The *MHC Class Ia* Genes in Chenfu's Treefrog (*Zhangixalus chenfui*) Evolved via Gene Duplication, Recombination, and Selection

Hu Chen †, Siqi Huang †, Ye Jiang, Fuyao Han, Qingyong Ni, Yongfang Yao, Huailiang Xu, Sudhanshu Mishra and Mingwang Zhang *

College of Animal Science and Technology, Sichuan Agricultural University, 211# Huimin Road, Chengdu 611130, China; aquaticpractitioner@stu.sicau.edu.cn (H.C.); seven17huang@dingtalk.com (S.H.); jiangye20170907@163.com (Y.J.); hanfuyao@126.com (F.H.); niqingyong@hotmail.com (Q.N.); yaoyongf@126.com (Y.Y.); huailxu@yahoo.com (H.X.); mishra.sudhanshu30@gmail.com (S.M.)
* Correspondence: mwzhangkiz@hotmail.com
† H.C. and S.H. equally contributed to this work.

check for updates

Simple Summary: Amphibians, the first terrestrial vertebrates, provide materials for adaptive evolutionary studies, such as the evolution of the major histocompatibility complex (MHC). To date, various MHC evolutionary mechanisms have been identified in frogs, but more research is needed to determine the evolutionary mechanisms of the frog MHC. The main purpose of this study was to evaluate polymorphisms in the *MHC class Ia* genes of the Chenfu's Treefrog. The *MHC class Ia* genes of the Chenfu's Treefrog have high polymorphism. The mechanisms responsible for the formation of the polymorphisms include gene duplication, recombination, and selection.

Abstract: The molecular mechanisms underlying the evolution of adaptive immunity-related proteins can be deduced by a thorough examination of the major histocompatibility complex (MHC). Currently, in vertebrates, there is a relatively large amount of research on MHCs in mammals and birds. However, research related to amphibian *MHC* genes and knowledge about the evolutionary patterns is limited. This study aimed to isolate the *MHC class I* genes from Chenfu's Treefrog (*Zhangixalus chenfui)* and reveal the underlying evolutionary processes. A total of 23 alleles spanning the coding region of *MHC class Ia* genes were identified in 13 individual samples. Multiple approaches were used to test and identify recombination from the 23 alleles. Amphibian *MHC class Ia* alleles, from NCBI, were used to construct the phylogenetic relationships in MEGA. Additionally, the partition strategy was adopted to construct phylogenetic relationships using MrBayes and MEGA. The sites of positive selection were identified by FEL, PAML, and MEME. In Chenfu's Treefrog, we found that: (1) recombination usually takes place between whole exons of *MHC class Ia* genes; (2) there are at least 3 loci for *MHC class Ia*, and (3) the diversity of genes in *MHC class Ia* can be attributed to recombination, gene duplication, and positive selection. We characterized the evolutionary mechanisms underlying *MHC class Ia* genes in Chenfu's Treefrog, and in so doing, broadened the knowledge of amphibian MHC systems.

Keywords: amphibian; paralog; immunogenetics; polymorphism; molecular evolution

1. Introduction

Genes of the major histocompatibility complex (MHC), as cell surface glycoproteins, are vital for adaptive immunoreactions. To distinguish between intracellular and extracellular antigens, there are two main subgroups of MHC: class I and class II. A molecule of *MHC class I* consists of a microglobulin chain (β_2 m chain) and a heavy chain (α chain). This molecule connects peptides from intracellular

pathogens and displays it at the surface of the cell where an immunoreaction is initiated upon recognition by CD8[+] T cells [1–3]. The heavy chain that is encoded by classical *MHC class I* genes has a cytoplasmic region, a transmembrane region, and three extracellular domains designated $\alpha1$, $\alpha2$, and $\alpha3$. There are seven exons in the *MHC class I* genes. Specifically, the leader peptide is encoded by exon 1; three extracellular domains are encoded by exons 2, 3, and 4 ($\alpha1$, $\alpha2$, and $\alpha3$, respectively); the transmembrane domain is encoded by exon 5; and the cytoplasmic tail is encoded by exons 6 and 7. The antigen-binding sites (ABS) are encoded by the second and third exons in the $\alpha1$ and $\alpha2$ domains, respectively [1,4]. Therefore, a high degree of genetic diversity exists in the second and third exons of the classical genes of *MHC class I* [5], and MHC genes have been the subject of evolutionary ecology studies in a wide variety of species over the last two decades [6].

In terms of research on the MHC, a lot of attention has been given to the mechanisms by which evolution can favor diversity in *MHC class I* genes. There are three mechanisms of particular interest, recombination [7–10], positive selection [11–14], and gene duplication [15,16]. Under the action of these three mechanisms, a high degree of gene polymorphism [17–19] and increased variation in loci numbers [15,20,21] exist in *MHC class I* genes. The high number of genetic polymorphisms of MHC genes is thought to offer organisms protection against a large spectrum of pathogens and therefore increases their fitness. This hypothesis is supported by the associations observed between disease susceptibility and MHC gene diversity in numerous vertebrates [22–24].

The evolution of *MHC class I* genes have mainly been studied in fish, mammals, and birds [6,10,25,26]. The relatively under studied group, amphibians, should be given special attention, not only because they are the first terrestrial vertebrates but because their unique living environment exposes them to a very different spectrum and diversity of microbes. Nonetheless, relatively few studies have explored amphibian *MHC class I* genes. The existing evolutionary studies on amphibian *MHC class I* genes have been reported in Urodela and Anura [27–31]. These studies have revealed that the evolution of *MHC class I* gene can be credited to recombination and positive selection [28,29]. Moreover, amphibian *MHC class I* genes commonly contain trans-species polymorphisms [29]. However, the evolutionary mechanisms responsible for amphibian *MHC class Ia* genes polymorphisms may be more complicated than previously thought [31,32]. Consequently, to gain a more systematic understanding of the evolution of *MHC class I* genes, more amphibians should be examined. Chenfu's Treefrog is a species of frog in the family Rhacophoridae endemic to China where it is found in Sichuan, Guizhou, Hubei, and Fujian provinces [33]. Its natural habitats are temperate forests, subtropical moist lowland forests, subtropical moist montane forests, subtropical moist shrubland, freshwater marshes, rural gardens, ponds, and irrigated land. The aim of this study was to identify *MHC class I* genes in the Chenfu's Treefrog. In addition, we investigated whether selective mechanisms and gene replication affected MHC mutations. This will help to further understand the evolution of the amphibian MHC class I.

2. Materials and Methods

2.1. Sampling and RNA Isolation

Thirteen healthy Chenfu's Treefrogs were collected from the Wawu Shan Nature Reserve in Sichuan, China. The experimental frogs were euthanized with 95% alcohol. The liver was dissected and frozen in liquid nitrogen, before freezing at −80 °C for storage until RNA extraction. Trizol[®] reagent was used to isolate the total RNA from each sample according to the manufacturer's protocol (TaKaRa, Otsu, Japan). At the wavelengths of 280 nm and 260 nm, the optical density absorption ratio (Bio-Rad) was measured to determine the RNA concentrations. Specimens with the rate of absorption from 1.7 to 2.2 were used for cDNA synthesis. RNA integrity was measured using gel electrophoresis. Sichuan Agricultural University's Animal Ethics Committee approved all experiments with live tree frogs with approval number, 71234-2015-0901.

