Evolution by Recombination and Transspecies Polymorphism in the MHC Class I Gene of *Xenopus laevis*

David H. Bos[1] *and Bruce Waldman*

School of Biological Sciences, University of Canterbury, Christchurch, New Zealand

The patterns of major histocompatibility complex (MHC) evolution involve duplications, deletions, and independent divergence of loci during episodes punctuated by natural selection. Major differences in MHC evolution among taxa have previously been attributed to variation in linkage patterns of class I and class II MHC genes. Here we characterize patterns of evolution in the MHC class Ia gene of *Xenopus laevis* in terms of polymorphism, recombination, and extent of transspecies polymorphism. We also compare these patterns to see if a correlation exists with linkage or separation of the MHC class I and class II regions as seen in amphibians and teleost fishes. In *X. laevis*, we find high levels of polymorphism. Also, genetic exchange is relatively frequent and occurs in intron II, reshuffling allelic forms of exons 2 and 3. Evolutionary relationships among class I alleles show an intermingling of alleles from divergent *Xenopus* species rather than a species-specific clustering. Results indicate that the patterns of evolution are similar to those found in salmonid fishes and are different from the mode of evolution seen in primates. Similar patterns of class Ia evolution in salmonid fishes and *X. laevis* suggest that nonlinkage of class I and class II regions alone is insufficient to explain some patterns of MHC evolution in salmonids.

Introduction

Major histocompatibility complex (MHC) class I and class II genes are found in all gnathostomes and encode structurally similar proteins that present antigenic peptides to T lymphocytes. Class II proteins are expressed mainly on specialized antigen-presenting cells and primarily function to bind peptides derived from extracellular pathogens. Class I proteins are expressed in almost all cells and are involved in monitoring the internal environment of the cell for foreign, mutated, or misfolded proteins. Class I genes comprise classical (class Ia) and nonclassical (class Ib) loci which differ in polymorphism, structure, function, and expression pattern. Aside from their important role in the immune system, the MHC genes are of particular interest because of their patterns of genetic diversity.

Class I and class II alleles have several exceptional features. They exhibit very high levels of allelic diversity and amino acid polymorphism (Parham and Ohta 1996). Some MHC allelic lineages also exhibit unusual longevity which predates the formation of species (Figueroa, Gunther, and Klein 1988; Lawlor et al. 1988). As a result, "transspecies polymorphisms" exist, whereby some MHC alleles from separate species cluster together in phylogenetic analysis to the exclusion of alleles from within the same species. Variations in linkage patterns, order of gene loci, and the number of gene family members resulting from tandem duplications have also been observed (Kelley, Walter, and Trowsdale 2005). Many of the features of the MHC have been attributed to the forces of balancing selection acting at the molecular level (Hughes and Yeager 1998).

In the typical pattern of MHC evolution, class I and class II families of genes evolve differently from each other, each with its own rate of duplication, divergence, longevity, and pattern of recombination. Specifically, class II transspecies polymorphisms extend further back in evolutionary history than in class I lineages (Bontrop et al. 1999; Vogel et al. 1999). Class II loci also have higher levels of polymorphism than class I loci. Additionally, allelic recombination has been characterized by the intralocus exchange of small minicassettes of nucleotides that can occur throughout the length of the gene and has no single prominent breakpoint (A. L. Hughes, M. K. Hughes, and Watkins 1993; Jakobsen, Wilson, and Easteal 1998). Until more recently, it was not known whether these well-established patterns of MHC evolution were also found in species other than the mammalian model organisms.

Despite many elements of conserved structure and function, MHC evolution in nonmammalian taxa differs from well-defined norms. For instance, class II gene evolution in birds differs from that in mammals; the loci have a relatively recent origin in birds (Edwards, Wakeland, and Potts 1995; Hess and Edwards 2002). In bony fishes, the class I and class II genes, linked in a common region in all other vertebrates, are found on separate chromosomes (Bingulac-Popovic et al. 1997). However, this appears to be a derived trait in bony fishes because both class I and class II genes of sharks are closely linked (Ohta et al. 2000). Class Ia lineages in salmonid fishes share transspecies polymorphism among divergent taxa, whereas class II alleles cluster in a species-specific manner, just opposite the pattern in mammalian model organisms (Shum et al. 2001). Salmonid fishes also have much higher levels of class Ia than class II polymorphism. In general, recombination plays a more prominent role in teleost class Ia evolution than in mammalian benchmarks (Shum et al. 2001; Consuegra et al. 2005). Intragenic recombination in salmonids typically involves entire exons, and a prominent breakpoint for genetic exchange is easily identifiable. Authors have speculated that these patterns of evolution might be due to the nonlinkage of class I and class II loci as seen in bony fishes (Shum et al. 2001).

Xenopus laevis, the African clawed frog, is the first ectothermic vertebrate from which class I proteins were isolated (Flajnik et al. 1984). In this species, there is a single

[1] Present address: Department of Forestry and Natural Resources, Purdue University.

Key words: transspecies evolution, recombination, exon shuffling, Akaike information criterion (AIC), *Xenopus*, MHC.

E-mail: dbos@purdue.edu.

