heart</td>
</tr>
<tr>
<td>SNC13.2</td>
<td>10</td>
<td>0</td>
<td>ND</td>
<td>Thymus, lung</td>
</tr>
<tr>
<td>SNC13.3</td>
<td>5</td>
<td>9</td>
<td>Yes</td>
<td>Thymus, spleen, heart, liver, skin</td>
</tr>
<tr>
<td>SNC13.5</td>
<td>1</td>
<td>0</td>
<td>ND</td>
<td>Brain</td>
</tr>
</tbody>
</table>

ND, no data.
in selectives pressures that drive distinct functions of class Ib compared with class Ia (30). To address the degree of SNC10 gene polymorphism, we amplified, analyzed, and aligned 13 α1 domain and 10 α2 domain nucleotide sequences for SNC10 that were obtained by 5’ RACE-PCR of a pool of 20 outbred S. tropicalis larvae. This analysis reveals that, like XNC10 (13), the α1 and α2 domains of SNC10 are not very polymorphic. Only six sites in the α1 domain, of 258 total sites, have a single-nucleotide polymorphism in 1 of the 14 sequences analyzed (Supplemental Fig. 4). Similarly, only two sites of 170 analyzed are mutated in the α2 domain. A larger sample size would be necessary to infer any more from this analysis, therefore we will simply conclude that as is expected for a typical class Ib gene, the coding region of SNC10 is not very polymorphic.

Southern blot hybridization with an XNC consensus α3 domain probe has shown that all Xenopus and Silurana species maintain a large number of class Ib genes (12). Furthermore, we have established that the XNC10 gene is conserved in all species tested of the genus Xenopus (13). To address whether other species of the genera Silurana and/or Xenopus maintain the XNC/SNC10 gene lineage, hybridization of a genomic Southern blot of representative lineages of class Ib genes from the XNC/SNC10 subfamily that are selectively maintained in individual species of the amphibian subfamily Xenopodinae. This is perhaps not unexpected because the different species are thought to have arisen by allopolyploidization.

**Early expression of SNC10 in the radiosensitive thymocytes of S. tropicalis larval thymus**

In X. laevis larvae, the XNC10 gene is primarily expressed by thymocytes (13). The striking molecular features and expression profile of XNC10 are suggestive of a function in T cell development in X. laevis larvae. If the orthology of SNC10 and XNC10 is indicative of a similar function that has been conserved over evolutionary time, then SNC10 expression should also preferentially occur in S. tropicalis larval thymus. To explore this possibility, we determined the expression pattern of SNC10 at adult (Fig. 5A) and early developmental stages (Fig. 5B, 5C) of S. tropicalis, focusing especially on the thymus. RT-PCR was performed on various organs pooled from 20 individual tadpoles (stage 56) using SNC10-specific primers. Like XNC10, SNC10 expression was found primarily in adult and larval thymus and to a lesser extent in the spleen (Fig. 5A, 5B). SNC10 was also markedly expressed in the gills and at low levels in the pancreas. All RT-minus controls were negative, and each amplified band was cloned and sequenced to verify its authenticity.

Another key feature of XNC10 expression in X. laevis tadpoles is its early onset of expression, at developmental stage 39 (13), which corresponds with the onset of thymic organogenesis (31). To address whether SNC10 is also expressed early during ontogeny in S. tropicalis thymus, its expression was assayed by RT-
In primates, for example, MHC class I genes diverged so recently—approximately 33–35 MYA—that even humans and New World monkeys, which diverged only 65 MYA, do not share functional genes. Similarly, in frogs and crystals, class Ib genes are restricted to the thymus. Thymus, therefore, is composed of transposable elements (14), and this could predispose certain genomic regions to reshuffling. The preserved pattern of multiple distinct class Ib gene lineages over a relatively long evolutionary time is suggestive of the selection of these molecules for performing important functions.