2.2. cDNA Amplification and Cloning

Prime Script RT Reagent Kit (TaKara, Dalian, China) was used for 1st strand cDNA synthesis, which contained mixture of equal amounts of liver RNA (2 μg) from the 13 individuals. We first amplified a partial Chenfu's Tree frog MHC (ZcMHC) sequence with primers ZcMHC-F and ZcMHC-R (Table 1). PCR condition was a touchdown PCR (2× Taq PCR Mastermix, Tiangen, Beijing, China): initial denaturation at 94 °C for 5 min; 10 cycles for 30 s at 94 °C, 68 °C for 30 s (−1 °C per cycle), and 72 °C for 2 min; 25 cycles for 30 s at 94 °C, 58 °C for 30 s, and 72 °C for 2 min; final extension at 72 °C for 10 min; and cooled to 4 °C. PCR products were excised from agarose gel (Universal DNA Purification Kit, TIANGEN, China), cloned into pMD19-T (TaKaRa, Dalian, China), and 30 clones were sequenced (BBI Life Science Corporation Company, Shanghai, China).

Table 1. Primer sequences and function used in this study.

Primer Name	Primer Sequence (5-3′)	Applications
ZcMHC-F	5′CTGCGSWAYTATKABACWGCAGTCTC 3′	ZcMHC cloning
ZcMHC-R	5′TYCAGRCTGCTGTGSTCCACAT 3′	ZcMHC cloning
ZcMHC-F1	5′CCGTCAGTCCGATGGTGCGA 3′	ZcMHC 3′RACE outer
ZcMHC-F2	5′GAAGTGCCAACAACCGAAGGAGAC 3′	ZcMHC 3′RACE inner
ZcMHC-R1	5′GAGCCTCGTTCATGGTGGGGATA 3′	ZcMHC 5′RACE outer
ZcMHC-R2	5′GAGTCACAGGACGAGCCTGATGA T3′	ZcMHC 5′RACE inner
ZcMHC-F3	5′GAGGTGACACACAGGAGGAGATGGA 3′	ZcMHC coding region PCR
ZcMHC-R3	5′CGGCATRGTGTACGCTCAGGCTKT 3′	ZcMHC coding region PCR

Primers were designed using primer premier 5.0 (PREMIER Biosoft International, Palo Alto, CA).

We then obtained full-length cDNA sequences of *MHC class Ia* by 5′ and 3′ RACE PCR (SMART RACE cDNA Amplification Kit, Clontech, CA, USA), with primers ZcMHC-F1, ZcMHC-F2, ZcMHC-R1, and ZcMHC-R2 (Table 1) designed based on the partial sequence, from the mixed RNA as above. The PCR condition was the touchdown PCR as above. PCR products were excised from agarose gel, cloned into pMD19-T, and 30 clones were sequenced as above.

In order to obtain individual allelic sequences, individual liver RNAs (2 μg each) were separately subjected to cDNA synthesis (Prime Script RT Reagent Kit, TaKaRa, Dalian, China) and touchdown PCR of coding sequences with primers ZcMHC-F3 and ZcMHC-R3 (Table 1) designed based on the full-length cDNA sequence. PCR products were cloned into pEASY-T5 Zero vector (TransGen Biotech, Beijing, China). We picked 30 clones up of each PCR product and sequenced as above.

2.3. Analysis of ZcMHC Class Ia

Sequence assembly was performed by DNASTAR version 7.1 (DNASTAR, Inc, Madison WI, USA). Using BLASTp and BLASTn (http://www.ncbi.nlm.nih.gov), the deduced protein sequences and nucleotides were studied. In addition, the ORF was predicted using the Open Reading Frame Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). With the use of NetOGlyc 3.1 Server (http://www.cbs.dtu.dk/services/NetOGlyc-3.1/) and NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc), posttranslational modification forecasts of N- and O-glycosylation sites were determined. Furthermore, transmembrane domain topology for proteins was predicted using the TMHMM Server v2.0 (http://www.cbs.dtu.dk/services/TMHMM/).

2.4. Sequence Alignment and Allele Identification

In order to accurately obtain the full-length ORF fragments (about 1100 bp), we used forward and reverse primers to obtain the sequence from both ends for each clone. Two ABI trace files from each clone were examined, edited, and assembled using DNASTAR 7.1. ClustalW in MEGA5 was used to align the assembled sequences. Only the sequences that appeared more than twice were considered alleles and included.

Newly identified alleles were translated into amino acid sequences and aligned with human HLA [5]. The identification of major amino acid sites in disulfide bond formation and peptide binding was permitted in the amino acid sequences.

2.5. Statistical Study

2.5.1. Recombination Detection

MEGA5 (https://www.megasoftware.net/ was used to compute the distances of Poisson-corrected amino acid and the average pairwise nucleotide distances (Kimura-2-parameter model, K2P). A total of 1000 bootstrap replicates were used to obtain standard errors of the estimates. A number of approaches were used to detect the recombination in our data: BOOTSCAN, RDP, CHIMAERA, MAXCHI, GENECONV, 3SEQ, and SISCAN, which were all implemented in the RDP4 (namely Recombination Detection Program version 4). The highest acceptable *p* value for events of recombination was set at 0.01 with a window size of 20 nucleotides in order to minimize the false-positive error rate. Only the breakpoints, identified by more than four approaches were considered valid.

2.5.2. Phylogenetic Analyses

The amphibian *MHC class Ia* alleles in the NCBI database (https://www.ncbi.nlm.nih.gov/) were used to construct phylogenetic relationships. *MHC class Ib* genes from *Xenopus laevis* and *Xenopus tropicalis* and *MHC class Ia* genes from *Polypedates megacephalus*, *Lithobates catesbeianus*, *Espadarana prosoblepon*, *Ambystoma mexicanum*, *Xenopus ruwenzoriensis*, *Zhangixalus omeimontis*, *Rana huanrenensis*, *Agalychnis callidryas*, *Smilisca phaeota*, and *Xenopus laevis*, were included in the phylogenetic analysis. MEGA5 was utilized to construct a neighbor-joining (NJ) tree in order to determine the general phylogenetic state of the genes of *MHC class I*, which were newly isolated. Later, according to the results of recombination tests, the sequences were partitioned into individual regions to deduce more precisely the phylogenetic trees of anuran *MHC class I* genes, given that the identification of true phylogenetic relationship can be hindered by events of recombination. For inclusion in the tree construction, around three alleles from each species with *MHC class I* gene sequences from the NCBI database were chosen as representatives. Moreover, according to the Akaike Information Criterion (AIC) in MrModeltest, the most suitable model of nucleotide substitution was selected. Analyses revealed that GTR+I was the most appropriate model for all domains (α1, α2, and α3). MrBayes 3.2 (http://nbisweden.github.io/MrBayes/index.html) was used to reconstruct the Bayesian inference trees. Until the standard deviation of the split frequencies became <0.01, two independent runs of four Metropolis with simulations of MCMC (namely Markov Chain Monte Carlo) (three of them were 'heated' at the temperature of 0.1) were conducted for millions of generations. From the cold chain, the first quarter of the sampled trees were abandoned by default. The rest of the trees were used to calculate the posterior probability of every bipartition and the major rule consensus tree. In addition, the matrix of nucleotide distances was utilized via MEGA5 to construct the neighbor-joining trees of individual exons. A total of 1000 bootstrap replicates were used to estimate support for the nodes in the obtained tree.