Mol. Biol. Evol. 23(1):137–143. 2006
doi:10.1093/molbev/msj016
Advance Access publication September 14, 2005

MHC class Ia locus with diploid inheritance patterns (Shum et al. 1993). This locus is highly polymorphic compared to mammals but is similar to the variation found in salmonid fishes. Another unusual aspect of *Xenopus* class Ia is the existence of two ancient allelic lineages (Flajnik et al. 1999). These lineages are very distinct as alleles belonging to different lineages are as divergent as MHC alleles from mouse and human. Linkage of class Ia and class II genes in *X. laevis* indicates a single MHC genomic region like many vertebrates, but unlike bony fishes (Nonaka et al. 1997).

Our motivation is to ascertain patterns of MHC class I evolution, by characterizing recombination, polymorphism, and extent of transspecies evolution in the class Ia gene of wild-caught *Xenopus* frogs. Previously established differences in the linkage of class I and class II regions between *Xenopus* and salmonid fishes also allow us to compare results and interpret these in light of hypotheses regarding the influence of linkage on MHC evolution. Differing patterns of class Ia evolution among these taxa would support the hypothesis that the mode of MHC evolution seen in salmonid fishes might be due to the nonlinkage of class I and class II genes. This mode of evolution was reported by Shum et al. (2001) and is distinguished by ancient class I lineages, high levels of polymorphism, and frequent recombination between the peptide-binding region (PBR)–coding exons. However, class Ia evolution in *X. laevis* that is similar to that of fishes supports the notion that nonlinkage of class Ia and class II genes alone may be insufficient to explain the mode of class Ia evolution reported for salmonids.

Materials and Methods

Data Collection

We extracted total RNA from *X. laevis* blood samples using the TRIzol protocol following the manufacturer's recommendations (Invitrogen). Total RNA was used to make cDNA in a Superscript One-Step reverse transcriptase–polymerase chain reaction (PCR) kit (Invitrogen, Carlsbad, CA), and first-strand synthesis was performed at 55°C for 25 min. Immediately after the first-strand synthesis, PCR was employed on cDNA with primers designed to amplify exons 2–4 of the MHC class Ia gene (forward primer: 5′-GTCACTCCCTGCGYTAYTAT-3′ and reverse primer: 5′-TTTCTCCTTCAGGCTGCTGT-3′). Primers were designed using Primer3 (Rozen and Skaletsky 2000) from known *X. laevis* sequences (Flajnik et al. 1999). The thermal profile used to amplify MHC fragments was optimized to minimize the occurrence of in vitro recombination (Judo, Wedel, and Wilson 1998). We cloned PCR products into the pCR 4 TOPO TA plasmid following the manufacturer's recommendations (Invitrogen), and recombinant DNA was transformed into TOP-10 *Escherichia coli* cells. *Escherichia coli* cells were plated onto Luria-Bertani (LB) agar and grown overnight at 37°C after which 6–10 individual colonies were picked and grown in LB broth at 37°C for 16 h.

A total of 5 ml of LB broth per cell matrix was removed, and plasmid DNA was extracted using alkaline lysis minipreps (Sambrook, Fritsch, and Maniatis 1989). We sequenced the MHC insert in both directions using BigDye v3.1 chemistry and an ABI 3730 automated sequencer. ABI trace files were edited using Bioedit (Hall 1999), and sequences were aligned using ClustalW (Thompson, Higgins, and Gibson 1994). Eleven new sequences were isolated from 11 *X. laevis* chromosomes; these sequences were independently verified from two to six separate colonies (GenBank accession numbers: DQ149596–DQ149606). Additional sequences were obtained but were not recovered multiple times and were excluded from the following analyses. New sequences were added to other known class Ia sequences of exons 2, 3, and 4 from frogs (GenBank accession numbers—*Xenopus tropicalis*: AY204558, AY204559 [*X. tropicalis* is also termed *Silurna tropicalis* but is listed herein as a congeneric *Xenopus* species due to strong monophyly of this species with other *Xenopus* {Evans et al. 2004}]; *Xenopus ruwenzoriensis*: AF497525–AF497528; and *X. laevis*, *Rana pipiens*, and a laboratory-bred interspecies hybrid of *X. laevis-Xenopus gilli*: AF185579–AF185588).

Statistical Analysis

We used various statistical methods to investigate evolutionary relationships and intragenic recombination. The program Maxchi was used to detect the occurrence of recombination events within *X. laevis* samples because it performs well in simulations and had a low false error rate (Posada and Crandall 2001*c*). The program RDP2 (Martin, Williamson, and Posada 2005) characterized intragenic recombination by identifying breakpoints and alleles created by recombination. The *P* value of significant differences used to infer recombination was set at 0.000005 with a window size of 20 nt to minimize the false-positive error rate (Martin, Williamson, and Posada 2005). To estimate population parameters (i.e., mutation [$\theta = 4N\mu$] and recombination [$\rho = 4Nr$]), we used the maximum likelihood (ML) method implemented in LDhat because this method uses a finite-sites model appropriate for highly divergent sequences (McVean, Awadalla, and Fearnhead 2002).

Evolutionary relationships were reconstructed among known *Xenopus* MHC class Ia sequences. Prior to phylogenetic analysis, we tested for the loss of information in these data due to saturation using the index of substitution saturation (I_{SS}) (Xia et al. 2003). We compared these values with the symmetric critical index of substitution saturation ($I_{SS.C}$) scores. For this procedure, we estimated the proportion of invariable sites and tested each data partition (see below) separately; each test included all sequences ($n = 27$). We used model-based algorithms to investigate evolutionary relationships among sequences (Bos and Posada 2005). The best approximating model of nucleotide evolution for these data was determined using Akaike's information criterion (AIC) (Akaike 1974). ML scores of candidate models were calculated using PAUP* 4.0 (Swofford 1998) and AIC scores computed in Modeltest (Posada and Crandall 1998). Employing the best approximating model, genetic distance and phylogenetic relationships were estimated using ML optimization.

