Developmental and Comparative Immunology 42 (2014) 311–322

Contents lists available at ScienceDirect

Developmental and Comparative Immunology

journal homepage: www.elsevier.com/locate/dci

Extensive diversification of MHC in Chinese giant salamanders *Andrias davidianus* (Anda-MHC) reveals novel splice variants

Rong Zhu, Zhong-yuan Chen, Jun Wang, Jiang-di Yuan, Xiang-yong Liao, Jian-Fang Gui, Qi-Ya Zhang *

State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China

ARTICLE INFO

Article history:
Received 6 September 2013
Revised 2 October 2013
Accepted 3 October 2013
Available online 14 October 2013

Keywords:
MHC of Chinese giant salamanders *Andrias davidianus* (Anda-MHC)
Genetic diversity
Alternative splice variant
Andrias davidianus ranavirus (ADRV)
Inducible expression

ABSTRACT

A series of MHC alleles (including 26 class IA, 27 class IIA, and 17 class IIB) were identified from Chinese giant salamander *Andrias davidianus* (Anda-MHC). These genes are similar to classical MHC molecules in terms of characteristic domains, functional residues, deduced tertiary structures and genetic diversity. The majority of variation between alleles is found in the putative peptide-binding region (PBR), which is driven by positive Darwinian selection. The coexistence of two isoforms in MHC IA, IIA, and IIB alleles are shown: one full-length transcript and one novel splice variant. Despite lake of the external domains, these variants exhibit similar subcellular localization with the full-length transcripts. Moreover, the expression of MHC isoforms are up-regulated upon *in vivo* and *in vitro* stimulation with *Andrias davidianus* ranavirus (ADRV), suggesting their potential roles in the immune response. The results provide insights into understanding MHC variation and function in this ancient and endangered urodele amphibian.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

The Chinese giant salamander *Andrias davidianus* belongs to one of the most primitive orders of urodele amphibians, the *Cryptobranchidae* (Zhang et al., 2003). It is the largest extant amphibian species, and found only in China (Wang et al., 2013). Being crown as a living fossil from 350 million years ago, and representing a transitional form that links aquatic to terrestrial organisms, it is considered to be a valuable model in studies on vertebrate evolution and biodiversity (Gao and Shubin, 2003; Robert and Cohen, 2011). The population has declined dramatically in the past five decades, and it has now been included in the list of Appendix I of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES, 2008) and in national class II protected species in China. Although the artificial breeding have been ongoing, the frequent outbreaks of infectious diseases pose a serious threat to the Chinese giant salamander population (Dong et al., 2011; Geng et al., 2011). *Andrias davidianus* ranavirus (ADRV), an emerging viral pathogen, is associated with mass mortality in farmed salamander (Chen et al., 2013; Zhang and Gui, 2012). To establish effective measures for the disease control, the development of genetic markers related to pathogen resistance is of great importance (Gui and Zhu, 2012).

One of the most ideal genetic markers in conservation programs is the major histocompatibility complex (MHC) (Sommer, 2005). It is a multigene family central to the vertebrate immune system by presenting self/non-self antigen peptides to T lymphocytes (Neefjes et al., 2011). A functional hallmark of MHC genes is extensive diversification concentrated in the peptide-binding region (PBR). MHC diversity is tightly linked to diseases resistance, and maintained through balancing selection mediated by host–pathogen co-evolution (Piertney and Oliver, 2006; Spurgin and Richardson, 2010). Hence, the well-characterized function of MHC in immune defense, alongside their outstanding diverse nature, makes them exceptional candidates to study patterns of adaptive genetic variation determining pathogen resistance, specially in species of conservation concern (Sutton et al., 2011).

To date, extensive research on MHC genes within amphibians has focused on model organisms such as the anuran *Xenopus* (Ohta et al., 2006) and the urodele axolotl (*Ambystoma mexicanum*) (Laurens et al., 2001). Compelling evidence showing high similarity in MHC structure and complexity between amphibians and mammals, and also close associations between MHC diversity and diseases resistance (Savage and Zamudio, 2011; Teacher et al., 2009). However, little information has so far been available about the functional role of MHC in the Chinese giant salamander. In our recent effort, a thymus cDNA library was constructed from Chinese giant salamander infected with ADRV. Screening for expressed sequence tags (ESTs) revealed a number of immune-related genes including those encoding MHC (Anda-MHC) and beta2-microglobuin (Anda-β2M). In this study, we embarked on investigating the variation and expression of these genes. The phylogenetic analyses were performed across vertebrate species, and the three-dimensional (3D) structures were predicted by homology modeling. The sequence polymorphism was examined, and the

* Corresponding author. Tel.: +86 27 68780792; fax: +86 27 68780123.
E-mail address: zhangqy@ihb.ac.cn (Q.-Y. Zhang).

