

Evolution by Gene Duplication, Recombination and Selection in MHC Class I Genes of *Odorrana margaretae*

(Evolusi melalui Penduaan Gen, Penggabungan Semula dan Gen Pemilihan Kelas I MHC *Odorrana margaretae*)

HU CHEN, XIAORONG TAN, FUYAO HAN, YONGFANG YAO, HUAILIANG XU & MINGWANG ZHANG*

ABSTRACT

*Amphibians have undergone catastrophic declines worldwide and the major histocompatibility complex (MHC) is an attractive candidate for investigating the link between adaptive variation and individual fitness. Nonetheless, little research has been conducted to study the amphibians MHC genes. Thus, MHC class Ia of *Odorrana margaretae*, a representative species of the *Odorrana*, was isolated and used to analyze the adaptive mechanism. We isolated alleles and determined evolution patterns of MHC class Ia in *O. margaretae*. The phylogenetic tree of MHC class Ia was reconstructed to understand the evolutionary relationship of Anura. In *O. margaretae* of MHC class Ia: there were more non-synonymous mutations compared to synonymous mutations; there are at least two loci; the recombination alleles took place for a large proportion (55.56%) of the alleles and recombination usually took place between whole exon of Ia genes of MHC class; positive selection sites were detected and most (10 of 19) located at the $\alpha 1$ and $\alpha 2$ domains belonged to the presumed ABS. The diversity of genes of MHC class Ia was led by recombination, gene duplication and positive selection. Trans-species polymorphisms of anuran genes of MHC class Ia were evident in anurans. Therefore, the MHC class Ia present demonstrative candidates for investigating the link between adaptive variation and individual fitness. We argue that knowledge of the MHC of other anuran lineage, especially focusing a genus such as *O. margaretae*, should provide an overall more complete picture of the organization of MHC in anurans.*

Keywords: Adaptive variation; anurans; evolution; individual fitness; major histocompatibility complex; polymorphism

ABSTRAK

*Amfibia telah menjalani bencana penurunan di seluruh dunia dan kompleks kehistoserasian utama (MHC) adalah calon yang menarik untuk mengkaji kaitan antara variasi mudah suai dan kecergasan individu. Walau bagaimanapun, terdapat kurang kajian yang dijalankan untuk mengkaji gen MHC amfibia. Oleh itu, MHC kelas Ia *Odorrana margaretae*, satu spesies yang mewakili *Odorrana*, dipencil dan digunakan untuk menganalisis mekanisme mudah suai. Kami memencil alel dan menentukan evolusi corak MHC kelas Ia dalam *O. margaretae*. Pokok filogenetik MHC kelas Ia telah dibina semula untuk memahami hubungan evolusi Anura. Dalam *O. margaretae* MHC kelas Ia: terdapat lebih banyak mutasi sinonim berbanding sinonim mutasi; terdapat sekurang-kurangnya dua loci; penggabungan semula alel berlaku bagi sebahagian besar (55.56%) daripada alel ini dan penggabungan semula ini biasanya berlaku antara keseluruhan ekson gen MHC kelas Ia; tapak pilihan positif dikesan dan kebanyakannya (10 daripada 19) terletak di dalam domain $\alpha 1$ dan $\alpha 2$ dipercayai kepunyaan ABS. Kepelbagaian gen kelas MHC Ia dipimpin oleh penggabungan semula, penduaan gen dan pemilihan positif. Polimorfisme trans-spesies gen anura MHC kelas Ia adalah jelas dalam anura. Oleh itu, MHC kelas Ia menunjukkan calon untuk mengkaji kaitan antara perubahan mudah suai dan kecergasan individu. Kami berpendapat bahawa pengetahuan daripada MHC keturunan anura lain, terutamanya menumpukan kepada genus seperti *O. margaretae* sepatutnya menyediakan satu gambaran lebih lengkap keseluruhan susunan MHC dalam anura.*

Kata kunci: Anura; evolusi; kecergasan individu; kompleks utama kehistoserasian; perubahan mudah suai; polimorfisme

INTRODUCTION

Proteins encoded by genes of the major histocompatibility complex (MHC) have an essential role in the adaptive immune response of vertebrates. These proteins are receptor molecules which bind small fragments of peptides derived from the processing of pathogens and present them to immune cells (Cardenas et al. 2005; Klein 1986). Two major groups of these proteins, class I and class II, are encoded by two subgroups of MHC genes which are

primarily responsible for the recognition of intracellular and extracellular pathogens, respectively (Cardenas et al. 2005). The MHC class I molecule is a heterodimer that consists of a heavy chain (α chain) and a microglobulin chain ($\beta 2m$ chain). The heavy chain encoded by MHC class I genes has a cytoplasmic region, a trans-membrane region and three extracellular domains designated $\alpha 1$, $\alpha 2$ and $\alpha 3$. The antigen binding sites (ABS), which are located in domains $\alpha 1$ and $\alpha 2$ and primarily encoded by the second

and third exon of MHC class I genes, are key regions responsible for the recognition of antigens (Hughes 1998; Klein 1986).

MHC polymorphisms have been commonly found in MHC genes, especially in ABS (Bernatchez & Landry 2003; Piertney & Oliver 2006). On the other hand, the ability of MHC genes to face various pathogens is believed to be mainly related to sequence variation among MHC alleles in the ABS (Potts & Wakeland 1990; Van Eden et al. 1983). In addition, the association between a high level of MHC polymorphisms and disease-resistant ability was confirmed by a great deal of evidence (Radwan et al. 2010; Spurgin & Richardson 2010; Trowsdale 2011; Wegner et al. 2003) and polymorphisms in the MHC can be used as an indicator of survival potential of a species (Edwards & Hedrick 1998; Sommer 2005). Therefore, understanding polymorphism and polymorphism maintenance mechanisms of MHC in a species or population is essential for understanding the process of adaptive evolution.