2.5.3. Positive Selection Identification

False inferences of positive selection were ascribed to recombination. However, according to breakpoints of recombination which had been deduced, false positive can be decreased to an acceptable degree with the partitioning strategy in consideration of separate un-recombinant segments [34]. With the purpose of testing whether positive selection could shape the evolutionary pattern of the sequences of *MHC class I* isolated in this study, three methods were utilized for pre-analysis partitioning. Firstly, the Codeml subroutine in PAML version 4 (http://abacus.gene.ucl.ac.uk/software/paml.html) was used to test for the existence of positive selection signals. This procedure could be used to generate the maximum possibility forecasts of ω (ω = dN/dS, which means the synonymous/non-synonymous substitution ratios) amongst codons. Corresponding to the different distributing patterns of ω, six different methods could be used for verification in our statistics, including M8 (positive selection with the beta distribution approximating ω variation), M7 (almost neutral with the beta distribution approximating ω variation), M3 (discrete), M2a (positive selection), M1a (almost neutral), and M0 (one ω). In order to test the existence of positive selection in M2a, M3, and M8 models, three likelihood ratio tests (LRTs) were conducted for comparison of the nested models (M7 vs. M8, M1a vs. M2a, and M0 vs.M3). In models M2a and M8, Bayes empirical Bayes (BEB) method was used to identify positively selected sites. However, provided that the Codeml power would be influenced by the preciseness of the inferred phylogenetic trees, the two other approaches to detect the signals of selection were also utilized. Those two approaches, MEME, and FEL, were conducted at the website Data monkey (http://www.datamonkey.org). The combined results from all available studies were used to obtain a more precise understanding of the sites under positive selection.