A traditional bifurcating phylogenetic tree may not accurately represent evolutionary relationships among a population sample because of genetic exchange (Posada and Crandall 2001*a*). Therefore, we reconstructed separate trees by partitioning these data into congruent segments with

Table 1
***Xenopus laevis* Recombinant Sequences at the *UAA* Locus**

Recombinant Sequence	Nucleotide Breakpoint	Potential Parent Sequence
Xela r	230	*Xela *08*/unknown
*Xela *03*	328	*Xela *06/Xela *11*
Xela j	252	*Xela *05/Xela *06*
*Xela *06*	252	*Xela *07/Xela *02*
*Xela *08*	252	*Xela f/Xela *02*
*Xela *11*	252	*lg-a/c1/Xela *02*

Table 2
Population Parameters of *Xenopus laevis* Nucleotide Sequences (n = 19 samples chromosomes)

Protein Domain	α1	α2	α3	Total
Number of nucleotide sites	255	279	247	781
Number of variable sites	105	75	29	209
Number of haplotypes	16	18	16	19
Per site nucleotide diversity (π)	0.165	0.091	0.022	0.089
Per locus mutation rate (θ)	n.d.	n.d.	n.d.	46.08
Per locus recombination rate (ρ)	n.d.	n.d.	n.d.	66.39

NOTE.—n.d., not determined.

shared evolutionary history on either side of a putative recombination breakpoint. ML (Felsenstein 1981) and Neighbor-Joining (NJ; Saitou and Nei 1987) reconstructions were performed to test hypotheses regarding lineage assortment among species. ML trees were compared to a range of a priori topologies corresponding to different levels of trans-species lineage sharing among taxa. Tree comparisons were performed using Shimodaira-Hasegawa tests (Shimodaira and Hasegawa 1999; Shimodaira 2002) implemented in the program package Consel (Shimodaira and Hasegawa 2001). One thousand bootstrap pseudoreplications were used to estimate support for nodes in the NJ tree, and >50% bootstrap support in resulting topologies is shown.

Results

Data

Overall, the data (including *Rana* outgroup samples) consist of 27 sequences and have 397 polymorphic sites out of a total of 781 nt; on average, alleles differ by 99.30 nt. The $I_{SS.C}$ represents the I_{SS} value beyond which data fail to recover a true phylogenetic tree. The α1 partition is 255 bp, and the α2/α3 domain partition is 526 bp; the respective I_{SS} values are 0.261 and 0.228. The $I_{SS.C}$ are 0.682 and 0.715, respectively, and these are significantly larger than the respective I_{SS} scores ($P < 0.001$).

Population Parameters

Intragenic recombination plays a prominent role in *X. laevis* MHC evolution and is responsible for the creation of a number of new alleles. Among *X. laevis* sequences (including the *X. laevis-X. gilli* hybrids), the number of alleles created through recombination is at least 6 out of 19, over 30% of alleles in this data set (table 1). The parameters of RDP2 were set conservatively to avoid false-positive identification of recombination events, so this number represents a minimum number of recombinant alleles. Estimation of the population parameters shows relatively high mutation and recombination rates (table 2). Estimated values indicate that past mutation and recombination events both operate on the same scale and play a major role in shaping variation at this locus.

The recombination breakpoint also is shared among alleles, indicating that recombination is not free. The breakpoint of the recombinant alleles indicates the likely crossover location of the genetic exchange, which commonly occurs in intron II in these data. The size of the DNA fragment involved in the exchanges typically encompasses about the first 255 nt or the entire α1 domain–coding exon (tables 1 and 2). This type of recombination leads to intralocus allelic "exon shuffling" that creates new arrangements of existing variation in the PBR. This pattern is very different from that seen in humans, where genetic exchange involves much smaller fragments.

Some other trends in the pattern of recombination in *X. laevis* are noteworthy. For instance, some alleles are involved in genetic exchange more often than others. The occurrence of genetic exchange often involves *Xela-UAA*02* in these data. This allele is involved in creating at least three new alleles; *Xela-UAA*06* could be a parent sequence for two additional recombinant alleles. A bias in alleles involved in recombination has also been detected in salmonid fishes (Shum et al. 2001). Inspection of recombinants reveals that two recombinant alleles (alleles *UAA*06* and *UAA*11*) have identical α1 domain sequences, but the α2 and α3 domains of these two alleles are different. Finally, the formation of recombinant alleles is not restricted to closely related alleles as two highly divergent alleles can recombine (e.g., UAA^f and *UAA*02*). The apparent ongoing genetic exchange results in a high level of recombination that is likely to affect the evolutionary relationships among alleles and different domains of alleles.