As mentioned before, the maintenance mechanisms of MHC polymorphism are very important for adaptive immunity of vertebrate species. Three major mechanisms have been proposed to explain the high polymorphisms of MHC genes, including gene duplication (Aguilar et al. 2004; Figueroa et al. 2001; Vincek et al. 1987), gene recombination (Consuegra et al. 2005; Jakobsen et al. 1998; Wang et al. 2010; Wutzler et al. 2012) and balancing selection (Garrigan et al. 2003; MCCAIRNS et al. 2011; Spurgin & Richardson 2010). Gene duplication provides additional genetic material for variations to accumulate. However, not all duplicate genes persist in the genome. In fact, some duplicate genes might be deleted or become dysfunctional due to deleterious mutations (Nei & Rooney 2005; Nei et al. 1997). Gene recombination within or between MHC loci also increase the sequence variation (Andersson & Mikko 1995; Begovich et al. 1992; Jakobsen et al. 1998; She et al. 1991; Zhao et al. 2013). Finally, balancing selection allows large numbers of alleles to remain in the genome for extended periods of time (Bernatchez & Landry 2003). Regardless of the mechanisms underlying MHC variation, the MHC polymorphisms are necessary for species or populations to fight disease (Bernatchez & Landry 2003; Edwards & Hedrick 1998; Jeffery & Bangham 2000; Spurgin & Richardson 2010).

One of the reasons amphibians have undergone catastrophic declines worldwide is due to disease (Carey et al. 1999; Pounds et al. 2006; Stuart et al. 2004). For example Ranavirus, which is found throughout China, has been responsible for amphibian declines (Daszak et al. 1999; Xu et al. 2010). Thus, as an indicator of disease resistance, MHC of amphibians is considered a rich research area. Nonetheless, little research has been conducted to study amphibian MHC genes. For example, there are only a few reports focusing on the Urodela and Anura MHC class I genes (Bos & Waldman 2006a, 2006b; Teacher et al. 2009). These studies showed that the evolution of

MHC class I genes were ascribed to the recombination and balancing selection maintenance mechanisms (Bos & Waldman 2006a, 2006b). Moreover, amphibian MHC class I genes commonly contain polymorphisms of trans-species (Bos & Waldman 2006a). It is not clear whether all amphibians fall under the rules shown by the evolution of a certain amount of species. Recent studies examining eight anuran species showed that along with the evolution of MHC class I genes, there were more complicated mechanisms and patterns (Kiemnec-Tyburczy et al. 2012; Zhao et al. 2013). Consequently, in order to gain a more systematic comprehension of the MHC genes evolution, more amphibian species should be evaluated.

As far as we know, studies on *Odorrana* MHCs specifically have not been published in literature, therefore, a study exploring *Odorrana* MHCs could prove to be very significant. For this study, *Odorrana margaretae*, a representative species for *Odorrana*, were obtained from Wawushan Mountain for research and analysis. To that end, a pair of primers was used to isolate MHC class I from the cDNA library of *O. margaretae*. Next, we analyzed the number, polymorphisms, phylogenetic status and polymorphism maintenance mechanisms of *O. margaretae* MHC class I alleles. We initially performed a preliminary investigation of the mechanisms underlying the formation and maintenance of a representative species of *Odorrana* MHC class I gene Polymorphisms. Thus, it is the intent of this study to fill the gap regarding *Odorrana* MHC research. The information presented here also offers insights to the evolution of anuran MHC class I genes.

MATERIALS AND METHODS

SAMPLING, RNA ISOLATION AND CDNA SYNTHESIS

O. margaretae specimens were collected from the Wawushan Natural Reserve in Sichuan, China. Taking into account the protection of wild resources such as *O. margaretae*, we have selected three representative sites in the Wawushan Natural Reserve and captured three, three, and four frogs, from each site, respectively. The frogs were then killed utilizing the double destroyed pulp method. The livers were quickly removed and frozen in liquid nitrogen for 20 min, then stored at -80°C until the RNA extraction phase. Trizol® reagent was used to isolate the total RNA from each sample according to the manufacturer's protocol (Takara, Japan). To determine the quality of RNA samples, wave lengths of 280 and 260 nm, the optical density absorption ratio (Bio-Rad) were used. Specimens with the ratio of absorption (OD260/OD280) from 1.7 to 2.2 were applied in the cDNA synthesis. With the use of reverse-transcribed (RT) Reagent Kit of PrimeScript™ (Takara, Dalian, China), 2 µg RNA from each sample was RT to cDNA with oligo(dT)18 primers according to the manufacturer's protocol. All procedures related to live frogs experiments were approved by Sichuan Agricultural University's Animal Ethics Committee.

DNA AMPLIFICATION AND CLONING

We selected the technique of touchdown PCR to amplify DNA, given its ability to improve the specificity and efficiency of the PCR reaction. The touchdown PCR was conducted under the following conditions: 95°C for 5 min; 26 cycles of 94°C for 30 s, 65°C (-0.5°C/cycle) for 30 s, and 72°C for 1 min; 10 cycles of 94°C for 30 s, 52°C for 30 s and 72°C for 1 min; and a final extension at 72°C for 7 min. Prior to the formal amplification of DNA, a pre-experiment was performed to determine whether the primers (forward primer 5'-CTGCGSWAYTATKABACWGCAGTCTC-3'; reverse primer 5'-TYCAGRCTGCTGTG STCCACAT-3') were able to work in all the samples. PCR was performed in a 50 µL volume solution containing 25 µL commercial TaqPremix (Takara, Dalian, China), 15 µL ddH₂O (TIANGEN, Beijing, China), 2.5 µL each primer (5 µM/µL) and 5 µL cDNA (20-50 ng/µL). The Universal DNA Purification kit (TIANGEN, Beijing, China) was used to excise and purify the PCR products from each sample.

Later, the products were connected to the pEASY-T5 Zero vector (TransGen Biotech, Beijing, China) and imported into Trans1-T1 competent cells (TransGen Biotech, Beijing, China). After a 1.5 h recovery process, the competent cells with PCR products were spread on the agar of Luria-Bertani (LB) with ampicillin and cultivated at 37°C for a night, according to manufacturer's instructions. The pEASY-T5 Zero vector contains a suicide gene and only recombinants can survive in ampicillin-containing LB plates. Furthermore, with the purpose of screening for the existence of the fragments of inserted DNA by PCR and agarose gel electrophoresis, positive clones were chosen randomly. With positive transformants, around 12 colonies were chosen for each specimen and every transformant was separately fostered for 8 h at 37°C in ampicillin-containing LB. The inserts DNA were sequenced at Sangon Biotech Co., Ltd (Shanghai, China).

SEQUENCE ALIGNMENT AND ALLELE IDENTIFICATION

Sequencing results of MHC genes were proofread, edited, aligned and assembled into contigs using DNASTAR 7.1 (Madison, USA). Two ABI trace files were used with a forward and reverse sequencing primer to sequence from both ends for each clone and assembled using the SeqMan Pro program in the DNASTAR 7.1 package. This step can reduce the false positive of nucleotide variation due to the sequencing process. ClustalW in MEGA5 was used to align and correct the assembly sequence. These sequences were then blasted onto the published anuran MHC class I genes in National Center for Biotechnology Information (NCBI).