3. Results

3.1. Identification and Sequencing of ZcMHC Class Ia cDNA

As shown in Figure 1, GenBank (https://www.ncbi.nlm.nih.gov/) accession number: KX021343, *ZcMHC class Ia* cDNA was obtained from the liver by RACE PCR and RT-PCR. A 1050 bp open reading frame (namely ORF), a 93 bp 5′-untranslated region, and a 165 bp -untranslated region were contained in the 1308 bp full-length *ZcMHC class Ia* cDNA. The 3′UTR contained a consensus polyadenylation signal sequence (AATAAA), which was 18 bp upstream from the poly A tail. The ORF encoded a 349-amino-acid precursor consisting of an α1 domain of 88 amino acids, a putative leader peptide of 18 amino acids, an α2 domain of 93 amino acids, and an α3 domain of 101 amino acids, followed by a 26 amino acid cytoplasmic region, and a 23 amino acid transmembrane region. According to the analyses of potential N-linked glycosylation sites, the α1 domain and cytoplasmic region contained potential N-linked glycosylation sites with amino acids NQTG and NNSD. Moreover, a potential O-linked glycosylation site was found at amino acid site 257.

```
-93                      CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGAAGTGTGAGGAGAAGCTGATGAGGTGACGCACAGGAGGAG
              10        20        30        40        50        60        70        80        90       100
1    ATGGAGCTCTGTGTGCTCACACTGATCTCTATGACTGTCATAGGAGCCCATTGTGATGGGCACTCTCTGCGGGTATTATTACACTGCAGTCTCGGACAAAGGATTG
1     M  E  L  C  V  L  T  L  I  S  M  T  V  I  G  A  H  C  D  G  H  S  L  R  Y  Y  Y  T  A  V  S  D  K  G  L
                         Leading peptide
             115       125       135       145       155       165       175       185       195       205
106  GATATACCAGAGTTCTCCATAGTTGGATATGTGGATGGACGAGAATTCATGAATTACAATAGTGAGTCTCATCAGGCTCGTCCTGTGACTCAATGGATGGAGAAG
36    D  I  P  E  F  S  I  V  G  Y  V  D  G  R  E  F  M  N  Y  N  S  E  S  H  Q  A  R  P  V  T  Q  W  M  E  K
                                                                                               α1 domain
             220       230       240       250       260       270       280       290       300       310
211  AATGAAGGACCCGAGTACTGGGAAGGACAGACACAGGTTTCCAAAGGCGCTGAACCTGTATTCAGACACAATGTCAGGACAGCGATGAGCCGCTTCAACCAGACT
71    N  E  G  P  E  Y  W  E  G  Q  T  Q  V  S  K  G  A  E  P  V  F  R  H  N  V  R  T  A  M  S  R  F  |N  Q  T
             325       335       345       355       365       375       385       395       405       415
316  GGAGGGATCCATACCTACCAGTGGATGTACGGCTGTGAGCTGAGAGATGACGGCAGCACTGCGGGGTATAATCAGGACGGATATGATGGGAGAGATTTTATGTCT
106   G| G  I  H  T  Y  Q  W  M  Y  G  C  E  L  R  D  D  G  S  T  A  G  Y  N  Q  D  G  Y  D  G  R  D  F  M  S
             430       440       450       460       470       480       490       500       510       520
421  CTGGACACACAGAGCGGGCTCTTTATCCCCACCATGAACGAGGCTCAGCTCACCACACAGAAATGGAACAGCCCGGAGATGAGAGACGGGGAGAGATACAAGAAT
141   L  D  T  Q  S  G  L  F  I  P  T  M  N  E  A  Q  L  T  T  Q  K  W  N  S  P  E  M  R  D  G  E  R  Y  K  N
                         α2 domain
             535       545       555       565       575       585       595       605       615       625
526  TATCTGCAGAATGAATGTATAGAATGGCTGAAGAAATTCATTGAGTACGGGAGAGAAGATCTGGAGAAGAGAGTGCGCCCAGAGGTGAAGGTGTGGGGCCGCCGT
176   Y  L  Q  N  E  C  I  E  W  L  K  K  L  I  E  Y  G  R  E  D  L  E  K  R  V  R  P  E  V  K  V  W  G  R  R
             640       650       660       670       680       690       700       710       720       730
631  CAGTCCGATGGTGCGACCAGACTTCAGTGCCTGGTGTACGGGTTTCACCCCCGAGCTGTGGATGTGAGGTGGATGAAGAATGGAGTAGACCATGCTCCTTCAGAT
211   Q  S  D  G  A  T  R  L  Q  C  L  V  Y  G  F  H  P  R  A  V  D  V  R  W  M  K  N  G  V  D  H  A  P  S  D
                         α3 domain
             745       755       765       775       785       795       805       815       825       835
736  GAAAGGACTCCGATTCTCCCCCATCCTGACGGCACCTATCAGATGAGAGTCGGTGTGGAAGTGCCAACAACCGAAGGAGACGCTTACTCCTGTCATGTGGAGCAC
246   E  R  T  P  I  L  P  H  P  D  G  T  Y  Q  M  R  V  G  V  E  V  P  T  T  E  G  D  A  Y  S  C  H  V  E  H
             850       860       870       880       890       900       910       920       930       940
841  AGCAGCCTGGGGGAAGAGACTCTCATTGTGACATGGGAACCTGACAATGGACTCTCAATGATTGTGATTATTGCCATTGTTTGTGCTGTTGTTGTCATCGTCGGC
281   S  S  L  G  E  E  T  L  I  V  T  W  E  P  D  N  G  L  S  M  I  V  I  I  A  I  V  C  A  V  V  V  I  V  G
                                                                                   trans-membrane region
             955       965       975       985       995      1005      1015      1025      1035      1045
946  GCTCTTGGCGGTGGAATCTTCCTCTACACAAAGAGGAGACGCGGCTACAAGACGGCCAGTACTTCAGATAACAATTCTGACTCCTCCAACAACCCGAAAGCCTGA
316   A  L  G  G  G  I  F  L  Y  T  K  R  R  R  G  Y  K  T  A  S  T  S  D  |N  N  S  D|  S  S  N  N  P  K  A  *
                                                                          cytoplasmic region
             1060      1070      1080      1090      1100      1110      1120      1130      1140      1150
1051 GCGTACACYATGCCGTGTTATGTGAGGACRCTCACAGGAGGAAGACGAGATATGACTTTTCTCYGTGCAGATTCTGCTCGCTTCTGTCTGCTTTTAAAAAYCTGAT
             1165      1175      1185      1195      1205
1156 GCAATAAATTTTAAAGAGAACCTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

Figure 1. Nucleotide and predicted amino acid sequences of ZcMHC class Ia. The regions are marked by underline and the region name. The stop codon is indicated by an asterisk. N-glycosylation sites are marked with squares. The polyadenylation signal (AATAAA) is underlined in the 3′ UTR.

3.2. Allele Characterization of ZcMHC Class Ia

We obtained 390 sequences using the ZcMHC-F3 and ZcMHC-R3 primers (Table 1) and identified 26 putative alleles of *MHC class I*. They contained the complete coding regions with lengths of 1047, 1050, or 1062 nucleotides. However, three of these sequences displayed stop codons at the 29 (2 sequences) and 32 (1 sequence) amino acid coding sites. As a result, 23 alleles without premature stop codons could be translated successfully to proteins. According to comparisons with HLA amino acid sequences (GenBank accession number: AAA76608.2), the four pivotal amino acid sites that form disulfide bridges (Cys203-Cys259 and Cys101-Cys164) were conserved in the 23 sequences (Figure 2).

Figure 2. Alignments of amino acid sequences translated from the 23 *MHC class I* alleles. Sites marked by black triangle represent the putative ABS deduced from structural information from human HLA molecules. The predicted positive selection sites were marked by the short vertical line. * refer to amino acids predicted to be under positive selection in the M8 model of PAML, with posterior probabilities of >95%. + marks positively selected sites identified by FEL and MEME.

This suggested that these isolated alleles may be functional. It was further shown in the phylogenetic analyses that the 23 alleles were placed at the *MHC class Ia* locus (Figure 3). Alleles were submitted with accession numbers (KX467504-KX467526) to GenBank.
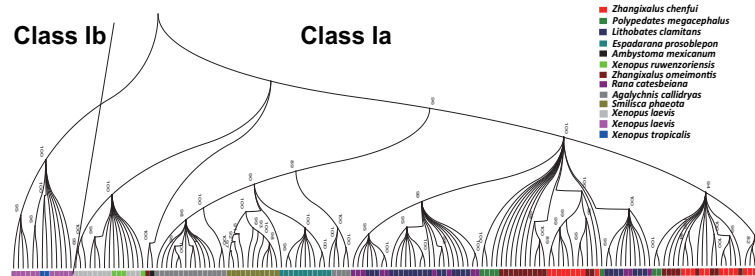
Figure 3. Phylogenetic relationships of anuran *MHC class Ia* and *Ib* genes. The neighbor-joining tree was constructed with MEGA5 using all available anuran *MHC class I* genes with an available length. Bootstrap values are indicated above the branches. Sample names are not shown due to the large size of the tree. Rather, the species to which these samples belong are denoted with different colors. The horizontal line separates *MHC class Ia* genes (to the right of the line) and *MHC class Ib* genes (to the left of the line).