Evolutionary Relationships

The evolutionary relationships among MHC class Ia alleles in *Xenopus* species were determined using ML estimation of genetic distances and NJ topology reconstruction. Tree reconstruction was done separately on two segments of the sequence, partitioning the nucleotide sequence fragment coding the α1 domain as one segment and the α2 and α3 domains together as the other segment. This partition was chosen to maximize the detection of different evolutionary histories due to genetic exchange and provides a means for confirming the presence of recombination in this data set. In the phylogenetic trees of the α1 and α2/α3 domains, recombinant alleles identified with the program RDP2 were found to be in different clades. These alleles moved across nodes with >50% bootstrap support to associate with different sets of alleles in each tree reconstruction. The translocation of alleles to different parts of the tree topology is consistent with patterns of recombination detected with RDP2.

The best approximating model selected for α1 domains differs from that chosen for the α2/α3 sequences (table 3). The best-fitting model for the α1 domain sequence evolution is (the transitional model with a gamma distribution for rates among sites TIM + Γ) (Posada and Crandall 2001*b*);

Table 3
Best Approximating Model Parameters for Data Partitions

MHC Domain	α1	α2/α3
Selected model	TIM + Γ	K81uf + Γ
Base frequency		
A	0.328	0.276
T	0.201	0.215
G	0.253	0.295
C	0.218	0.214
Substitution rates		
A↔C	1.00	1.00
A↔G	2.43	1.89
A↔T	1.41	0.75
C↔G	1.41	0.75
C↔T	3.89	1.89
G↔T	1.00	1.00
Rate variation		
α parameter	0.576	0.466

for the α2/α3 sequence fragment, the K81uf + Γ (Kimura 1981) model was the best fit. Differences between data partitions are also found in tree topology, although the conservative Shimodiara-Hasagawa test indicates that the trees are not significantly different ($P > 0.05$).

The topology showing relationships among α1 domain sequences shows mixing of alleles from different species (fig. 1). Pairs of alleles from a species form well-supported groups in some cases, but both *X. tropicalis* and *X. ruwenzoriensis* α1 domains are intermingled together with *X. laevis* sequences. One group of *X. laevis* alleles is closely related and forms a tight cluster that has 100% bootstrap support; this group comprises five recombinant sequences. Most well-supported clades are near the tips of the tree and are comprised of only a few sequences; branches in the more basal parts of the tree are typically shorter than many terminal branches.

The evolutionary relationships reconstructed for α2/α3 domain sequences were different in some ways from the α1 domain sequence topology (fig. 2). For instance, some alleles segregated by species rather than being intermingled. In this tree, *X. tropicalis* alleles are basal in the topology and paraphyletic with respect to a clade of *X. laevis*, *X. gilli*, and *X. ruwenzoriensis* alleles. This topology establishes an incomplete separation of *X. tropicalis* alleles and monophyly of other sequences of the ingroup. All *X. ruwenzoriensis* alleles form a monophyletic cluster nested within a larger clade of *X. laevis* and *X. gilli*. There is a closely related group that forms a tight cluster similar to that seen in the α1 domain tree, but the cluster is comprised of different sequences and contains only one recombinant allele. This tree also has a mixture of both long and short terminal branches, but compared to the α1 domain tree, branches on the α2/α3 tree are much shorter.

Several a priori hypotheses were compared to the ML and NJ topologies for both the α1 and α2/α3 sequence partitions. A priori hypotheses are not exhaustive but are designed to gauge the level of transspecies allelic sharing among progressively more distantly related species. For instance, transspecies polymorphism could be confined to two relatively closely related species, such as *X. laevis* and *X. ruwenzoriensis*, or it may extend to more distantly related taxa, such as *X. laevis* and *X. tropicalis*. Hypotheses are designed to gauge similarity of the observed level of

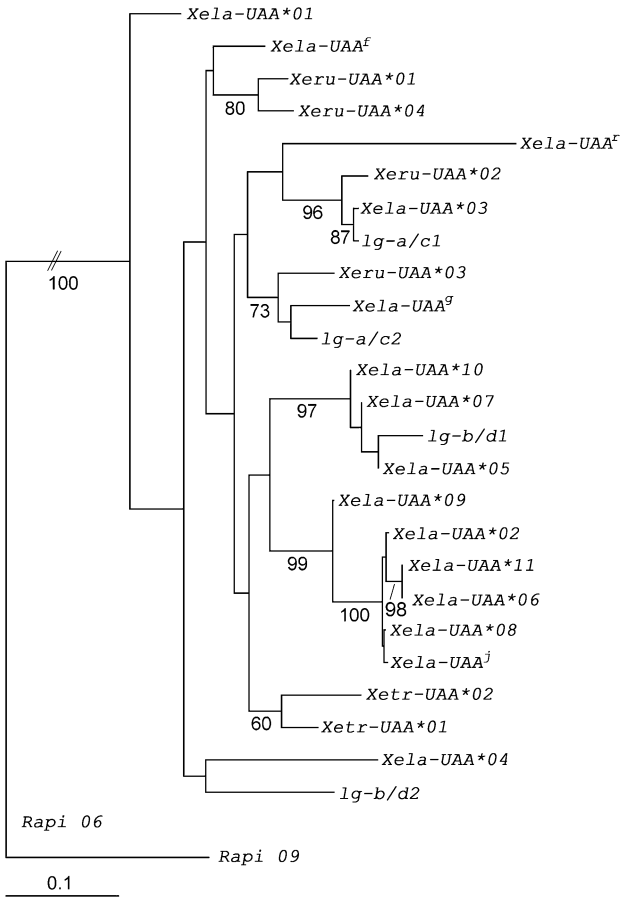

FIG. 1.—Evolutionary relationships of *Xenopus* class Ia sequences using α1 domain sequences. Numbers indicate bootstrap support for nodes. All branches shown to the scale at the bottom left of the figure, except the branch leading to the outgroup, which was shortened for graphical clarity of the remaining branches of the tree. *Rapi*, *Rana pipiens*; *Xela*, *Xenopus laevis*; *Xetr*, *Xenopus tropicalis*; *Xeru*, *Xenopus ruwenzoriensis*; and *lg*, *X. laevis-Xenopus gilli* laboratory hybrid.