In order to avoid false positives, strict selection rules for alleles were performed in the data analysis. At the first screening level, only the same sequence that appeared five or more times in our sequenced database was considered an allele. The second level of filtering was performed on the remaining data. The sequence was defined as an allele that presented in two or more individuals and more than

four times in all. After the two screenings, the remaining sequences were considered as false positives and therefore did not participate in subsequent data analysis. The newly gained putative alleles were translated into sequences of amino acids and aligned with chains of human leukocyte antigen (HLA). According to comparisons between HLA amino acid sequences of human-beings and newly gained peptides (GenBank accession number: AAA76608.2), only those sequences whose disulfide bridge were consistent with HLA (Cys203-Cys259 and Cys101-Cys164 in HLA) were considered as alleles.

DATA ANALYSIS

MEGA5 was employed to calculate the average pairwise nucleotide distances (Kimura-2-parameter model, K2P) and the Poisson-corrected amino acid distances among alleles were computed over all sequence pairs, Exon 2, Exon 3 and Exon 4 of nucleotide and $\alpha 1$, $\alpha 2$ and $\alpha 3$ domains of amino acid. 1000 bootstrap replicates were utilized to obtain standard errors of the estimates. A number of approaches had been adopted to identify the recombination in our data which were BOOTSCAN (Kosakovsky Pond et al. 2006), RDP (Martin & Rybicki 2000), CHIMAERA (Posada & Crandall 2001), MAXCHI (Smith 1992), GENECONV (Sawyer 1989), 3SEQ (Boni et al. 2007) and SISCAN (Gibbs et al. 2000) in the Recombination Detection Program v4 (RDP4) program. Another method was applied using the online program genetic algorithm recombination detection (GARD) (Kosakovsky Pond et al. 2006) to identify the recombination at the Datamonkey website (<http://www.datamonkey.org/>). The highest acceptable p-value for events of recombination was set at 0.01 with 20 nucleotides window size in order to minimize the false-positive error rate. Only the breakpoints which were identified with more than four approaches were regarded as valid.

Since there is plenty of data on Anura in the NCBI database, we screened the data before reconstructing the phylogenetic tree. Mega was then used to reconstruct the phylogenetic tree for all anuran MHC class Ia sequences that could be aligned, including the sequences from our experiments. Next, we selected one allele for each branch of the same species for the next stage of analysis. MEGA5 and MrBayes v3.2.2 were utilized to reconstruct a neighbor-joining (NJ) tree and Bayesian inference trees in order to determine the general phylogenetic state of the genes of MHC class I which had been newly isolated. Considering recombination can hinder the identification of true phylogenetic relationships, the sequences were partitioned as individual regions to deduce more precise phylogenetic trees of MHC anuran species class I genes. According to the Akaike Information Criterion (AIC) in MrModeltest, the most suitable model of nucleotide substitution was selected. MrBayes was used to reconstruct the Bayesian inference trees of each region. Moreover, the matrix of nucleotide distances was utilized via MEGA5 to reconstruct the neighbor-joining trees of individual exon.

False inference of positive selection was ascribed to recombination. Fortunately, according to breakpoints of recombination which had been deduced, false positives can be decreased to an acceptable degree by partitioning strategy in consideration of separate un-recombinant segments (Scheffler et al. 2006). With the purpose of testing whether affirmative selection could shape the evolutionary pattern of the sequences of MHC class I which had been just isolated, three methods had been utilized with partitioning pre-analyses. First, the Codeml subroutine in PAML version 4 was employed to test the existence of positive selection signals. However, provided that the Codeml power would be influenced by the precision of the inferred phylogenetic trees, the other two approaches used to detect the signals of selection were utilized as well. The two approaches, including MEME (Murrell et al. 2012) and FEL (Kosakovsky Pond & Frost 2005) could be conducted at the website of Datamonkey (<http://www.datamonkey.org>). The combined results from all the current studies can be used to obtain a more precise, comprehensive analysis of the sites under positive selection.

RESULTS

ALLELE CHARACTERIZATION OF *O. MARGARETAE* CLASS IA

We obtained 120 sequences through sequencing and identified 104 sequences as an effective sequence by alignment of Anura MHC in the NCBI database. Some of the 104 sequences showed large deletions in Exon 3 and could not be translated into functional proteins. These sequences with missing were discarded from the dataset and 91 sequences were eventually used for allelic analysis. As described previously, nine alleles were identified under the stringent allele screening conditions. These alleles consisted of 61 sequences out of the 91 sequences and had three sequence lengths, 785 bp (allele 7, 8 and 9), 782 bp (allele 1, 2, 3, 4 and 5) and 879 bp (allele 6), respectively. According to comparisons between HLA amino acid sequences of human-beings (GenBank accession number: AAA76608.2) and predicted peptides of nine alleles, forming disulfide bridges (Cys125-Cys188 and Cys227-Cys164 in HLA of human-beings), the four pivotal amino acid sites were conserved in the 9 sequences (Figure 1, marked with '^'). Alleles were submitted with accession numbers into GenBank (Have not yet acquired).

NUMBER OF LOCI AND GENETIC DIVERSITY

We identified one allele in four individuals, two alleles in five individuals and three alleles in one individual. Given our stringent allele screening conditions, the number of individual alleles is greatly underestimated. Therefore, it is reasonable to have more than three alleles in an individual. In addition, *O. margaretae* is a diploid organism and one locus can only produce two alleles. Taken together, we hypothesized that *O.*

margaretae possesses two loci of MHC class Ia gene at least. In addition, phylogenetic analysis showed that *O. margaretae* MHC alleles were clearly divided into three branches with a high degree of confidence (greater than 98%) (Figure 2). In general, alleles at the same locus are usually clustered together. This shows that the *O. margaretae* may have three of MHC class Ia gene loci. For these two reasons, we hypothesized that *O. margaretae* owns two MHC class Ia loci at least.