3.3. Number of Loci and Genetic Diversity

In this study, we identified 2 alleles in 3 individuals, 3 alleles in 4 individuals, 4 alleles in 4 individuals, 5 alleles in 1 individual, and 6 alleles in 1 individual. At most, six alleles were observed in each Zc*MHC class Ia* gene. Since Chenfu's Treefrog is diploid [33], we hypothesized that it possesses 2 to 3 loci for *MHC class Ia*. Additionally, phylogenetic analysis showed that all alleles were clustered into three branches (Figure 3). Therefore, Chenfu's Treefrogs have at least three *MHC class Ia* loci.

Comparatively, high genetic diversity was observed in the sequences of *MHC class Ia* in Chenfu's Treefrogs (Table 2). In Chenfu's Treefrogs, the $\alpha 3$ domain had the lowest diversity whereas the $\alpha 1$ domain had the highest. Since the amino acid chain of the leading peptide, cytoplasmic region, and the trans-membrane region was shorter, there is a larger standard error in these regions. In addition, in comparison with the degree of divergence amongst amino acid chains, less divergence was observed amongst nucleotide sequences.

Table 2. Average nucleotide and amino acid distances among the 23 newly isolated alleles.

Amino Acid.	K2P Nucleotide Distance	Standard Error	Poisson-Corrected Amino Acid Distance	Standard Error
complete sequence	0.075	0.005	0.130	0.013
leading peptide	0.014	0.007	0.042	0.021
$\alpha 1$ domain	0.170	0.020	0.295	0.040
$\alpha 2$ domain	0.089	0.012	0.163	0.030
$\alpha 3$ domain	0.014	0.003	0.030	0.009
transmembrane region	0.062	0.020	0.144	0.059
cytoplasmic region	0.062	0.018	0.126	0.045

3.4. Recombination Detection

The recombination test was performed before the selection test, given that false inference of positive selection can be generated by recombination. The RDP4 program was used and three alleles were identified as products of recombination (Table 3). The actual number of recombinants is likely more than three as conservative parameters were used to avoid miscalculations. Nonetheless, these three alleles represented a marked proportion (13.04%) of the alleles. The RDP4 program was used to detect the five breakpoints (234, 314, 764, 993, and 1046).

Table 3. Recombination test using the RDP4 program.

Recombinant Sequence	Nucleotide Breakpoint	Potential Parental Sequences	Methods
*Poch-UA*05*	234,764	*Poch-UA*07/Poch-UA*18*	RDP, GENECONV, BootScan, MaxChi, Chimaera, Siscan, 3Seq
*Poch-UA*09*	314,1046	Unknown(*Poch-UA*14*)/*Poch-UA*11*	GENECONV, BootScan, MaxChi, Chimaera, Siscan, 3Seq
*Poch-UA*18*	234,993	Unknown(*Poch-UA*07*)/*Poch-UA*10*	BootScan, MaxChi, Chimaera, Siscan, 3Seq

Notes: Unknown was used to indicate an indeterminate source of reorganization.

3.5. Selection Detection

Since recombination mainly occurs at the boundary of different MHC domains, in order to avoid the influence of recombination on the selection signal, we divided *MHC class I* into six parts: leading peptide, α1 domain, α2 domain, α3 domain, transmembrane region, and cytoplasmic region. In order to screen for selection signals, three different codon-oriented maximal likelihood approaches were applied, FEL, PAML, and MEME, despite each approach identifying a different number of positively selected sites (Figure 2 and Table 4). In total, 21 sites of Chenfu's Treefrog *MHC class Ia* alleles were detected as under positive selection, eight of which were identified by more than one program. Of the 21 sites, 13 were assigned to the α1 and α2 domains. In these domains, excluding two sites on domain α1 and four sites on domain α2, the seven remaining sites all belonged to the presumed ABS [5] (Figure 2). We also tested for positively selected sites in the leading peptide, α3 domain, transmembrane region, and cytoplasmic region. We observed no positively selected sites in the leading peptide, two positively selected sites in the α3 domain, two positively selected sites in the transmembrane region, and three positively selected sites in the cytoplasmic region.

Table 4. Summary statistics for codon sites undergoing positive selection identified by different methods.

Method	α1 Domain				α2 Domain										α3 Domain		Transmembrane Region		Cytoplasmic Region		
	11	42	80	83	7	23	25	34	37	42	63	67	74	84	85	99	5	9	1	17	23
FEL [a]		+	+					+			+	+						+		+	
MEME [b]	+	+	+	+		+		+	+		+	+	+	+			+		+	+	
Codeml [c]		*			**	*	**	*			**	**	**		**	**			*		*

Notes: Numbers correspond to the alignment shown in Figure 2. [a] Amino acid sites identified by FEL are identified with +. [b] Amino acid sites identified by MEME are identified with +. [c] Amino acids were identified in model M8 by the Bayes empirical Bayes procedure using Codeml. Only sites that were predicted to have undergone positive selection with a posterior probability >95% are presented in the table. * indicates a posterior probability is >95%, and ** indicates posterior probability is >99%.

3.6. Phylogenetic Analyses

The 23 Chenfu's Treefrog alleles and previously published *MHC class I* genes with comparable coverage from amphibians (see Methods) were used to construct the phylogenetic tree and resolve the phylogenetic status of the new alleles. We found that genes of *MHC class Ia* and *MHC class Ib* clustered separately and that the cluster of *MHC class Ia* genes contained all Chenfu's Treefrog alleles (Figure 3).

Furthermore, to reconstruct a more accurate phylogenetic relationship, the aligned sequences were partitioned to individual domain regions, given that the exon exchanges took place between exons 3/4 and 2 and that mutation ratios were different between the α1, α2, and α3 domains. No specific topology of species was observed in the phylogenetic trees of the α1, α2, and α3 domains (Figure 4) when neighbor-joining (NJ) trees were constructed using synonymous mutations.
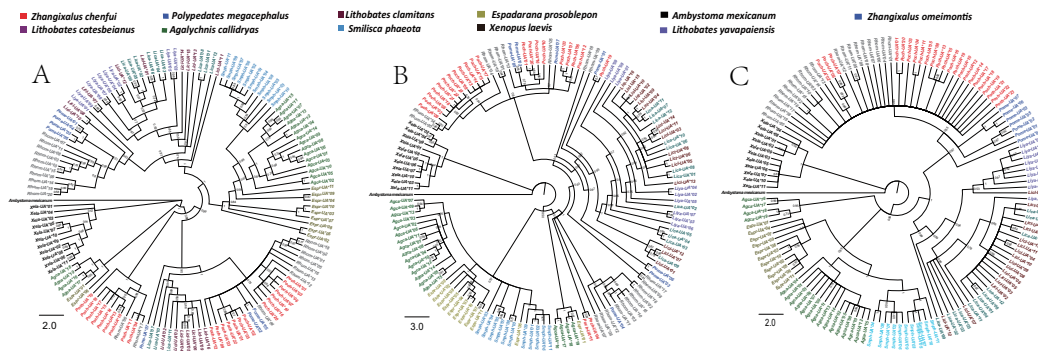
Figure 4. Phylogenetic relationships of α1 (**A**), α2 (**B**) and α3 (**C**) domains (analysis by MEGA). All of the branches are proportional to the scale shown at the bottom left of the figure. Alleles belonging to different species are marked using different colors.

Trans-species polymorphisms of anuran *MHC class Ia* genes were evident considering that *MHC class Ia* alleles shared common ancestry (Figure 3). Alleles from different taxa in the phylogenetic trees of the α1, α2, and α3 domains were no less similar to each other than the alleles from the same species. A part of the Rhacophoridae alleles in the tree of α1 domains branched from the common ancestors Centrolenidae, Hylidae, Ranidae, and Rhacophoridae species; the other portion was bifurcated into two clusters (Figures 4A and 5A). Though the tree showed a star-like topology because of the extensive non-synonymous mutations, a mixture of alleles from various species in the tree of the α2 domains could be observed as well (Figures 4B and 5B, Table 1). The tree constructed for the α3 domains conformed to the species tree as there were four clusters which were consistent with the anuran taxa: the cluster of Ranidae as the sister clade of Rhacophoridae was formed by alleles from *Lithobates clamitans*, *Lithobates yavapaiensis*, and *Lithobates catesbeianus*; the Rhacophoridae cluster with *Zhangixalus omeimontis* and *Polypedates megacephalus* was formed by the newly obtained alleles from Chenfu's Treefrog; a well-supported clade was formed by alleles from Xenopus; alleles from *Agalychnis callidryas* and *Smilisca phaeota* belonged to the Hylidae cluster and combined with alleles from *Espadarana prosoblepon* (Figures 4C and 5C). However, a portion of the *Agalychnis callidryas* alleles formed an independent branch. In conclusion, trans-species polymorphisms were observed in α1 and α2 domains of *MHC class I* genes, and species-specific clustering was found in α3 domain of *MHC class I* genes.