transspecies polymorphism with constraints that represent various combinations of species that share alleles. Representative hypotheses include (1) no transspecies polymorphisms or reciprocal monophyly of species ((*X. laevis*),(*X. ruwenzoriensis*),(*X. tropicalis*),(*R. pipiens*)), (2) unconstrained *X. laevis* (*X. laevis*, (*X. ruwenzoriensis*),(*X. tropicalis*),(*R. pipiens*)), (3) transspecies polymorphism among two closely related species or unconstrained *X. laevis* and *X. ruwenzoriensis* (*X. laevis*, *X. ruwenzoriensis*, (*X. tropicalis*),(*R. pipiens*)), and (4) monophyletic *X. laevis* ((*X. laevis*), *X. ruwenzoriensis*, *X. tropicalis*, (*R. pipiens*)). Note that the ML tree serves to represent another hypothesis, namely, that transspecies polymorphism can occur throughout the genus *Xenopus*. Results of the Shimodaira-Hasegawa tests for both data partitions show that there is no significant difference between the ML and NJ trees (table 4). However, hypotheses 1 and 4 are significantly different from unconstrained optimal trees for both data partitions. For the α1 domain, hypothesis 2 is also significantly different from the optimal tree. For the α2/α3 data partition, tree constraints for hypotheses 2 and 3 resulted in the same phylogenetic reconstruction. Only hypothesis 3, where the only

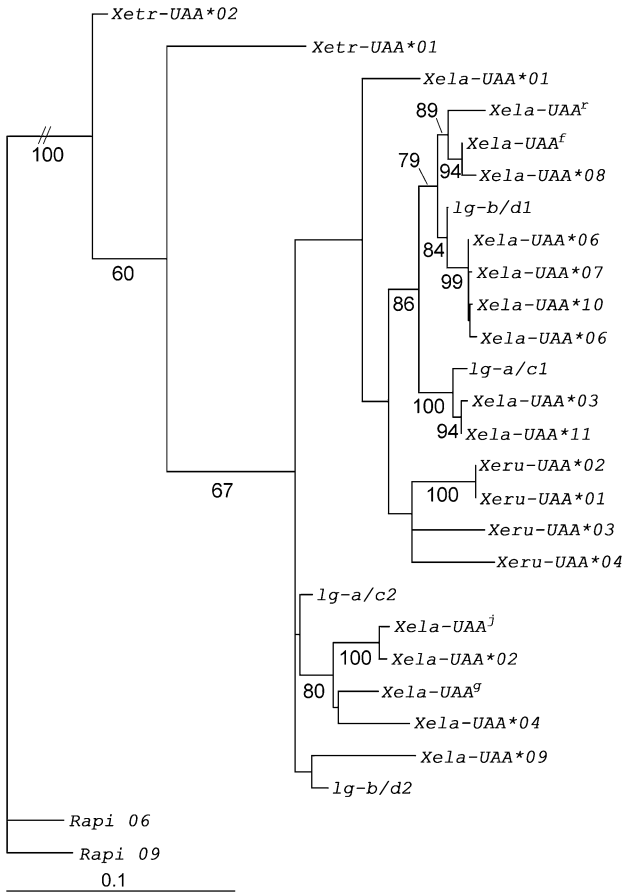

FIG. 2.—Evolutionary relationships of *Xenopus* class Ia sequences with *Rana* outgroup using α2 and α3 domain sequences. Numbers indicate bootstrap support for nodes. All branches scaled as in figure 1.

constraint is placing *R. pipiens* as an outgroup to *Xenopus*, is not significantly different from the unconstrained ML and NJ topologies for all data partitions.

Discussion

Recombination

Intragenic recombination plays a prominent role in creating and maintaining diversity in the class Ia gene of *X. laevis*. Here, a clear pattern of reticulate evolution is characterized by high rates and phylogenetic inconsistencies between the α1 domain and sites in the remaining 3′ exons of the gene. The nature of *X. laevis* recombination events is in contrast to the reticulate evolution seen among alleles of humans for three class I loci. Also, the pattern seen here would not be expected from in vitro chimera formation, where random breakpoints would be expected; therefore, we interpret this result to be due to in vivo allelic recombination. Although we have identified one recombination breakpoint in these sequences, the location of the other endpoint remains uncertain. However, given strong linkage disequilibrium with functionally related low molecular mass polypeptide (LMP) and transporter associated with antigen processing (TAP) loci (Namikawa et al. 1995; Ohta et al. 2003), genetic exchange events are expected to be much smaller than the distance separating these loci (Andolfatto and Nordborg 1998).