0145-305X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.dci.2013.10.001

Table 1
Random sites model estimates for MHC IA, IIA and IIB alleles of the Chinese giant salamander.

Groups	Model	*P*	*lnL*	Parameter estimates	LRT	2Δ*lnL*	Positively selected sites
MHC IA	M0: one-ration	1	−2684.2561	κ = 2.6518, ω = 0.8073			
	M1: nearly neutral	1	−2648.4741	κ = 2.4850, p_0 = 0.5622, p_1 = 0.4378, ω_0 = 0.0000, ω_1 = 1.0000			
	M2: positive selection	3	−2626.1863	κ = 2.8482, p_0 = 0.5182, p_1 = 0.3740, p_2 = 0.1078, ω_0 = 0.0000, ω_1 = 1.0000, **ω_2 = 6.3411**	M2 vs M1	44.5757	90L, **122V**, **124Y**, 129R, 139F, **182E**, 189E, 193A, 196K
	M3: discrete	5	−2625.6300	κ = 2.8574, p_0 = 0.5823, p_1 = 0.3465, p_2 = 0.0712, ω_0 = 0.0000, ω_1 = 1.5611, **ω_2 = 8.2568**	M3 vs M0	117.2522	
	M7: beta	2	−2649.3795	κ = 2.5164, p = 0.0051, q = 0.0051			
	M8: beta & omega	4	−2626.2206	κ = 2.8415, p_0 = 0.8871, p_1 = 0.1129, p = 0.050, q = 0.0075, **ω = 6.0823**	M8 vs M7	46.3178	90L, **122V**, **124Y**, 129R, 139F, **182E**, 189E, 193A, 196K
MHC IIA	M0: one-ration	1	−1394.5989	κ = 2.8839, ω = 1.0114			
	M1: nearly neutral	1	−1394.3163	κ = 2.8043, p_0 = 0.1782, p_1 = 0.8218, ω_0 = 0.0000, ω_1 = 1.0000			
	M2: positive selection	3	−1392.9339	κ = 2.8416, p_0 = 0.9488, p_1 = 0.0000, p_2 = 0.9488, ω_0 = 0.7422, ω_1 = 1.0000, **ω_2 = 5.8360**	M2 vs M1	2.7648	9P, 18G, 32I, 87M
	M3: discrete	5	−1392.9339	κ = 2.8416, p_0 = 0.4599, p_1 = 0.4889, p_2 = 0.0512, ω_0 = 0.7422, ω_1 = 0.7422, **ω_2 = 5.8360**	M3 vs M0	3.3300	
	M7: beta	2	−1394.3779	κ = 2.8398, p = 0.0473, q = 0.0050			
	M8: beta & omega	4	−1392.9349	κ = 2.8417, p_0 = 0.9492, p_1 = 0.0508, p = 99.0000, q = 33.9983, **ω = 5.8612**	M8 vs M7	2.8859	9P, 18G, 32I, 87M
MHC IIB	M0: one-ration	1	−1533.1951	κ = 3.0802, ω = 2.9798			
	M1: nearly neutral	1	−1533.7786	κ = 2.5054, p_0 = 0.4423, p_1 = 0.5578, ω_0 = 0.0000, ω_1 = 1.0000			
	M2: positive selection	3	−1506.2896	κ = 3.6550, p_0 = 0.4120, p_1 = 0.5011, p_2 = 0.0870, ω_0 = 1.0000, ω_1 = 1.0000, **ω_2 = 27.5920**	M2 vs M1	54.9780	42A, **59L**, **78F**, **79V**, **98I**, **101D**, **102A**, **103R**, **109Y**,161S
	M3: discrete	5	−1506.1608	κ = 3.7159, p_0 = 0.1577, p_1 = 0.7646, p_2 = 0.0777, ω_0 = 0.0000, ω1 = 1.6150, ω_2 = **34.7115**	M3 vs M0	54.0686	
	M7: beta	2	−1533.9187	κ = 2.4607, p = 0.0050, q = 0.0050			
	M8: beta & omega	4	−1506.2896	κ = 3.6550, p_0 = 0.9131, p_1 = 0.0870, p = 2.8720, q = 0.0050, **ω = 27.5919**	M8 vs M7	55.2581	42A, **59L**,**78F**, **79V**, **98I**, **101D**, **102A**, **103R**, **109Y**, 161S