The *O. margaretae* MHC class I sequences determined in both nucleotide and amino acid displayed a high genetic diversity (Table 1). The complete sequence K2P nucleotide distance of the nine alleles was 0.154 and K2P nucleotide distances of Exon 2, Exon 3 and Exon 4 were 0.299, 0.142 and 0.058, respectively. On the other hand, the full sequence Poisson-corrected amino acid distance of the nine allele sequences was 0.235 and the Poisson-corrected amino acid distance of $\alpha 1$ domain, $\alpha 2$ domain and $\alpha 3$ domain was 0.402, 0.238 and 0.095, respectively. Whether in the complete sequence or in each partition, Poisson-corrected amino acid distance levels were higher than the K2P nucleotide distance level. This observation suggests that there were more non-synonymous mutations compared with synonymous mutations in the newly isolated alleles, which indicated a large number of positive selections in these sequences.

RECOMBINATION DETECTION

The recombination test was performed before positive selection measure given false inferences of positive selection can be generated by recombination. RDP v.4.70 was used to analyze recombination and found that alleles *Odma-09*, *Odma-01*, *Odma-02*, *Odma-07* and *Odma-06* with breakpoint at codon 262, 235, 183, 295 and 219, respectively, were recombinant products (Table 2). The actual amount of recombinants was likely to be more than three as conservative parameters were established to avoid miscalculations. It is worth reminding that the recombination alleles took place for a large proportion (55.56%) of the alleles of *O. margaretae* MHC class Ia genes in this study.

SELECTION DETECTION

As previously mentioned, partitioning strategy considering separate recombinant segments based on deduced recombination breakpoints can reduce false positives to an acceptable level. According to studies which utilized PAML, MEME and FEL though the positive selected sites identified by different approaches were different (Figure 1 & Table 3), *O. margaretae* alleles of MHC class Ia went through positive selection. To have undergone positive selection over one program was adopted to identify a total of 8 of the 23 sites. However, most of the positive selection sites (19 of 23) are located in the $\alpha 1$ and $\alpha 2$ domains where the ABS is located. Most of the positive selection sites (10 of 19) located at the $\alpha 1$ and $\alpha 2$ domains belonged to the presumed

TABLE 3. Summary statistics for codon sites undergoing positive selection identified by different methods

Codon	Codeml ($p>0.950$)	MEME ($p<0.05$)	FEL (Bayes Factor >99)
$\alpha 1$ domain			
20	0.991		1401.87
50		0.024	
56	0.973		
57	0.959		414.916
60	0.997		
61	0.998		3088.07
62			129.367
75		0.042	
$\alpha 2$ domain			
5	1.000		
7	1.000		37330.1
23	1.000		1477.09
25	1.000		
42	0.977		
61	1.000		
67	1.000		99.8888
74	1.000		
77	1.000	0.005	
78		0.008	
84		0.000	
$\alpha 3$ domain			
9	0.991		18648.7
13			828.191
38			613.27
41			1208.7

ABS (Bjorkman et al. 1987a, 1987b) excluding four sites on domain $\alpha 1$ and five sites on domain $\alpha 2$ (Figure 1).

PHYLOGENETIC ANALYSES

The selection criteria for the alleles used to reconstruct the phylogenetic tree are as described above. Screened alleles (Table 4) and nine alleles from this study were used to reconstruct phylogenetic trees to resolve the phylogenetic status of Anura. It was found that genes of MHC class Ia and MHC class Ib clustered severally and clusters of MHC class Ia gene contained all the nine *O. margaretae* alleles (Figure 2). The phylogenetic tree showed that the MHC alleles of *O. margaretae* were divided into three portions, one section was clustered with the *Rana* including *Rana yavapaiensis* and *R. catesbeiana*; one part was clustered into two independent branches; one portion was mixed with *Rana*, *Polypedates* and *Rhacophorus* including *Rana clamitans*, *Rhacophorus chenfui*, *Polypedates megacephalus* and *R. omeimontis*. The two MHC class Ib sequences of *Xenopus tropicalis* and *X. laevis* were independent of all the MHC class Ia sequences and clustered a single branch.

TABLE 4. The sequences downloaded from GenBank in this study

Species name	Allele number	GenBank number
<i>Rhacophorus omeimontis</i>	<i>Rhom-UA*18</i>	KC261658.1
	<i>Rhom-UA*17</i>	KC261660.1
	<i>Rhom-UA*15</i>	KC261658.1
	<i>Rhom-UA*12</i>	KC261655.1
<i>Rhacophorus chenfui</i>	<i>Poch-UA*19</i>	KX467522
	<i>Poch-UA*13</i>	KX467516
	<i>Poch-UA*12</i>	KX467515
	<i>Poch-UA*07</i>	KX467510
<i>Polypedates megacephalus</i>	<i>Pome-UA*07</i>	KC261643.1
	<i>Pome-UA*04</i>	KC261640.1
	<i>Pome-UA*02</i>	KC261638.1
<i>Smilisca phaeota</i>	<i>Smph-UA*09</i>	JQ679388.1
	<i>Smph-UA*03</i>	JQ679382.1
<i>Rana yavapaiensis</i>	<i>Liya-UA*08</i>	JQ679378.1
	<i>Liya-UA*07</i>	JQ679377.1
	<i>Liya-UA*06</i>	JQ679376.1
<i>Rana clamitans</i>	<i>Licl-UA*13</i>	JQ679367.1
	<i>Licl-UA*08</i>	JQ679362.1
	<i>Licl-UA*05</i>	JQ679359.1
<i>Rana catesbeiana</i>	<i>Lica-UA*12</i>	JQ679354.1
	<i>Lica-UA*08</i>	JQ679350.1
	<i>Lica-UA*06</i>	JQ679348.1
	<i>Lica-UA*02</i>	JQ679344.1
<i>Espadarana prosoblepon</i>	<i>Espr-UA*06</i>	JQ679336.1
	<i>Espr-UA*05</i>	JQ679335.1
	<i>Espr-UA*01</i>	JQ679331.1
<i>Agalychnis callidryas</i>	<i>Agca-UA*01</i>	JQ679312.1
<i>Xenopus laevis</i>	<i>Xela-UA*05</i>	DQ149600.1
<i>Xenopus tropicalis</i>		NM_001247995.1
<i>Xenopus laevis</i>		L20731.1

Taking into account the recombination between Exons 2, 3 and 4 as well as the difference of mutation rate and function between these exon, a portion of the *O. margaretae* MHC class Ia alleles branched from common ancestors of *Rana* species; the other portion branched from common ancestors of *Polypedates* and *Rhacophorus* species. In the tree of Exon 3, *O. margaretae* MHC class Ia alleles were mixing of different species and exhibited a star-like topology, likely due to the informative mutations. The tree reconstructed for Exon 4 was seemingly consistent with species tree because there were four clusters that corresponded to the anuran taxa: alleles from *L. yavapaiensis*, *L. clamitans* and *L. catesbeianus* formed the Ranidae cluster; alleles from *Rhacophorus omeimontis*, *Rhacophorus chenfui* and *Polypedates megacephalus* belonged to the *Rhacophorus* and *Polypedates* cluster; and one portion of the alleles obtained from *O. margaretae* formed the *Odorrana* cluster (Figure 3). Notably, a large number of MHC class Ia alleles from *O. margaretae* were scattered in the clusters of *Rana*, *Polypedates* and *Rhacophorus*.