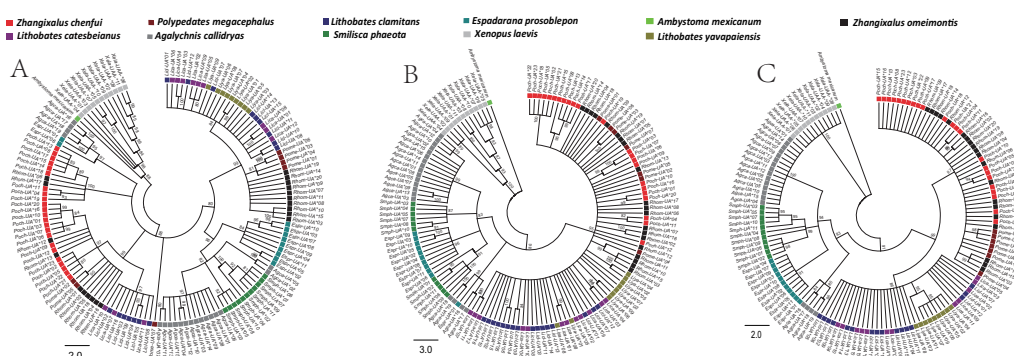
Figure 5. Phylogenetic relationships of α1 (**A**), α2 (**B**) and α3 (**C**) domains. Bayesian posterior probabilities are presented. All of the branches are proportional to the scale shown at the bottom left. Alleles belonging to different species are marked using different colored.

4. Discussion

Full-length cDNA of *MHC class I* genes was isolated successfully from Chenfu's Treefrogs. Afterward, we obtained 23 alleles from 13 individuals with full-length ORFs. Phylogenetic analyses suggested that all of the newly isolated alleles were contained within the *MHC class Ia* loci. Subsequent

analysis showed that: (1) alleles clustered into three branches in the phylogenetic analyses and displayed a comparatively high genetic diversity; (2) the putative amino acids of all the alleles included those forming disulfide bonds; (3) three alleles were the products of recombination; and (4) the alleles showed positive selection on many putative antigen binding sites (ABS).

As far as we know, previous research focusing on the MHC of anurans was mainly focused on α1, α2, and α3 domains [31,32]. Therefore, this study has provided a more comprehensive information regarding the alleles of a species in the order Anura (Chenfu's Treefrog). This enabled some interesting features to be discovered in the transmembrane and cytoplasmic regions. First of all, we found two positively selected sites in the transmembrane region and three positively selected sites in the cytoplasmic region. Second, we found breakpoints in the boundaries of both the transmembrane and cytoplasmic regions. Third, we located a potential N-linked glycosylation site in the cytoplasmic region of alleles. These findings suggest that further studies should focus on these areas. Furthermore, we observed N-glycosylation in the α1 region, which is important for assembling α-chains and transporting molecules of *MHC class I* through the Golgi apparatus and endoplasmic reticulum (ER) [35].

A great number of vertebrates have been observed to have high variation in MHC loci numbers, including fish [9,36], reptiles [37], birds [2,10,26], and mammals [15], which usually have more than one locus. For instance, three loci of *MHC class I* are expressed in Atlantic salmon [38], at least six loci are expressed in Blue tits [10], and 17 loci in a Cichlid species [39] but only two loci in Chinese sturgeon and paddlefish [9], as well as chickens [40]. According to prior research, a simple yet similar pattern for amphibians has shown that only one locus is usually expressed in Anura [30,41–44], whereas more than one locus is expressed in Caudata [27]. Researchers with less or insufficient information predicted that the diversification of amphibian loci of *MHC class I* took place sometime after the divergence of the Anura/Caudata but before that of the Pipidae/Ranidae [41] and that the number of loci of *MHC class I* was different amongst the anuran species [31,32].

We concluded that there are three *MHC class Ia* loci in Chenfu's Treefrog, given that six alleles were apparent in some Chenfu's Treefrog individuals included in this study and the knowledge that this species is diploid [33]. Due to our stringent criterion of allele identification, a large amount of sequence data (37 sequences) was discarded. Because alleles with low frequencies were removed, the actual number of loci of *MHC class Ia* in Chenfu's Treefrog may have been underestimated. It is notable that more than three loci of *MHC class Ia* were observed in Chenfu's frog. Prior research has shown that the Anura, including two Ranidae species (*Rana temporaria* and *Lithobates pipiens*) [41–43] and two Pipidae species (*Xenopus tropicalis and Xenopus laevis*) [30,44], expressed only one *MHC class Ia* locus, making three *MHC class Ia* loci quite uncommon in anuran species. However, according to recent studies on the evolution of the anuran MHC, non-model anuran species exhibited 2 or 3 for *MHC class Ia* loci [31,32]. In addition, a recent study showed that there are at least four loci in *Bufo gargarizans* and *Hyla japonica* [45]. Our study adds an additional anuran species with 3 *MHC class Ia* loci, which supports the ubiquity of multiple loci in the Auran MHC class Ia.

The variation in number of amphibian *MHC class I* loci appears to be more complex. After all available information on amphibian *MHC class I* gene loci was combined, it was found that species in the same family can have different MHC loci numbers (Figure 6) [46]. This supports that the variation in loci number in anurans is species-specific. During the evolution of amphibian *MHC class I* genes, which is the birth-and-death model of evolution, the occurrence of gene loss/duplication can be used to explain the pattern of mixed loci numbers [47]. However, inaccurate estimations of loci numbers in every species can also contribute to this pattern, as the two causes are compatible. Furthermore, previous studies have reported differences in number of MHC loci amongst populations or individuals of the same species [16,31,48]. Therefore, it is too early to draw conclusions regarding these variations until further studies are conducted to preclude the possibility that the number of loci was underestimated. We propose that future studies should be conducted to establish accurate numbers for anuran *MHC class Ia* gene loci in order to understand its history of duplication. In addition, availability of whole genome sequence would also help with elucidating MHC copy number.
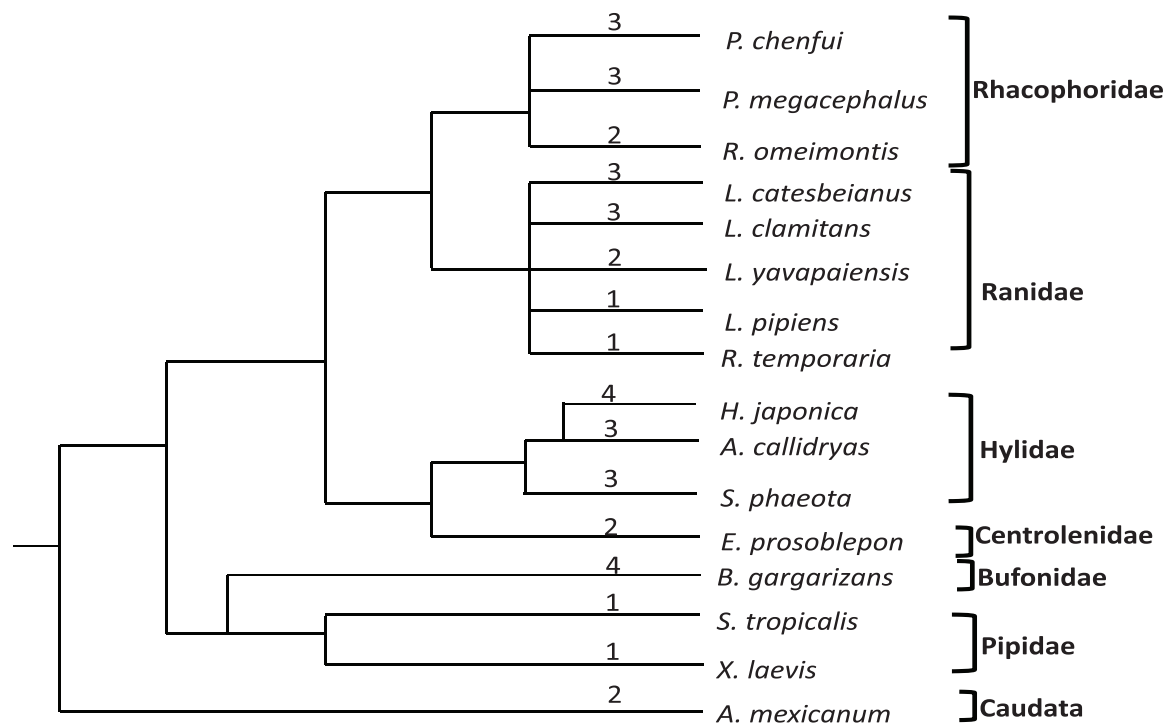
Figure 6. Numbers of putative *MHC class Ia* loci in amphibian species. The simplified phylogenetic tree of amphibian species with a deduced number of *MHC class Ia* loci based on the amphibian pedigree shown in Didinger et al., Zhao et al., and Roelants et al. The putative number of loci is shown above the corresponding branches.

Recombination is one of the most significant drivers of high diversity of *MHC class I* genes and generation of novel alleles [7–10]. Here, recombination was found to play a significant role in the evolution of *MHC class Ia* genes of Chenfu's Treefrog. Recombination generated 13.04% of the alleles. Furthermore, a large portion of the recombination breakpoints was found near the regional boundary, and the other portions were located near the two insertion mutation sites. It was found that the DNA was exchanged primarily between domains α1, α2, and α3. The recombination pattern is similar to that observed in the genes of *MHC class I* in bony fish and other anurans [8,9,28,49]. By looking at the pattern of exon shuffling, *MHC class I* genes of teleost and anuran species may share similar organizations. One previous study reported that intron 2 of *MHC class I* gene in bony fish contained numerous repeated elements and had long coverage [50]. These two gene organization forces played an important role in making intron 2 sequence recombination hot spots. Nonetheless, the intron 2 sequences of Rhacophoridae genes of *MHC class I* was not available for study. According to the genetic information from the entire genome of *Xenopus tropicalis*, it appears that Rhacophoridae species tend to have a long intron 2 regions of the *MHC class I* genes [28]. However, recombination analysis revealed breakpoints located near the two insertion mutation sites, which suggests that there are more complicated mechanisms and patterns involved in the recombination of Chenfu's Treefrog genes of *MHC class I*. This indicates that the recombinational pattern in Chenfu's Treefrog is similar to the mammals which show reticulate recombination patterns [51]. To gain more systematic and comprehensive information regarding recombination in amphibian *MHC class I* genes, a larger number of amphibian species should be evaluated.