2△*lnL*: Log likelihood difference between models using the c2-test, P: number of free parameters for the ω ratios, κ: transition/transversion rate, ω: Ratio of non-synonymous to synonymous nucleotide substitution, p_n: proportion of sites that fall into the ω_n site class, p, q: shape parameters of the β function (for models M7 and M8), Positively selected sites with posterior probability >0.95 are in bold.

selection pattern was test to explore the underlying mechanisms shaping MHC diversity. The expression profiles of MHC genes upon stimulation with ADRV were investigated to elucidate their immune signification. Additionally, three alternative transcripts lacking the extracellular domains (α1/α2, α2, and β2) were isolated from MHC IA, IIA and IIB, respectively. The expression and subcellular distribution patterns were further compared between the full-length and truncated transcripts. Our study provide important genetic information for conservation of the endangered Chinese giant salamander.

2. Materials and methods

2.1. Salamanders, cells, and virus

Chinese giant salamanders weighing about 30 g were obtained from a farm in Jiangxi, China. The salamanders were maintained at 22 °C in an aerated freshwater tank for 2 weeks, and no clinical signs were observed prior to experiment. Chinese giant salamanders thymus (CGST) cells were cultured at 20 °C in medium 199 with 10% FBS, and *Epithelioma papulosum* cyprini (EPC) cells were maintained at 25 °C in medium 199 with 10% FBS. *Andrias davidianus* ranavirus (ADRV) was isolated from the Chinese giant salamanders and propagated in EPC cells as described previously (Chen et al., 2013).

Table 2
Primers used in this study.

Primer name	Sequence (5′–3′)	Usage
β-Actin-F	CCACTGCTGCCTCCTCTT	Real-time PCR
β-Actin-R	GCAATGCCTGGGTACATG	
MHC IA-F1	GGACTTCATCAGCCTCCACA	
MHC IA-R1	AGGTTCCGTCAGGGTTCG	
β2M-F1	TTTGCTCCTGCTGGTGGT	
β2M-R1	AGATAAGGGTGTTCGGTTTT	
MHC IIA-F1	CTGCTGTCACCGTGTTCC	
MHC IIA-R1	ACTGGTTGCGTGTTCTTCA	
MHC IIB-F1	CACTTCCTGAACGGCTCG	
MHC IIB-R1	CGGGCGGGTAAAAGTCT	
MHC IA-F2	ATGACGTCCCGGACCACTCT	Complete ORF of MHC
MHC IA-R2	GGCGGAGGCGGTGCTGGA	
MHC IIA-F2	ATGGCTGCAGTCCGGTGC	
MHC IIA-R2	TTTCATTCTTCTTCTGTGACTTGGG	
MHC IIB-F2	ATGCGTCCCCCCCTCAATCC	
MHC IIB-R2	TAGGCAACATCATTCTGAGGGA	
MHC IA-F3	CCCTCGAGATGACGTCCCGGACCACT	Subcellular localization
MHC IA-R3	CGCGGATCCGGCGGAGGCGGTGCTGGA	
β2M-F2	CCCTCGAGCTATGGGTACCACTGTGAGG	
β2M-R2	CGGAATTCCAGCTTGGGATCCCAGGTGTG	
MHC IIA-F3	CCCTCGAGATGGCTGCAGTCCGGTGC	
MHC IIA-R3	CGGGATCCTTTCATTCTTCTTCTGTG	
MHC IIB-F3	CCCTCGAGCTATGCGTCCCCCCCTCAATCC	
MHC IIB-R3	CGGGATCCGGCAACATCATTCTGAGGGA	

2.2. Gene cloning and structure modeling

ESTs of MHC IA, β2-Microglobuin (β2M), MHC IIA, and MHC IIB were retrieved from a SMART cDNA library, which was made by mRNA derived from ADRV-infected thymus of the Chinese giant salamanders. To obtain the full-length cDNA, specific primers (Table 2) were designed based on the EST sequences, and RACE-PCR was performed according to the previous report (Zhu et al., 2008). The splice variants were amplified with primers designed to amplify the entire coding sequence of MHC. All PCR products