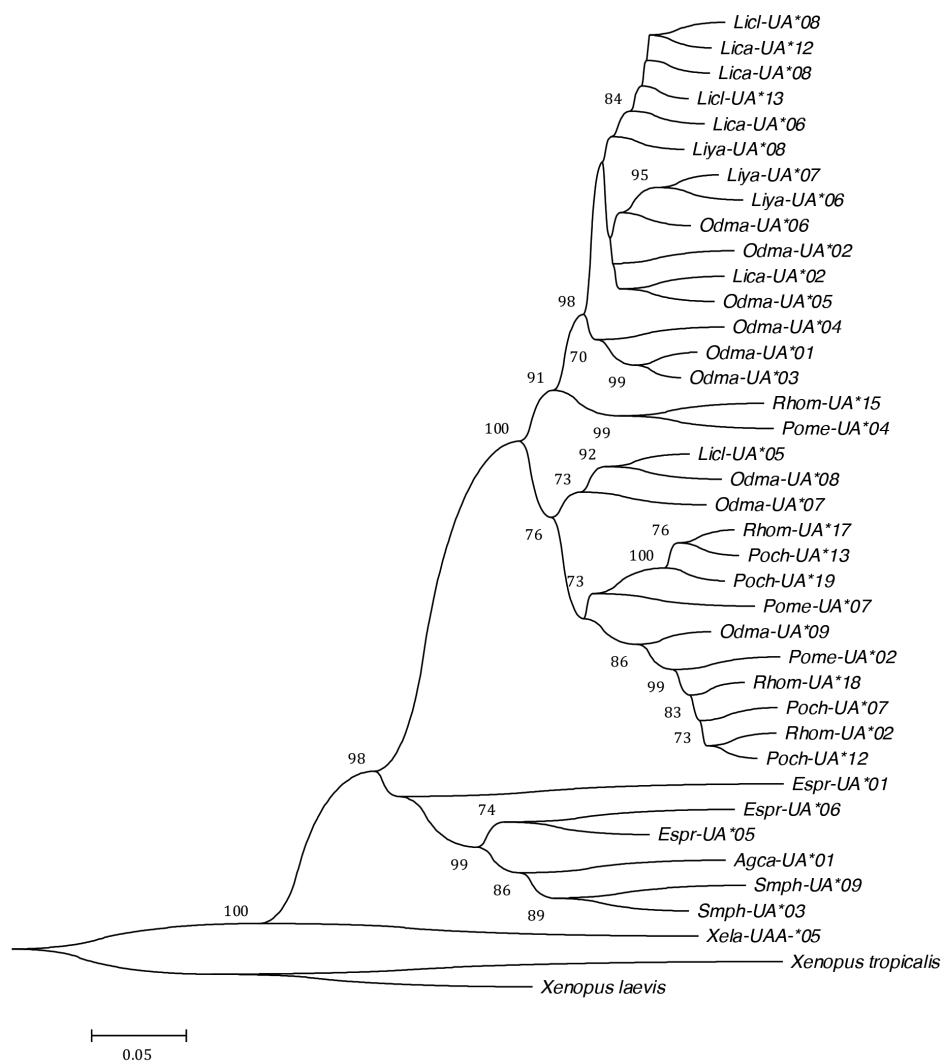


FIGURE 2. Phylogenetic relationships of anuran MHC class Ia and Ib genes. The neighbor-joining tree was constructed with MEGA5 using all available anuran MHC class I genes with a sufficient length. Bootstrap values are indicated above the branches

DISCUSSION

Nine of the MHC class Ia alleles of *O. margaretae* were obtained from 10 individuals by using a pair of universal primers designed by Zhao et al. (2013). Subsequently, the polymorphism of these 9 alleles as well as the formation mechanism of polymorphisms were analyzed by using a series of bioinformatics analysis software. In addition, we examined the evolutionary relationship of the anuran MHC class Ia alleles and the evolutionary status of the MHC class Ia alleles of *O. margaretae*. Data showed that: MHC class Ia alleles of *O. margaretae* have a high degree of polymorphism both in its nucleotide and amino acid sequences; polymorphism of *O. margaretae* MHC class Ia genes was closely related to gene recombination and positive selection; and *O. margaretae* MHC class Ia genes have a common ancestor with *Rana*, *Polypedates* and *Rhacophorus*.

The MHC class Ia alleles of several anuran species, such as *Rana*, *Agalychnis*, *Espadarana*, *Smilisca* (Kiemnec-

Tyburczy et al. 2012), *Polypedates* and *Rhacophorus* (Zhao et al. 2013), were identified and found to be highly polymorphic. However, the *Odorrana* (approximately 50 species) is an important part of the anuran and no reports of MHC class Ia gene have been reported in *Odorrana*. In addition, the taxonomy of the *Odorrana* has been controversial (Dubois 1992; Matsui 1994). In this study, we argue that the isolation and analysis of the MHC class Ia alleles of an *Odorrana* species, namely *O. margaretae*, can provide the basic data for understanding the evolution of the *Odorrana*. The data from this study show that the MHC class Ia of *O. margaretae* has a high degree of polymorphism and this result is similar to that found in the *Rana*, *Agalychnis*, *Espadarana*, *Smilisca* (Kiemnec-Tyburczy et al. 2012), *Polypedates* and *Rhacophorus* (Zhao et al. 2013). The K2P nucleotide distances of MHC class Ia Exon 2, 3, and 4 ranged from 0.076 to 0.202, the Poisson-corrected amino acid distances was between 0.134 and 0.287 in these anuran (Kiemnec-Tyburczy et al. 2012; Zhao et al. 2013). The ubiquitous Poisson-corrected amino

species in the range of one to three in anuran (Kiemnec-Tyburczy et al. 2012; Zhao et al. 2013).