In this study, using various approaches, it has been shown that Chenfu's Treefrog *MHC class I* genes have experienced positive selection. With positive selection occurring in the 23 alleles of *MHC class I*, positive selection signals and sites were detected by multiple codon-oriented approaches. The sites, identified by more than one approach, were regarded robust and made up the subset of

putative ABS from the human HLA crystal structure [5]. This result indicated that, despite some differences, the Chenfu's Treefrog *MHC class I* molecules and human HLA may share some similarity in three-dimensional structure. Similar results have been observed in many other species [9,31,52]. It can be concluded that the genes of *MHC class I* genes tend to undergo pathogen-mediated positive selection due to their ABS function.

Given that there is common ancestry in the alleles of *MHC class Ia* from Rhacophoridae, Ranidae, Centrolenidae, and Hylidae (Figure 1), trans-species polymorphisms of anuran genes of *MHC class Ia* were evident. Specifically, the α1 and α2 domains of anuran *MHC class I* genes are where the trans-species polymorphisms were observed (Figure 5). The species-specific cluster was not observed in the phylogenetic trees of α1 and α2 (Figure 5) when NJ trees were constructed using synonymous mutations. Later, positive selection was identified to match the pattern of trans-species polymorphism. To be more specific, α1 and α2 containing major regions for antigen binding and recognition will experience a strong positive selection force. By encoding the domains, nucleotide sequences can be retained for long periods of time and the origin of alleles with sufficient importance predates the divergence of species. On the contrary, α3 functions as a link and an anchor and experiences relatively little selection pressure. Consequently, a precise phylogenetic relationship of these species can be established by examining the commonality in the α3 domain. Previous studies have also reported differences in patterns of evolution for different exons of the same gene [9,32,53].

The mixture of alleles present in various species as trans-species polymorphisms speaks to the long history of those genes in the MHC, predating the emergence of those species [54]. Therefore, the time of MHC gene formation can be inferred from the divergence time of related species. Before the divergences of Centrolenidae/Hylidae and Rhacophoridae/Ranidae (Figure 5A), based on this research, the allelic lineages of *MHC class Ia* found in Chenfu's Treefrog already existed. The oldest allelic lineages of *MHC class Ia* in Anura have been predicted to have originated around 70 MYA (million years ago) as the two divergence events date back to 68–91 and 69–72 MYA [55]. The time that the alleles of *MHC class I* genes have been persisted in Anura is longer than that of certain other vertebrates [56].

5. Conclusions

In summary, *MHC class Ia* genes were isolated successfully from *Chenfu's Treefrog*, which were found to contain three loci for *MHC class Ia* genes. Because of exchanges of entire exons and positive selection acting predominantly on the ABS, frequent gene duplication and recombination played significant roles in the evolution of the genes of Rhacophoridae *MHC class I*. Our knowledge of the evolution of the anuran *MHC* genes has been expanded upon by the inclusion of Rhacophoridae *MHC class I* genes.

References

1. Klein, J. *Natural History of the Major Histocompatibility Complex*; Wiley: New York, NY, USA, 1986.
2. Leclaire, S.; Strandh, M.; Dell'Ariccia, G.; Gabirot, M.; Westerdahl, H.; Bonadonna, F. Plumage microbiota covaries with the major histocompatibility complex in blue petrels. *Mol. Ecol.* **2019**, *28*, 833–846. [CrossRef] [PubMed]
3. Wieczorek, M.; Abualrous, E.T.; Sticht, J.; Álvaro-Benito, M.; Stolzenberg, S.; Noé, F.; Freund, C. Major histocompatibility complex (MHC) class I and MHC class II proteins: Conformational plasticity in antigen presentation. *Front. Immunol.* **2017**, *8*, 292. [CrossRef] [PubMed]
4. Hughes, A.M. Natural selection at major histocompatibility complex loci of vertebrates. *Annu. Rev. Genet.* **1998**, *32*, 415–435. [CrossRef] [PubMed]
5. Bjorkman, P.J.; Saper, M.; Samraoui, B.; Bennett, W.S.; Strominger, J.L.; Wiley, D. Structure of the human class I histocompatibility antigen, HLA-A 2. *Nature* **1987**, *329*, 506–512. [CrossRef] [PubMed]
6. Piertney, S.; Oliver, M. The evolutionary ecology of the major histocompatibility complex. *Heredity* **2006**, *96*, 7–21. [CrossRef] [PubMed]
7. Jakobsen, I.B.; Wilson, S.R.; Easteal, S. Patterns of reticulate evolution for the classical class I and II HLA loci. *Immunogenetics* **1998**, *48*, 312–323. [CrossRef]
8. Consuegra, S.; Megens, H.J.; Schaschl, H.; Leon, K.; Stet, R.J.M.; Jordan, W.C. Rapid Evolution of the MH Class I Locus Results in Different Allelic Compositions in Recently Diverged Populations of Atlantic Salmon. *Mol. Biol. Evol.* **2005**, *22*, 1095. [CrossRef]
9. Wang, D.; Zhong, L.; Wei, Q.; Gan, X.; He, S. Evolution of *MHC class I* genes in two ancient fish, paddlefish (*Polyodon spathula*) and Chinese sturgeon (*Acipenser sinensis*). *FEBS Lett.* **2010**, *584*, 3331–3339. [CrossRef]
10. Wutzler, R.; Foerster, K.; Kempenaers, B. *MHC class I* variation in a natural blue tit population (*Cyanistes caeruleus*). *Genetica* **2012**, *140*, 349–364. [CrossRef]
11. Spurgin, L.G.; Richardson, D.S. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proc. R. Soc. Lond. Ser. B* **2010**, *277*, 979–988. [CrossRef]
12. Garrigan, D.; Hedrick, P.W. Perspective: Detecting adaptive molecular polymorphism: Lessons from the MHC. *Evolution* **2003**, *57*, 1707. [CrossRef] [PubMed]
13. Mccairns, R.J.; Bourget, S.; Bernatchez, L. Putative causes and consequences of MHC variation within and between locally adapted stickleback demes. *Mol. Ecol.* **2011**, *20*, 486–502. [CrossRef] [PubMed]
14. Aguilar, A.; Roemer, G.; Debenham, S.; Binns, M.; Garcelon, D.; Wayne, R.K. High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 3490–3494. [CrossRef] [PubMed]
15. Vincek, V.; Nizetić, D.; Golubić, M.; Figueroa, F.; Nevo, E.; Klein, J. Evolutionary expansion of *MHC class I* loci in the mole-rat, Spalax ehrenbergi. *Mol. Biol. Evol.* **1987**, *4*, 483–491.
16. Figueroa, F.; Mayer, W.E.; Sato, A.; Zaleska-Rutczynska, Z.; Hess, B.; Tichy, H.; Klein, J. *MHC class I* genes of swordtail fishes, Xiphophorus: Variation in the number of loci and existence of ancient gene families. *Immunogenetics* **2001**, *53*, 695–708. [CrossRef]
17. Gaudieri, S.; Dawkins, R.L.; Habara, K.; Kulski, J.K.; Gojobori, T. SNP profile within the human major histocompatibility complex reveals an extreme and interrupted level of nucleotide diversity. *Genome Res.* **2000**, *10*, 1579–1586. [CrossRef]
18. Meyer, D.; Thomson, G. How selection shapes variation of the human major histocompatibility complex: A review. *Ann. Hum. Genet.* **2001**, *65*, 1–26. [CrossRef]