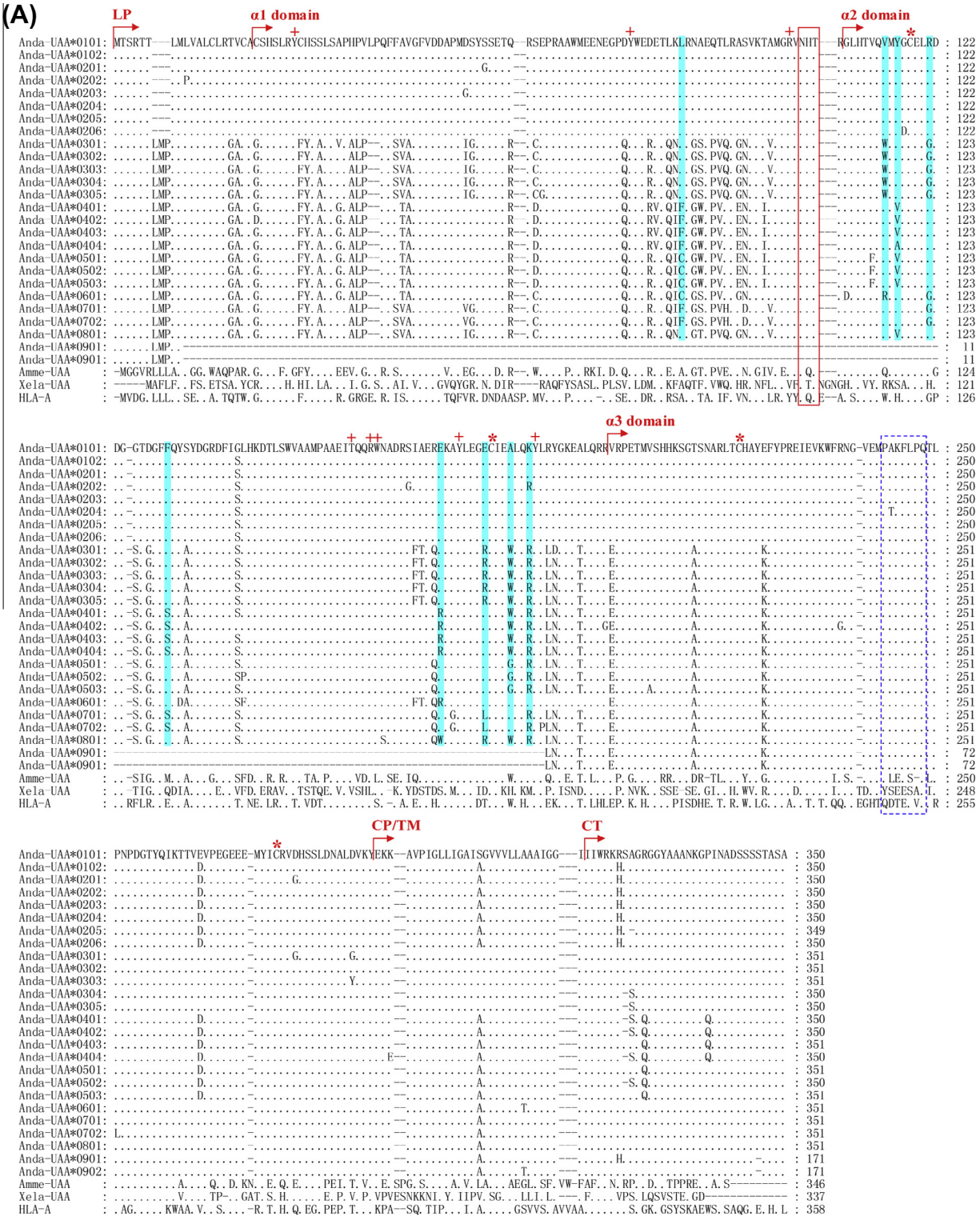

Fig. 1. Multiple alignments of MHC IA (A), β2M (B), MHC IIA (C) and MHC IIB (D). Characteristic motifs for MHC genes including leader peptide (LP), α domain (α1, α2, α3), β domain (β1, β2), connecting peptide (CP), transmembrane (TM) and cytoplasmic (CT) domains are marked above the sequences. The putative peptide-binding region (α1/α2, α1, β1) contain most variable sites and positively selected sites that are highlighted with shading. N-glycosylation sites, conserved cysteines, and peptide anchor residues are denoted by boxes, asterisks, and plus signs, respectively. Ig-like motif in β2M, CD8 binding loop in IA, and CD4 binding loop in IIB are indicated as dashed boxes. GxxxGxxGxxxG/GxxGxxxGxxxxxxG motif, identical amino acids, and gaps are represented by hashes, dots, and dashes, respectively. Species included are axolotl (Amme), Frog (Xela), and human (Hosa/HLA).

were ligated into the pMD18-T vector (Takara) and identified by sequencing.