Recombination is also known to play a role in MHC polymorphism in anuran (Consuegra et al. 2005; Jakobsen et al. 1998; Wang et al. 2010; Wutzler et al. 2012). Recombination in current research plays a significant role in the progress of MHC class Ia genes in *O. margaretae* and generates 55.56% of the alleles. Recombination was also found in other anuran species (Kiemnec-Tyburczy et al. 2012; Zhao et al. 2013). For example, in *Polypedates megacephalus* and *Rhacophorus omeimontis*, the five recombinant alleles accounted for a relatively large proportion (18.5%) of MHC class Ia alleles (Zhao et al. 2013). In six non-model frogs, three measures detected a significant signal of recombination (Kiemnec-Tyburczy et al. 2012). It is noteworthy that a large number of recombinant breakpoints were put near the exon boundary. This has been found in a large number of anuran, such as *Rana clamitans*, *Polypedates megacephalus*, *Rhacophorus chenfu*, *R. omeimontis*, *X. laevis* and *O. margaretae* (Kiemnec-Tyburczy et al. 2012; Zhao et al. 2013). The recombination pattern is similar to that observed in the genes of MHC class I in bony fishes (Shum et al. 2001). According to prior research, intron 2 of genes of MHC class I in bony fishes contains numerous repeated elements and shows long coverage (Shum et al. 2001). These conditions showed that according to the genetic information from the entire genome of *X. tropicalis*, it could be deduced that Rhacophoridae species tended to have a big intron 2 regions in the genes of MHC class I (Bos & Waldman 2006b). These results suggested that the possible mechanism for recombination is the exchange between entire exon. For example, in *X. laevis*, two recombinant alleles (Xela-UAA*06 and UAA*11) have corresponding a1 domain sequences but different a2 and a3 domains (Bos & Waldman 2006b), which was analogous to other studies in the anuran (Kiemnec-Tyburczy et al. 2012; Zhao et al. 2013). In general, recombination is central to increasing genetic diversity in anuran and the breakpoints are usually located at the boundary of the exon.

Natural selection is a key process that acts on MHC genes in most vertebrate (Bernatchez & Landry 2003; Hughes 1998). Evidence has shown that *O. margaretae* genes of MHC class I experience natural selection. Firstly, in comparison with the divergence of amino acid sequence, the nucleotide sequence divergence of the genes was decreased in consistent with the positive selection signal. Secondly, under positive selection in the nine alleles of *O. margaretae* MHC class I, natural selection signals and sites have been detected by lots of codon-oriented approaches. The sites identified by over one approach were regarded robust and they were the subset of putative ABS from the human HLA crystal structure (Bjorkman et al. 1987a, 1987b). This result is concordant with that found in the alleles of MHC class Ia in other species (Bos & Waldman 2006a; Kiemnec-Tyburczy et al. 2012; Zhao et al. 2013). It can be concluded that the genes of MHC class I tend to undergo pathogen-mediated balancing selection given the

ABS function. However, there are some sites that showed significant signatures of positive selection that were not located in the human ABS (Bos & Waldman 2006a; Kiemnec-Tyburczy et al. 2012; Zhao et al. 2013). The underlying cause may be due to differences in the crystal structure of the MHC in anuran versus human. Previous studies have shown the differences between the crystal structures of chicken and human class I molecules (Koch et al. 2007). Hence, the crystal structures of frogs MHC are necessary to decode the functional role of the sites under positive selection.

Given that the taxonomy of the *Odorrana* has been controversial (Dubois 1992; Matsui 1994), MHC class Ia alleles from the *O. margaretae* and alleles selected from other species were used to reconstruct phylogenetic trees to solve the evolutionary status of the *O. margaretae*. While evolutionary status of *O. margaretae* MHC class I sequences in this study is also unclear, our results are not entirely unexpected, especially given trans-species polymorphisms of anuran genes of MHC class Ia were evident in anuran (Kiemnec-Tyburczy et al. 2012; Zhao et al. 2013b). To be more specific, Exon 2 and Exon 3 phylogenetic trees of anuran genes of MHC class I, the trans-species polymorphisms can be observed and species-specific topology could not be observed (Figure 3) (Zhao et al. 2013). This finding suggests that concerted evolution play a role in the trans-species polymorphism of anuran of MHC class Ia (Zhao et al. 2013). Finally, balancing selection tended to be related to the pattern of trans-species polymorphism. To be more specific, Exon3 and Exon 4 contained major regions for ABS and recognition experience strong balancing selection (Figure 3) (Zhao et al. 2013). The functions of $\alpha 3$ expressed by Exon 4 are a linker and an anchor, suffering from fewer constraints of selection. As consequence, the precise phylogenetic relationship of these species can be shown in the pedigree of Exon 4 commonly. Prior researches have shown the differences in the patterns of evolution for different exon of the same gene (Kiemnec-Tyburczy et al. 2012; Nonaka et al. 2011; Wang et al. 2010).

In summary, alleles of MHC class Ia were isolated successfully from an *Odorrana* species: *O. margaretae*, which were considered to contain two locus of MHC class Ia in *O. margaretae* genome. Because of the exchanges of whole exon and positive selection acting predominantly on the ABS, frequent gene duplications of recombination were regarded as significant in the evolution of genes of anuran MHC class I. Based on this study, our knowledge of the anuran MHC genes evolution was broadened by an *Odorrana* species: *O. margaretae* MHC class I.

ACKNOWLEDGEMENTS

The capture activity was approved by the Administration of Wawushan Nature Reserve. This work was supported by grants from the China National Natural Science Foundation (No. 31470444) to M.W Zhang, Foundation of Education Department of Sichuan to M.W. Zhang (13ZA0262) and

the Innovative Research Team in University of Sichuan Bureau of Education (14TD0002). All procedures related to live frogs experiments were approved by Sichuan Agricultural University's Animal Ethics Committee (SA20160725).