This pattern of class Ib evolution within the Xenopodinae subfamily is in stark contrast to what is seen in mammals where different families or orders usually do not have orthologous genes. In primates, for example, MHC class I genes diverged so recently that even humans and New World monkeys, which diverged only ~33–35 MYA, do not share functional genes (35, 36). Similarly,
MHC class I genes from two marsupial species, which separated ~48 MYA from a common ancestor, show no orthologous relationships (37). The turnover rate of MHC class I genes in mammals is therefore very high. Class Ib gene evolution in the Xenopodinae subfamily of amphibians appears to have taken a different route compared with that of mammals. Various attempts to date the origin of Xenopodinae using extant species have provided an estimate that ranges from 50.4 to 81.3 MYA (9, 38), which is why we decided to use the average time of 65 MYA. Although the evolutionary relationships of some individual class Ib genes are established within mammals (e.g., CD1, HLA-E/Qa1), the conservation of 9 of 14 S. tropicalis and X. laevis class Ib lineages over this relatively long period of evolution is to our knowledge the first case in vertebrates of such preservation of a group of class Ib genes.

The morphological resemblance of S. tropicalis and X. laevis may suggest a corresponding genetic stability that would explain a higher overall conservation of class Ib locus compared with that of mammals. However, we think this is unlikely. Even though there is a small degree of morphological difference between these two amphibian species, they differ far more from each other in their protein sequences than do organisms with large morphological differences, such as humans and chimpanzees (39, 40). Although S. tropicalis and X. laevis have many morphological similarities, numerous studies of their mitochondrial DNA have demonstrated the large degree of divergence between species of the Silurana and Xenopus genera (9, 38, 41). This divergence between the Xenopodinae subfamily does not appear to be due to drastic differences in rates of nucleotide substitution between amphibians and mammals because when rates of nucleotide substitution have been compared between pairs of X. laevis duplicated genes and the orthologous loci of humans and rodents, it has been determined that, at the amino acid level, the X. laevis genes and the mammalian genes are under similar constraints (41). Collectively, these results suggest that morphological evolution and protein sequence evolution can proceed at contrasting rates.

All extant species of frogs from the Xenopodinae subfamily are confined to the continent of Africa. The antiquity of the Xenopodinae subfamily of amphibians is highlighted by the discovery of a fossil identified as a species from the Xenopus genus in South America (42). This indicates that evolutionary radiation with the Xenopodinae subfamily of amphibians must have preceded the separation of the African and South American continents more than 100 MYA (10). The substantial degree of divergence between X. laevis and S. tropicalis (43) is therefore one of the reasons why the maintenance of multiple and distinct class Ib lineages in these two species of frogs is suggestive of important conserved functions. One possible driving force behind the conservation of multiple class Ib lineages in two species of the Xenopodinae subfamily of amphibians might be their exposure to relatively similar pathogens. There is some overlap of the natural geographical niches that S. tropicalis and X. laevis occupy in the continent of Africa (9). These organisms may have therefore needed to develop a way to defend themselves against a common set of pathogens for which certain class Ib lineages may be specialized to recognize. Furthermore, in contrast to most mammals where there are three class Ia genes per haplotype, a large number of class Ib gene lineages in species of the amphibian subfamily Xenopodinae might serve to compensate for the presence of only one class Ia gene.

Nonfunctionalization of S. tropicalis class Ib gene lineage

The particular expression profile of the XNC11 gene (e.g., high expression by three different lymphoid tumors derived from independent spontaneous thymic tumors of genetically unrelated animals, very weak expression by thymocytes, and no expression elsewhere) suggests an association with tumorigenesis in X. laevis (13). In case XNC11 would, for example, provide some survival signal during thymocyte differentiation as some class Ib molecules do in mammals (44), one can conceive that a deregulated overexpression of the gene would be an advantageous step in neoplastic transformation. Alternatively, the aberrant expression of XNC11 in thymic tumors may just be related to aberrant regulation or the stage of thymic differentiation from which the tumor developed. Notably, no XNC11 homologue can be found in S. tropicalis, which makes us speculate that genes of the X. laevis XNC11 subfamily may not have been conserved in the Silurana lineage because of their potential association with tumors.