Multiple alignments of the deduced amino acid sequences were generated using Clustal X program. Phylogenetic tree was constructed using the Neighbor-Joining method with MEGA 5 (Tamura et al., 2007). The three-dimensional structures were predicted by homology modeling using Swiss-Model server (http://www.expasy.org/swissmod/Swiss-Model.html) and the figures were visualized by DeepView software (Kiefer et al., 2008). The crystal structures of human and mouse orthologous genes were used as template models (PDB entries: 2BCK, 3OV6, 3LQZ, and 1I3R).

2.3. Polymorphism and selection pattern study

To analyze the polymorphism of MHC genes, total RNA was extracted from the intestines of eight individuals using Trizol reagent (Invitrogene). cDNA was then synthesized with random primer and M-MLV reverse system (Promega), and ORFs of MHC genes were amplified with Pfu polymerase (Transgene). An average of ten positive clones per each individual were sequenced.

Selection pattern was estimated using codon-based models implemented in the CODEML program from PAML v4 (Yang, 2007). Six models allowing for different levels of selection among sites were tested: M0, M1a, M2a, M3, M7, and M8 model. The neutral models (M0, M1a, and M7) against its nested models allowing for positive selection (M3, M2a, and M8) were compared using the likelihood ratio tests (LRT). The LRT is a comparison of twice the log-likelihood difference ($2\Delta lnL$) to a χ^2 distribution with degrees of freedom equal to the differences in the number of parameters between the corresponding models (M0 vs M3, M1a vs M2a, and M7 vs M8). In the models M2a and M8, the positively selected sites (posterior

Fig. 1 (*continued*)