REFERENCES

- Aguilar, A., Roemer, G., Debenham, S., Binns, M., Garcelon, D. & Wayne, R.K. 2004. High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *PNAS* 101(10): 3490-3494.
- Andersson, L. & Mikko, S. 1995. Generation of MHC class II diversity by intra- and intergenic recombination. *Immunol. Rev.* 143: 5-12.
- Begovich, A.B., McClure, G.R., Suraj, V.C., Helmuth, R.C., Fildes, N., Bugawan, T.L., Erlich, H.A. & Klitz, W. 1992. Polymorphism, recombination, and linkage disequilibrium within the HLA class II region. *J. Immunol.* 148(1): 249-258.
- Bernatchez, L. & Landry, C. 2003. MHC studies in nonmodel vertebrates: What have we learned about natural selection in 15 years? *J. Evol. Biol.* 16: 363-377.
- Bjorkman, P., Saper, M., Samraoui, B., Bennett, W., Strominger, J. & Wiley, D. 1987a. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* 329(6139): 512-518.
- Bjorkman, P.J., Saper, M., Samraoui, B., Bennett, W.S., Strominger, J.L. & Wiley, D. 1987b. Structure of the human class I histocompatibility antigen, HLA-A 2. *Nature* 329: 506-512.
- Boni, M.F., Posada, D., Feldman, M.W. & Boni, M.F. 2007. An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics* 176: 1035-1047.
- Bos, D. & Waldman, B. 2006a. Polymorphism, natural selection, and structural modeling of class Ia MHC in the African clawed frog (*Xenopus laevis*). *Immunogenetics* 58: 433-442.
- Bos, D.H. & Waldman, B. 2006b. Evolution by recombination and transspecies polymorphism in the MHC class I gene of *Xenopus laevis*. *Mol. Biol. Evol.* 23: 137-143.
- Cardenas, P.P., Suarez, C.F., Martinez, P., Patarroyo, M.E. & Patarroyo, M.A. 2005. MHC class I genes in the owl monkey: Mosaic organisation, convergence and loci diversity. *Immunogenetics* 56: 818-832.
- Carey, C., Cohen, N. & Rollins-Smith, L. 1999. Amphibian declines: An immunological perspective. *Dev. Comp. Immunol.* 23: 459-472.
- Consuegra, S., Megens, H.J., Schaschl, H., Leon, K., Stet, R. & Jordan, W. 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.
- Daszak, P., Berger, L., Cunningham, A.A., Hyatt, A.D., Green, D.E. & Speare, R. 1999. Emerging infectious diseases and amphibian population declines. *Emerging Infectious Diseases* 5(6): 735-748.
- Dubois A. 1992. Notes sur la classification des Ranidae (Amphibiens Anoures). *Bulletin Mensuel de la Société Linnéenne de Lyon* 61: 305-352.
- Edwards, S.V. & Hedrick, P.W. 1998. Evolution and ecology of MHC molecules: From genomics to sexual selection. *Trends Ecol. Evol.* 13(8): 305-311.
- Evans, B.J., Kelley, D.B., Tinsley, R.C., Melnick, D.J. & Cannatella, D.C. 2004. A mitochondrial DNA phylogeny of African clawed frogs: Phylogeography and implications for polyploid evolution. *Mol. Phylog. Evol.* 33(1): 197-213.
- Figuroa, F., Mayer, W.E., Sato, A., Zaleska-Ruczynska, Z., Hess, B., Tichy, H. & Klein, J. 2001. Mhc class I genes of swordtail fishes, *Xiphophorus*: Variation in the number of loci and existence of ancient gene families. *Immunogenetics* 53: 695-708.
- Flajnik, M.F. & Kasahara, M. 2001. Comparative genomics of the MHC: Glimpses into the evolution of the adaptive immune system. *Immunity* 15(3): 351-362.
- Garrigan, D., Hedrick, P.W. & Mitton, J. 2003. Perspective: Detecting adaptive molecular polymorphism: Lessons from the MHC. *Evolution* 57: 1707-1722.
- Gibbs, M.J., Armstrong, J.S. & Gibbs, A.J. 2000. Sister-scanning: A Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* 16: 573-582.
- Haddad, C.F., Pombal Jr., J.P. & Batistic, R.F. 1994. Natural hybridization between diploid and tetraploid species of leaf-frogs, genus *Phyllomedusa* (Amphibia). *J. Herpetol.* 28(4): 425-430.
- Hughes, A.M. 1998. Natural selection at major histocompatibility complex loci of vertebrates. *Annual Review of Genetics* 32: 415-435.
- Jakobsen, I.B., Wilson, S.R. & Easta, S. 1998. Patterns of reticulate evolution for the classical class I and II HLA loci. *Immunogenetics* 48: 312-323.
- Jeffery, K.J.M. & Bangham, C.R.M. 2000. Do infectious diseases drive MHC diversity? *Microbes & Infection* 2: 1335-1341.
- Kaufman, J., Milne, S., Göbel, T.W., Walker, B.A., Jacob, J.P., Auffray, C., Zoorob, R. & Beck, S. 1999. The chicken B locus is a minimal essential major histocompatibility complex. *Nature* 401: 923-925.
- Kiemnec-Tyburczy, K., Richmond, J., Savage, A., Lips, K. & Zamudio, K. 2012. Genetic diversity of MHC class I loci in six non-model frogs is shaped by positive selection and gene duplication. *Heredity* 109: 146-155.
- Klein, J. 1986. *Natural History of the Major Histocompatibility Complex*. New York: Wiley.
- Koch, M., Camp, S., Collen, T., Avila, D., Salomonsen, J., Wallny, H.J., van Hateren, A., Hunt, L., Jacob, J.P. & Johnston, F. 2007. Structures of an MHC class I molecule from B21 chickens illustrate promiscuous peptide binding. *Immunity* 27: 885-899.
- Kosakovsky Pond, S.L., Posada, D., Gravenor, M.B., Woelk, C.H. & Frost, S.D. 2006. Automated phylogenetic detection of recombination using a genetic algorithm. *Molecular Biology & Evolution* 23: 1891-1901.
- Kosakovsky Pond, S.L. & Frost, S.D. 2005. Not so different after all: A comparison of methods for detecting amino acid sites under selection. *Molecular Biology & Evolution* 22(5): 1208-1222.
- Martin, D. & Rybicki, E. 2000. RDP: Detection of recombination amongst aligned sequences. *Bioinformatics* 16(6): 562-563.
- Matsui, M. 1994. A taxonomic study of the *Rana narina* complex, with description of three new species (Amphibia: Ranidae). *Zool. J. Linn. Soc.* 111: 385-415.
- McCairns, R.J.S., Bourget, S. & Bernatchez, L. 2011. Putative causes and consequences of MHC variation within and between locally adapted stickleback demes. *Mol. Ecol.* 20: 486-502.
- Miller, K.M. & Withler, R.E. 1998. The salmonid class I MHC: Limited diversity in a primitive teleost. *Immunol. Rev.* 166: 279-293.