Table 4
Statistical Comparison of Phylogenetic Hypotheses for *Xenopus* Class Ia Alleles

Topology	ln L	δ ln L	S-H[a]	W S-H[b]
α1 Domain trees				
ML	−1844.420	—	—	—
3	−1844.905	0.5	0.850	0.812
NJ	−1858.748	14.3	0.491	0.281
2	−1920.962	76.5	0.000	0.000
4	−2009.004	164.6	0.000	0.000
1	−2058.469	214.0	0.000	0.000
α2/α3 Domain trees				
ML	−3003.147	—	—	—
3	−3003.382	0.2	0.848	0.861
2	−3003.382	0.2	0.848	0.877
NJ	−3003.710	0.6	0.743	0.727
1	−3148.052	144.9	0.000	0.000
4	−3168.448	165.3	0.000	0.000

[a] Shimodaira-Hasegawa (1999) test.
[b] Weighted Shimodaira-Hasegawa (2002) test.

The high rate of genetic exchange in *X. laevis* could be influenced by large intron size, as seen in salmonid fishes (Shum et al. 2002). We attempted to PCR amplify the intron intervening exons 2 and 3 of the class Ia gene but were unable to recover any specific amplification product. However, bioinformatic exploration of a draft of the *X. tropicalis* genome indicates that a very large intron separates exons 2 and 3 of a class I gene. The largest intron in a class I gene was reported by Shiina et al. (2005); however, the draft genome of *X. tropicalis* indicates that the corresponding intron is much larger (D. H. Bos, unpublished data). If intron size in a closely related species is an indication of class I intron size in *X. laevis*, then it is probable that recombination rates in the class I gene are influenced by these genomic features. The high recombination rate and the location of recombination events support the notion that these events may relieve possible Hill-Robertson interference (Hill and Robertson 1966; Otto and Barton 1997) on the evolution of PBR domains with differing selective pressures (Kaufman et al. 1992).

Phylogenetics of *Xenopus* MHC

Class I alleles fall into lineages that are long lasting and predate certain speciation events within the genus *Xenopus*. These conclusions are upheld when evolutionary relationships are reconstructed with either α1 domain sequences or α2/α3 domain sequences. However, we note that although the sequences may be useful in terms of the test of substitution saturation, our conclusions are based on partitioned data, resulting in short sequences which have less power to resolve and support relationships. In reconstructing evolutionary analysis on MHC genes, other authors have used the "total evidence" approach (Kluge 1989) to establish the species distribution of allelic lineages (e.g., Shum et al. 2001). Such an approach using ML methods on these data also confirm the above conclusions (data not shown).

Based on our reconstruction of evolutionary relationships, transspecies sharing of allelic lineages among certain

divergent *Xenopus* species takes place. Comparison of this hypothesis with scenarios that constrain allelic separation among various species groups confirms this result. Class Ia transspecies evolution in *Xenopus* extends to species thought to be much more evolutionarily divergent than the species among which allele-lineage sharing is commonly found in primate class I genes. While the class Ia sequences from all *Xenopus* species form a monophyletic clade, *X. tropicalis* appears to have diverged from other *Xenopus* species prior to the formation of the extant class Ia lineages. This is not surprising because the divergence time of *X. tropicalis* and the *Xenopus* common ancestor is estimated between 50 and 81 MYA (Evans et al. 2004). Undoubtedly, the transspecies mode of evolution in *Xenopus* is strongly influenced by natural selection acting on the class Ia locus (D. H. Bos and B. Waldman, unpublished data), which tends to extend allele retention.

Pattern of Class Ia MHC Evolution in *X. laevis*

MHC class Ia evolution in *X. laevis* is more similar to MHC evolution in salmonid fishes than mammals. Salmonid fishes and *Xenopus* frogs show similarity in at least three aspects of MHC evolution: (1) levels of polymorphism exceeding that found in primates, (2) a distinct pattern of genetic exchange, and (3) class Ia lineages that are maintained for long periods of time. All three of these characteristics differ from the patterns of class Ia evolution found in primates. For instance, primates exhibit no allele sharing among species thought to last share a common ancestor approximately 35 MYA (Vogel et al. 1999), and recombination events are mostly spread throughout the gene sequence and involve very short sequence tracts (Jakobsen, Wilson, and Easteal 1998).

The pattern of recombination, polymorphism, and transspecies allele sharing of class Ia sequences described in salmonid fishes, and now found in *X. laevis*, may be the result of certain genomic features of the MHC region. Shum et al. (2001) suggested that these features of evolution might be due to the separation of MHC class I and class II regions onto different chromosomes. The nonlinkage of class I and class II regions may influence patterns of MHC evolution by altering the potential Hill-Robertson constraints or selection for conserved haplotype blocks and linkage disequilibrium. However, class Ia evolution in *X. laevis* and salmonid fishes shows similarities in polymorphism, recombination, and transspecies polymorphism. These similarities support the idea that the separation of the class I and class II regions onto different chromosomes is alone insufficient to account for these patterns of MHC evolution.

Other factors may play a role in determining class Ia gene evolution. For instance, both salmonid and *X. laevis* class Ia and class I–processing pathway genes are located close to one another (Namikawa et al. 1995; Takami et al. 1997; Ohta et al. 1999). A likely result of this linkage is that distinctive allelic associations exist among class Ia and class I–processing genes in *Xenopus* frogs and other taxa but are not known in primates and mice (Joly et al. 1998; Kaufman 1999; Ohta et al. 2003). Therefore, it is possible that the pattern of evolution common to *Xenopus* is due in part to the number of class Ia loci, linkage, and possible coevolution of this suite of genes. Additionally, the location of the class Ia gene in a "central" rather than "distal" position within the MHC region may influence patterns of evolution, as suggested by Nonaka et al. (1997).