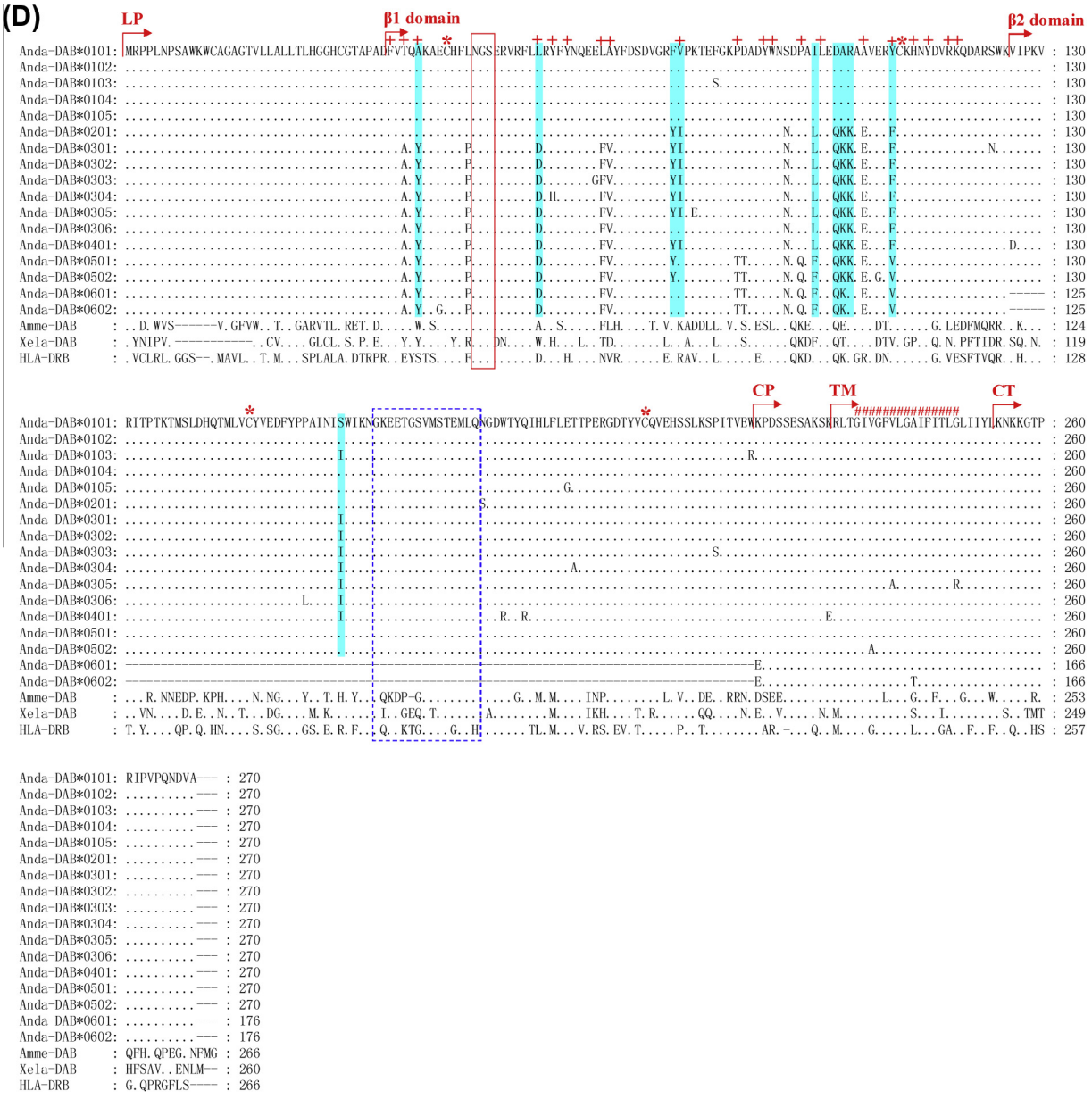

Fig. 1 (*continued*)

probabilities >0.95) were identified with Bayes Empirical Bayes (BEB) approach (Yang et al., 2005).

2.4. Expression profiles analyses

For *in vivo* stimulation experiment, two groups of six salamanders were intraperitoneally injected with 400 μl serum-free medium 199 containing ADRV ($1 \times 10^{6.5}$ $TCID_{50}$/ml) or 400 μl sterilized saline (0.9% Sodium Chloride) as control. On 12 days post-infection, the total RNA from five tissues including liver, spleen, kidney, intestine, and thymus were isolated. Quantitative real-time PCR was performed with SYBR green real-time PCR master mix reagents kit (ToYoBo). All primers used for PCR were list in Table 2. The amplifications were conducted on StepOne real-time PCR system (Applied Biosystems) and the cycle condition was as follows: 95 °C for 5 min, followed by 40 cycles of 95 °C for 15 s, and 60 °C for 1 min. The melting curve analysis of PCR products from 60 °C to 95 °C were performed after PCR. β-Actin was used as an endogenous control to normalize. All samples were tested in triplicate and the relative expression levels of MHC genes were determined with the comparative CT method (Zhu et al., 2013).

For *in vitro* stimulation experiment, CGST cells cultured in 25 cm^2 culture plates were incubated with ADRV at a multiplicity of infection (MOI) of 0.05, or with 250 μg/ml LPS (Sigma). The cells treated with FBS-free 199 were used as control in parallel. At various times (6, 12, 24, 48 and 72 h) post-infection, the cellular RNA were extracted for real-time PCR analysis as described above.

For tissue distribution analysis of MHC isoforms, three healthy salamanders were sampled to isolate total RNA from eight tissues including liver, spleen, kidney, thymus, intestine, heart, muscle and testis. PCR using primers for the complete ORF (Table 2) was performed, with conditions as follows: 95 °C for 5 min, 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s for 30 cycles, followed by 72 °C for 10 min.

For expression analysis of different isoforms after viral infection, RT-PCR was performed on total RNA exacted from spleen, thymus, and CGST cells infected with ADRV. The PCR products were subjected to agarose gel (1.5%) electrophoresis stained with

ethidium bromide. Densitometric analysis was performed using GeneTools 4.01 program (Syngene). The relative expression levels of MHC genes were calculated as relative to that of β-Actin in each sample using the formula: intensity of the target gene band/intensity of its corresponding β-Actin band (Chen et al., 2012).

2.5. Subcellular localization

To generate C-terminal fusion of MHC or the splice variants with Enhanced Green Fluorescent Protein (EGFP) or Red Fluorescent Protein (DsRed), their entire ORFs were cloned into pEGFP or pDsRed2 vector (Clontech, USA). The plasmids were confirmed by sequencing analysis. All primers used for constructions were list in Table 2. Transfection was performed according to the previous reports (Zhu et al., 2013). Typically, each well of EPC cells seeded in 6-well plates overnight was transfected with the mixture containing 4 μg of indicated plasmids and 8 μl of lipofectamine 2000 (Invitrogen) in 1 ml FBS-free 199 medium. 6 h later, the transfection mixture was replaced with 2 ml 10% FBS-containing 199 medium. At 24 h post-transfection, the cells were washed with PBS, fixed in 4% paraformaldehyde (PFA) for 30 min, and stained with Hoechst33342 (Sigma, USA) for 5 min. The cells were then visualized under a Leica DM IRB fluorescence microscope (objective 100×).

3. Results and discussion

3.1. Characteristics of gene sequences

Two types of MHC genes, namely, class I and class II genes, together with β2M, were identified from the Chinese giant salamander with the typical structure features of the MHC family.