- Murrell, B., Wertheim, J.O., Moola, S., Weighill, T., Scheffler, K. & Pond, S.L.K. 2012. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* 8: e1002764.
- Nei, M., Gu, X. & Sitnikova, T. 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proceedings of the National Academy of Sciences* 94: 7799-7806.
- Nei, M. & Rooney, A.P. 2005. Concerted and birth-and-death evolution of multigene families. *Annu. Rev. Genet.* 39: 121-152.
- Nonaka, M.I., Aizawa, K., Mitani, H., Bannai, H.P. & Nonaka, M. 2011. Retained orthologous relationships of the MHC Class I genes during euteleost evolution. *Mol. Biol. Evol.* 28(11): 3099-3112.
- Nonaka, M., Namikawa, C., Kato, Y., Sasaki, M., Salter-Cid, L. & Flajnik, M.F. 1997. Major histocompatibility complex gene mapping in the amphibian *Xenopus* implies a primordial organization. *Proceedings of the National Academy of Sciences* 94: 5789-5791.
- Piertney, S. & Oliver, M. 2006. The evolutionary ecology of the major histocompatibility complex. *Heredity* 96: 7-21.
- Posada, D. & Crandall, K.A. 2001. Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *P.N.A.S.* 98(24): 13757-13762.
- Potts, W.K. & Wakeland, E.K. 1990. Evolution of diversity at the major histocompatibility complex. *Trends Ecol. Evol.* 5: 181-187.
- Pounds, J.A., Bustamante, M.R., Coloma, L.A., Consuegra, J.A., Fogden, M.P., Foster, P.N., La Marca, E., Masters, K.L., Merino-Viteri, A & Puschendorf, R. 2006. Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature* 439: 161-167.
- Ptacek, M.B., Gerhardt, H.C. & Sage, R.D. 1994. Speciation by polyploidy in treefrogs: Multiple origins of the tetraploid, *Hyla versicolor*. *Evolution* 48(3): 898-908.
- Radwan, J., Biedrzycka, A. & Babik, W. 2010. Does reduced MHC diversity decrease viability of vertebrate populations? *Biol. Conserv.* 143: 537-544.
- Sato, A., Klein, D., Sülthmann, H., Figueroa, F., O'hUigin, C. & Klein, J. 1997. Class I Mhc genes of cichlid fishes: Identification, expression, and polymorphism. *Immunogenetics* 46: 63-72.
- Sawyer, S. 1989. Statistical tests for detecting gene conversion. *Molecular Biology & Evolution* 6: 526-538.
- Scheffler, K., Martin, D.P. & Seoighe, C. 2006. Robust inference of positive selection from recombining coding sequences. *Bioinformatics* 22: 2493-2499.
- She, J.X., Boehme, S.A., Wang, T.W., Bonhomme, F. & Wakeland, E.K. 1991. Amplification of major histocompatibility complex class II gene diversity by intraexonic recombination. *P.N.A.S.* 88: 453-457.
- Shum, B.P., Guethlein, L., Flodin, L.R., Adkison, M.A., Hedrick, R.P., Nehring, R.B., Stet, R.J., Secombes, C. & Parham, P. 2001. Modes of salmonid MHC class I and II evolution differ from the primate paradigm. *The Journal of Immunology* 166: 3297-3308.
- Smith, J.M. 1992. Analyzing the mosaic structure of genes. *J. Mol. Evol.* 34: 126-129.
- Sommer, S. 2005. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in Zoology* 2:16. <https://doi.org/10.1186/1742-9994-2-16>.
- Spurgin, L.G. & Richardson, D.S. 2010. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society of London B: Biological Sciences* DOI: 10.1098/rspb.2009.2084
- Stuart, S.N., Chanson, J.S., Cox, N.A., Young, B.E., Rodrigues, A.S., Fischman, D.L. & Waller, R.W. 2004. Status and trends of amphibian declines and extinctions worldwide. *Science* 306: 1783-1786.
- Teacher, A.G.F., Garner, T.W.J. & Nichols, R.A. 2009. 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* 4: e4616.
- Trowsdale, J. 2011. The MHC, disease and selection. *Immunol. Lett.* 137: 1-8.
- Van Eden, W., Devries, R. & Van Rood, J. 1983. The genetic approach to infectious disease with special emphasis on the MHC. *Dis. Mark.* 1: 221-242.
- Vincek, V., Nizetić, D., Golubić, M., Figueroa, F., Nevo, E. & Klein, J. 1987. Evolutionary expansion of Mhc class I loci in the mole-rat, *Spalax ehrenbergi*. *Mol. Biol. Evol.* 4: 483-491.
- Wang, D., Zhong, L., Wei, Q., Gan, X. & He, S. 2010. Evolution of MHC class I genes in two ancient fish, paddlefish (*Polyodon spathula*) and Chinese sturgeon (*Acipenser sinensis*). *FEBS Lett.* 584: 3331-3339.
- Wegner, K.M., Reusch, T.B.H. & Kalbe, M. 2003. Multiple parasites are driving major histocompatibility complex polymorphism in the wild. *J. Evol. Biol.* 16: 224-232.
- Wenyuan, C., Zisu, W., Xizhong, W., Yuhua, Y. & Qiling, S. 1983. A comparative study of the karyotypes from six species of frogs in Sichuan. *Zoological Research* 4: 83-88.
- Westerdahl, H., Wittzell, H., von Schantz, T. & Bensch, S. 2004. MHC class I typing in a songbird with numerous loci and high polymorphism using motif-specific PCR and DGGE. *Heredity* 92: 534-542.
- Wutzler, R., Foerster, K. & Kempnaers, B. 2012. MHC class I variation in a natural blue tit population (*Cyanistes caeruleus*). *Genetica* 140(7-9): 349-364.
- Xu, K., Zhu, D.Z., Wei, Y., Schloegel, L.M., Chen, X.F. & Wang, X.L. 2010. Broad distribution of ranavirus in free-ranging *Rana dybowskii* in Heilongjiang, China. *EcoHealth* 7: 18-23.
- Zhao, M., Wang, Y., Shen, H., Li, C., Chen, C., Luo, Z. & Wu, H. 2013a. Evolution by selection, recombination, and gene duplication in MHC class I genes of two Rhacophoridae species. *BMC Evol. Biol.* 13: 113. <https://doi.org/10.1186/1471-2148-13-113>.

College of Animal Science and Technology
Sichuan Agricultural University
211# Huimin Road, Chengdu
China

*Corresponding author; email: mwzhangkiz@hotmail.com

Received: 28 February 2018

Accepted: 4 June 2018