Acknowledgments

Martin Flajnik and Neil Gemmell provided useful comments on this manuscript, and Louis Du Pasquier provided ideas and invaluable technical assistance in the laboratory and in the animal breeding and care facility. Scott Edwards and two anonymous reviewers provided constructive comments on an earlier version of this manuscript. This research was supported by the Marsden Fund (Royal Society of New Zealand) and a Ph.D. scholarship from the University of Canterbury.

Literature Cited

Akaike, H. 1974. A new look at the statistical model identification. IEEE Trans. Automat. Control **19**:716–723.

Andolfatto, P., and M. Nordborg. 1998. The effect of gene conversion on intralocus associations. Genetics **148**:1397–1399.

Bingulac-Popovic, J., F. Figueroa, A. Sato, W. S. Talbot, S. L. Johnson, M. Gates, J. H. Postlethwait, and J. Klein. 1997. Mapping of MHC class I and class II regions to different linkage groups in the zebrafish, *Danio rerio*. Immunogenetics **46**: 129–134.

Bontrop, R. E., N. Otting, N. de Groot, and G. G. M. Doxiadis. 1999. Major histocompatibility complex class II polymorphisms in primates. Immunol. Rev. **167**:339–350.

Bos, D. H., and D. Posada. 2005. Using models of nucleotide evolution to build phylogenetic trees. Dev. Comp. Immunol. **29**:211–227.

Consuegra, S., H.-J. Megens, H. Schaschl, K. Leon, R. J. M. Stet, and W. C. Jordan. 2005. Rapid evolution of the MH class I locus results in different allelic compositions in recently diverged populations of Atlantic salmon. Mol. Biol. Evol. **22**:1095–1106.

Edwards, S. V., E. K. Wakeland, and W. K. Potts. 1995. Contrasting histories of avian and mammalian MHC genes revealed by class II B sequences from songbirds. Proc. Natl. Acad. Sci. USA **92**:12200–12204.

Evans, B. J., D. B. Kelley, R. C. Tinsley, D. J. Melnick, and D. C. Cannatella. 2004. A mitochondrial DNA phylogeny of African clawed frogs: phylogeography and implications for polyploid evolution. Mol. Phylogenet. Evol. **33**:197–213.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. **17**:368–376.

Figueroa, F., E. Gunther, and J. Klein. 1988. MHC polymorphism pre-dating speciation. Nature **335**:265–267.

Flajnik, M. F., J. Kaufman, P. Riegert, and L. Du Pasquier. 1984. Identification of class I major histocompatibility complex encoded molecules in the amphibian *Xenopus*. Immunogenetics **20**:134–143.

Flajnik, M. F., Y. Ohta, A. S. Greenberg, L. Salter-Cid, A. Carrizosa, L. Du Pasquier, and M. Kasahara. 1999. Two ancient allelic lineages at the single classical class I locus in the *Xenopus* MHC. J. Immunol. **163**:3826–3833.

Hall, T. 1999. Bioedit: a user-friendly biological sequence alignment editor and analysis program for Window 95/98/NT. Nucleic Acids Symp. Ser. **41**:95–98.

Hess, C. M., and S. V. Edwards. 2002. The evolution of the major histocompatibility complex in birds. Bioscience **52**:423–431.

Hill, W. G., and A. Robertson. 1966. The effect of linkage on limits to artificial selection. Genet. Res. **8**:269–294.

Hughes, A. L., M. K. Hughes, and D. I. Watkins. 1993. Contrasting roles of interallelic recombination at the HLA-A and HLA-B loci. Genetics **133**:669–680.

Hughes, A. L., and M. Yeager. 1998. Natural selection at major histocompatibility complex loci of vertebrates. Ann. Rev. Genet. **32**:415–435.

Jakobsen, I. B., S. R. Wilson, and S. Easteal. 1998. Patterns of reticulate evolution for the classical class I and II HLA loci. Immunogenetics **48**:312–323.

Joly, E., A. F. Le Rolle, A. L. Gonzalez, B. Mehling, J. Stevens, W. J. Coadwell, T. Hunig, J. C. Howard, and G. W. Butcher. 1998. Co-evolution of rat TAP transporters and MHC class I RT1-A molecules. Curr. Biol. **8**:169–172.

Judo, M. S. B., A. B. Wedel, and C. Wilson. 1998. Stimulation and suppression of PCR-mediated recombination. Nucleic Acids Res. **26**:1819–1825.

Kaufman, J. 1999. Co-evolving genes in MHC haplotypes: the "rule" for nonmammalian vertebrates? Immunogenetics **50**:228–236.

Kaufman, J., R. Anderson, D. Avila, J. Engberg, J. Lambris, J. Salomonsen, K. Welinder, and K. Skjodt. 1992. Different features of the MHC class I heterodimer have evolved at different rates. J. Immunol. **142**:1532–1546.

Kelley, L., J. Walter, and J. Trowsdale. 2005. Comparative genomics of major histocompatibility complexes. Immunogenetics **56**:683–695.

Kimura, M. 1981. Estimation of evolutionary distances between homologous nucleotide sequences. Proc. Natl. Acad. Sci. USA **78**:454–458.