Twenty-six unique MHC IA alleles were isolated from 8 individuals and designated as Anda-UAA*0101–Anda-UAA*0902 (GenBank Accession Nos.: KF611820–KF611845), including 24 full-length sequences and 2 truncated ones (Fig. 1A). The full-length transcripts encode proteins ranging in length from 349 to 351 amino acids. They possess all characteristic domains present

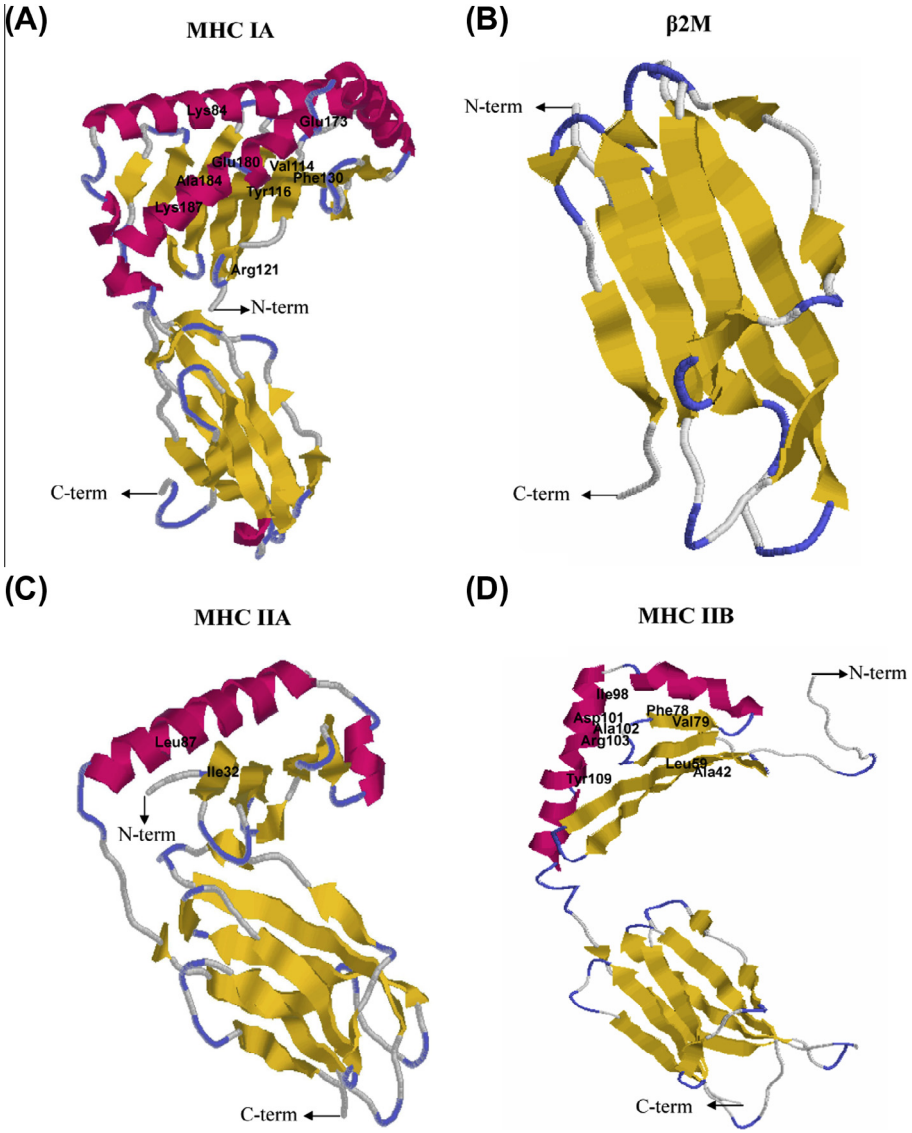

Fig. 2. Homology modeling of Anda-MHC genes. Ribbon cartoon diagram of MHC IA (A), β2M (B), MHC IIA (C), and MHC IIB (D) displaying the typical structural conformations of vertebrate MHC genes. α-helices are shown as helical ribbons (pink), and β-strands are represented as thick arrows (yellow). The respective polymorphic residues are indicated. Amino and carboxy termini are denoted as "N-term" and "C-term". (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