In X. laevis, transcripts of the XNC11 subfamily are most similar to the XNC5 subfamily (13). The S. tropicalis orthologue of the X. laevis XNC5 subfamily is the SNC5 gene. It is interesting to note that the SNC5 subfamily is monogenic. In fact, most class Ib gene lineages in S. tropicalis are monogenic (Fig. 2). Although the loss of duplicated genes over time does not need to be the result of a selection, the maintenance of many monogenic Xenopodinae class Ib gene lineages together with a few multigenic ones, all located in the same genetic region, suggests to us that some class Ib genes are under pressure to remain a single copy in the genome. In this regard, the targeted diploidization of the class Ib locus in polyplid species as a control of gene copy number is consistent with the idea that, as for class Ia, some class Ib genes cannot tolerate gene duplication because of a critical nonredundant dose-dependent function. For example, because there is only one copy of SNC5 in the S. tropicalis genome, it is possible that in X. laevis, XNC11 originally duplicated from XNC5, and the consequence of having more than one XNC5-like gene was a predisposition to tumors. This interesting possibility remains to be explored.

Stability of the divergent XNC/XNC10 Xenopodinae class Ib gene lineage

The XNC/XNC10 subfamily forms an independent phylogenetic clade that is intermediate between S. tropicalis and X. laevis class Ib and class Ia sequences. This topology suggests an independent evolutionary origin of XNC/XNC10 relative to other Xenopodinae class Ib lineages. XNC10 and XNC10 might have thus become fixed to serve a critical function early during Xenopodinae evolution. Notably, in both S. tropicalis and X. laevis there is only one gene representing this class Ib lineage, and genomic Southern analysis strongly suggests that the monogenic and diploidized state of this subfamily has also been maintained in other species, including polyploids. The fact that this subfamily is old and the gene dosage has been maintained to one is a strong argument for an important, nonredundant function.

Expression of XNC10 and XNC10 is associated with thymocyte differentiation during early larval development at the onset of thymic organogenesis. This is of particular interest given the intriguing issue, which is not yet fully elucidated, that X. laevis larvae have thymus-dependent circulating mature CD8 T cells but lack consistent class Ia surface expression, especially in the thymus, until metamorphosis (45). One can postulate that the relatively high level of SNC10 and XNC10 expressed in thymus, from early developmental stages, compensates for the lack or low level of class Ia expression. As such, it may be the case that in larvae, certain class Ib molecules, such as XNC10 and/or SNC10, are primarily and critically involved in regulating the differentiation of CD8 T cells needed for rapid protection during early development, considering that tadpoles are free swimming by 3 dpf. Larval sublethal gamma-irradiation experiments have shown that
in the thymus, XNC10 and SNC10 are predominately expressed by the radiosensitive hematopoietic population. In mice, there are several examples of thymic T cells being selected on class Ib molecules expressed by hematopoietic cells, in the absence of class Ia. For example, the class Ib molecules TL and CD1d are expressed on cortical double positive thymocytes and seem to be the key to the development of a subset of γδ as well as NKT cells, respectively (46). These findings are consistent with a convergence of function within amphibian and mammalian class Ib. It will be fascinating to explore the possibility that because space and time are limited in Xenopus and Silurana larvae, they may preferentially generate class Ib-restricted “innate” T cells, which have the intrinsic property of an activated phenotype thus requiring much less expansion and time to carry out effector functions (47). This may constitute a critical defense strategy early in life when an organism’s adaptive immune system is not yet fully competent. Having cells around with an accelerated ability to detect and control danger might serve to manage a potential infection before it is allowed to establish. Furthermore, one of the useful features of X. laevis larvae as a potential model to study class Ib function is that this model is a natural system that has minimal contribution from class Ia.

Acknowledgments
We thank Dr. Martin Flajnik for invaluable discussions. We also thank an anonymous reviewer for helpful suggestions and Hristina Nedelkovska for critical reading of the manuscript and useful suggestions. The expert animal husbandry of David Albright and Tina Martin is greatly appreciated.

Disclosures
The authors have no financial conflicts of interest.

References