Kluge, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicurates* (Boidae, Serpentes). Syst. Zool. **38**:7–25.

Lawlor, D. A., F. E. Ward, P. D. Ennis, A. P. Jackson, and P. Parham. 1988. *HLA-A* and *HLA-B* polymorphism predate the divergence of humans and chimpanzees. Nature **335**:268–271.

Martin, D. P., C. Williamson, and D. Posada. 2005. RDP2: recombination detection and analysis from sequence alignments. Bioinformatics **21**:260–262.

McVean, G., P. Awadalla, and P. Fearnhead. 2002. A coalescent-based method for detecting and estimating recombination from gene sequences. Genetics **160**:1231–1241.

Namikawa, C., L. Salter-Cid, M. F. Flajnik, Y. Kato, M. Nonaka, and M. Sasaki. 1995. Isolation of Xenopus *LMP-7* homologues: striking allelic diversity and linkage to MHC. J. Immunol. **155**:1964–1971.

Nonaka, M., C. Namikawa, Y. Kato, M. Sasaki, L. Salter-Cid, and M. F. Flajnik. 1997. Major histocompatibility complex gene mapping in the amphibian *Xenopus* implies a primordial organization. Proc. Natl. Acad. Sci. USA **94**:5789–5791.

Ohta, Y., K. Okamura, E. C. McKinney, S. Bartl, K. Hashimoto, and M. F. Flajnik. 2000. Primitive synteny of vertebrate histocompatibility complex class I and class II genes. Proc. Natl. Acad. Sci. USA **97**:4712–4717.

Ohta, Y., S. J. Powis, W. J. Coadwell, D. E. Haliniewski, Y. Liu, H. Li, and M. F. Flajnik. 1999. Identification and mapping of *Xenopus* TAP2 genes. Immunogenetics **49**:171–182.

Ohta, Y., S. J. Powis, R. L. Lohr, M. Nonaka, L. Du Pasquier, and M. F. Flajnik. 2003. Two highly divergent ancient allelic lineages of the transporter associated with antigen processing (TAP) gene in *Xenopus*: further evidence for co-evolution among MHC class I region genes. Eur. J. Immunol. **33**:3017–3027.

Otto, S. P., and N. H. Barton. 1997. The evolution of recombination: removing the limits to natural selection. Genetics **147**:879–906.

Parham, P., and T. Ohta. 1996. Population biology of antigen presentation by MHC class I molecules. Science **272**:67–74.

Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14**:817–818.

———. 2001*a*. Intraspecific gene genealogies: trees grafting into networks. Trends Ecol. Evol. **16**:37–45.

———. 2001*b*. Selecting models of nucleotide substitution: an application to human immunodeficiency virus (HIV-1). Mol. Biol. Evol. **18**:897–906.

———. 2001*c*. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. Proc. Natl. Acad. Sci. USA **98**:13757–13762.

Rozen, S., and H. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. Pp. 365–386 *in* S. Krawetz and S. Misener, eds. Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, N.J.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

Sambrook, E., F. Fritsch, and T. Maniatis. 1989. Molecular cloning. Cold Spring Harbor Press, Cold Spring, N.Y.

Shiina, T., J. M. Dijkstra, S. Shimizu et al. (15 co-authors). 2005. Interchromosomal duplication of major histocompatibility complex class I regions in rainbow trout (*Oncorhynchus mykiss*), a species with a presumably recent tetraploid ancestry. Immunogenetics **56**:878–893.

Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. **51**:492–508.

Shimodaira, H., and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. **16**:1114–1116.

———. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics **17**:1246–1247.

Shum, B. P., D. Avila, L. Du Pasquier, M. Kasahara, and M. F. Flajnik. 1993. Isolation of a classical MHC class I cDNA from an amphibian. Evidence for only one class I locus in the *Xenopus* MHC. J. Immunol. **151**:5376–5386.

Shum, B. P., L. A. Guethlein, L. R. Flodin, M. A. Adkinson, R. P. Hedrick, R. B. Nehring, R. J. M. Stet, C. Secombes, and P. Parham. 2001. Modes of salmon MHC class I and II evolution differ from the primate paradigm. J. Immunol. **166**: 3297–3308.

Shum, B. P., P. M. Mason, K. E. Magor, L. R. Flodin, R. J. M. Stet, and P. Parham. 2002. Structures of two major histocompatibility complex class I genes of the rainbow trout (*Oncorhynchus mykiss*). Immunogenetics **54**:193–199.

Swofford, D. L. 1998. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4.0. Sinauer, Sunderland, Mass.

Takami, K., Z. Zaleska-Rutczynska, F. Figueroa, and J. Klein. 1997. Linkages of *LMP*, *TAP*, and *RING3* with *Mhc* class I rather than class II genes in the zebrafish. J. Immunol. **159**:6052–6060.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

Vogel, T. U., D. T. Evans, J. A. Urvater, D. H. O'Connor, A. L. Hughes, and D. I. Watkins. 1999. Major histocompatibility complex class I genes in primates: co-evolution with pathogens. Immunol. Rev. **167**:327–337.

Xia, X., Z. Xie, M. Salemi, L. Chen, and Y. Wang. 2003. An index of substitution saturation and its application. Mol. Phylogenet. Evol. **26**:1–7.

Scott Edwards, Associate Editor

Accepted August 31, 2005