19. Pavelko, K.D.; Mendez-Fernandez, Y.; Bell, M.P.; Hansen, M.J.; Johnson, A.J.; David, C.S.; Rodriguez, M.; Pease, L.R. Nonequivalence of classical *MHC class I* loci in ability to direct effective antiviral immunity. *PLoS Pathog.* **2012**, *8*, e1002541. [CrossRef]
20. Kelley, J.; Walter, L.; Trowsdale, J. Comparative genomics of major histocompatibility complexes. *Immunogenetics* **2005**, *56*, 683–695. [CrossRef]
21. Nonaka, M.; Namikawa, C.; Kato, Y.; Sasaki, M.; Salter-Cid, L.; Flajnik, M.F. Major histocompatibility complex gene mapping in the amphibian Xenopus implies a primordial organization. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 5789–5791. [CrossRef]
22. Rodriguez, A.; Garcia-Montijano, M. MHC diversity and differential exposure to pathogens in kestrels (Aves: Falconidae). *Mol. Ecol.* **2010**, *19*, 691–705.

23. May, S.; Zeisset, I.; Beebee, T.J. Larval fitness and immunogenetic diversity in chytrid-infected and uninfected natterjack toad (*Bufo calamita*) populations. *Conserv. Genet.* **2011**, *12*, 805–811. [CrossRef]

24. Sommer, S. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Front. Zool.* **2005**, *2*, 16. [CrossRef] [PubMed]

25. Bernatchez, L.; Landry, C. MHC studies in nonmodel vertebrates: What have we learned about natural selection in 15 years? *J. Evol. Biol.* **2003**, *16*, 363–377. [CrossRef] [PubMed]

26. Roved, J.; Hansson, B.; Tarka, M.; Hasselquist, D.; Westerdahl, H. Evidence for sexual conflict over major histocompatibility complex diversity in a wild songbird. *Proc. R. Soc. Lond. Ser. B* **2018**, *285*, 20180841. [CrossRef] [PubMed]

27. Sammut, B.; Du Pasquier, L.; Ducoroy, P.; Laurens, V.; Marcuz, A.; Tournefier, A. Axolotl MHC architecture and polymorphism. *Eur. J. Immunol* **1999**, *29*, 2897–2907. [CrossRef]

28. Bos, D.H.; Waldman, B. Evolution by recombination and transspecies polymorphism in the *MHC class I* gene of Xenopus laevis. *Mol. Biol. Evol.* **2006**, *23*, 137–143. [CrossRef]

29. Bos, D.; Waldman, B. Polymorphism, natural selection, and structural modeling of class Ia MHC in the African clawed frog (*Xenopus laevis*). *Immunogenetics* **2006**, *58*, 433–442. [CrossRef]

30. Ohta, Y.; Goetz, W.; Hossain, M.Z.; Nonaka, M.; Flajnik, M.F. Ancestral organization of the MHC revealed in the amphibian Xenopus. *J. Immunol.* **2006**, *176*, 3674–3685. [CrossRef]

31. Zhao, M.; Wang, Y.; Shen, H.; Li, C.; Chen, C.; Luo, Z.; Wu, H. Evolution by selection, recombination, and gene duplication in *MHC class I* genes of two Rhacophoridae species. *BMC Evol. Biol.* **2013**, *13*, 113. [CrossRef]

32. Kiemnec-Tyburczy, K.; Richmond, J.; Savage, A.; Lips, K.; Zamudio, K. Genetic diversity of *MHC class I* loci in six non-model frogs is shaped by positive selection and gene duplication. *Heredity* **2012**, *109*, 146. [CrossRef] [PubMed]

33. Dejun, X.L.W.G.L. Studies on the Karyotypes of Rhacoohorus Chenfui and Rana Japonica Japonica. *Acta Acad. Med. Zunyi* **1990**, *2*, 008.

34. Scheffler, K.; Martin, D.P.; Seoighe, C. Robust inference of positive selection from recombining coding sequences. *Bioinformatics* **2006**, *22*, 2493–2499. [CrossRef] [PubMed]

35. Lankat-Buttgereit, B.; Tampë, R. The transporter associated with antigen processing TAP: Structure and function. *FEBS Lett.* **1999**, *464*, 108–112. [CrossRef]

36. Yamaguchi, T.; Dijkstra, J.M. Major Histocompatibility Complex (MHC) Genes and Disease Resistance in Fish. *Cells* **2019**, *8*, 378. [CrossRef]

37. Miller, H.C.; Belov, K.; Daugherty, C.H. *MHC class I* genes in the tuatara (Sphenodon spp.): Evolution of the MHC in an ancient reptilian order. *Mol. Biol. Evol.* **2006**, *23*, 949–956. [CrossRef]

38. Lukacs, M.F.; Harstad, H.; Bakke, H.G.; Beetz-Sargent, M.; McKinnel, L.; Lubieniecki, K.P.; Koop, B.F.; Grimholt, U. Comprehensive analysis of *MHC class I* genes from the U-, S-, and Z-lineages in Atlantic salmon. *BMC Genom.* **2010**, *11*, 154. [CrossRef]

39. Sato, A.; Klein, D.; Sültmann, H.; Figueroa, F.; O'hUigin, C.; Klein, J. Class I Mhc genes of cichlid fishes: Identification, expression, and polymorphism. *Immunogenetics* **1997**, *46*, 63–72. [CrossRef]

40. Kaufman, J.; Milne, S.; Göbel, T.W.; Walker, B.A.; Jacob, J.P.; Auffray, C.; Zoorob, R.; Beck, S. The chicken B locus is a minimal essential major histocompatibility complex. *Nature* **1999**, *401*, 923–925. [CrossRef]

41. Teacher, A.G.F.; Garner, T.W.J.; Nichols, R.A. Evidence for Directional Selection at a Novel Major Histocompatibility Class I Marker in Wild Common Frogs (*Rana temporaria*) Exposed to a Viral Pathogen (Ranavirus). *PLoS ONE* **2009**, *4*, e4616. [CrossRef]

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

43. Flajnik, M.; Ohta, Y.; Namikawa-Yomada, C.; Nonaka, M. Insight into the primordial MHC from studies in ectothermic vertebrates. *Immunol. Rev.* **1999**, *167*, 59–67. [CrossRef] [PubMed]

44. Goyos, A.; Sowa, J.; Ohta, Y.; Robert, J. Remarkable conservation of distinct nonclassical *MHC class I* lineages in divergent amphibian species. *J. Immunol.* **2011**, *186*, 372–381. [CrossRef] [PubMed]

45. Didinger, C.; Eimes, J.A.; Lillie, M.; Waldman, B. Multiple major histocompatibility complex class I genes in Asian anurans: Ontogeny and phylogeny. *Dev. Comp. Immunol.* **2017**, *70*, 69–79. [CrossRef] [PubMed]

46. Roelants, K.; Gower, D.J.; Wilkinson, M.; Loader, S.P.; Biju, S.; Guillaume, K.; Moriau, L.; Bossuyt, F. Global patterns of diversification in the history of modern amphibians. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 887–892. [CrossRef]

47. Nei, M.; Gu, X.; Sitnikova, T. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 7799–7806. [CrossRef]

48. Xu, S.; Ren, W.; Zhou, X.; Zhou, K.; Yang, G. Sequence polymorphism and geographical variation at a positively selected MHC-DRB gene in the finless porpoise (Neophocaena phocaenoides): Implication for recent differentiation of the Yangtze Finless porpoise? *J. Mol. Evol.* **2010**, *71*, 6–22. [CrossRef]

49. Lau, Q.; Igawa, T.; Komaki, S.; Satta, Y. Characterisation of major histocompatibility complex class I genes in Japanese Ranidae frogs. *Immunogenetics* **2016**, *68*, 797–806. [CrossRef]

50. Shum, B.P.; Guethlein, L.; Flodin, L.R.; Adkison, M.A.; Hedrick, R.P.; Nehring, R.B.; Stet, R.J.; Secombes, C.; Parham, P. Modes of salmonid *MHC class I* and II evolution differ from the primate paradigm. *J. Immunol.* **2001**, *166*, 3297–3308. [CrossRef]

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

52. Adams, E.J.; Cooper, S.; Thomson, G.; Parham, P.; Adams, E. Common chimpanzees have greater diversity than humans at two of the three highly polymorphic *MHC class I* genes. *Immunogenetics* **2000**, *51*, 410–424. [CrossRef] [PubMed]

53. Nonaka, M.I.; Aizawa, K.; Mitani, H.; Bannai, H.P.; Nonaka, M. Retained orthologous relationships of the *MHC class I* genes during euteleost evolution. *Mol. Biol. Evol.* **2011**, *28*, 3099–3112. [CrossRef] [PubMed]

54. Klein, J.; Sato, A.; Nagl, S.; O'hUigín, C. Molecular trans-species polymorphism. *Annu. Rev. Ecol. Syst.* **1998**, *29*, 1–21. [CrossRef]

55. Roelants, K.; Haas, A.; Bossuyt, F. Anuran radiations and the evolution of tadpole morphospace. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 8731–8736. [CrossRef] [PubMed]

56. Piontkivska, H.; Nei, M. Birth-and-death evolution in primate *MHC class I* genes: Divergence time estimates. *Mol. Biol. Evol.* **2003**, *20*, 601–609. [CrossRef] [PubMed]