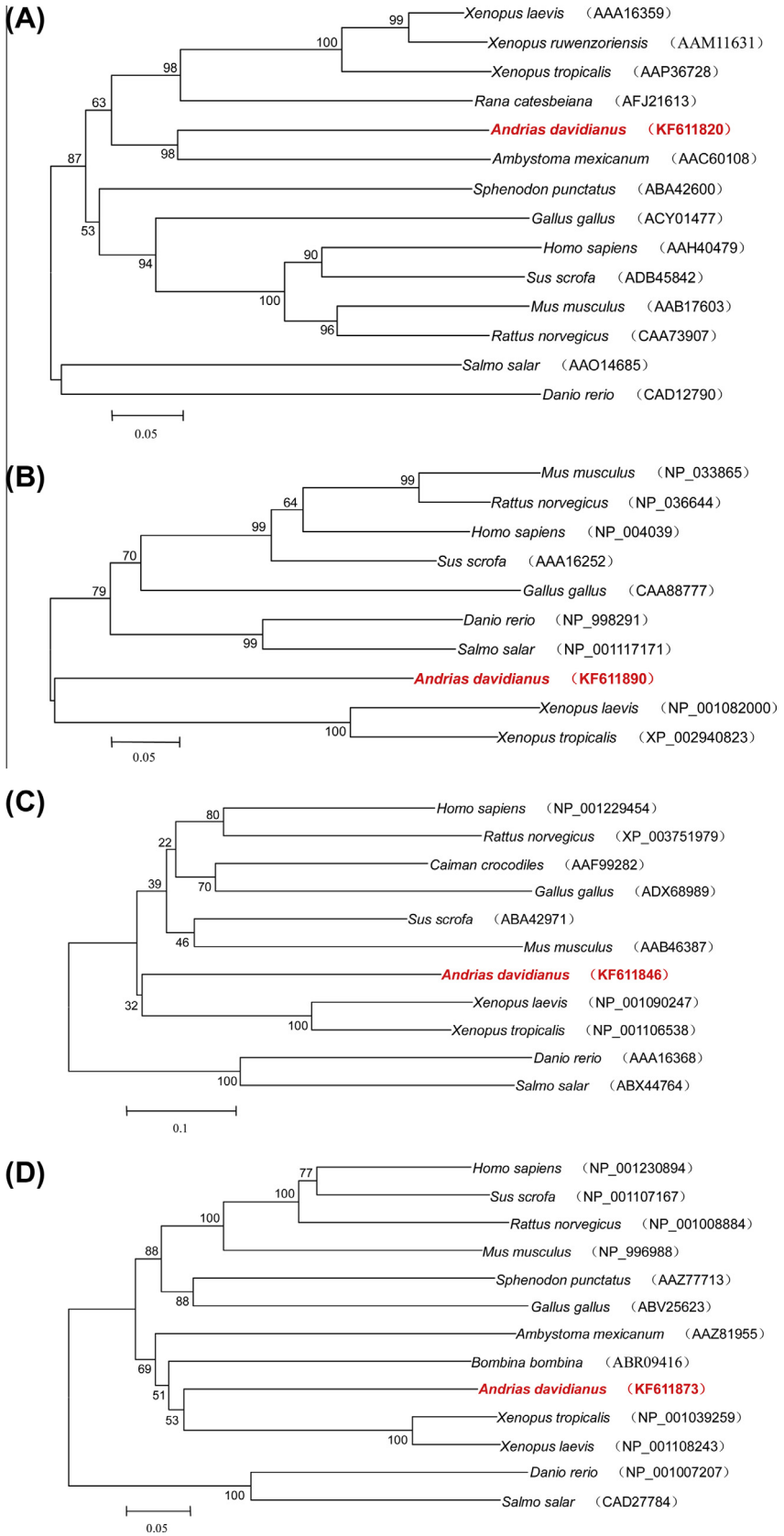

Fig. 3. Phylogenetic analysis of MHC IA (A), β2M (B), MHC IIA (C) and MHC IIB (D) in representative vertebrates. The trees are based on the peptide-binding domains and mature protein regions for MHC and β2M genes, respectively. Anda-MHC sequences are tightly clustered with those of other amphibians. The bootstrap confidence values shown at the nodes of the tree derived from 1000 replicates. GenBank accession numbers are indicated in parentheses after each sequence name.

in MHC IA, including a leader peptide, two highly variable domains (α1 and α2) involved in peptide binding, and a relatively conserved immunoglobulin (Ig)-like domain (α3), followed by connecting peptide/transmembrane domain/cytoplasmic domain (CP/TM/CT). One N-glycosylation site (N^{104}HT) is observed in the α1 domain, and four conserved cysteine residues (C^{118}/C^{181}/C^{220}/C^{275}) are found in the α2 and α3 domains, which are responsible for forming intrachain disulfide bonds. Furthermore, the residues of Y26, Y77, R102, T159, R162, W163, Y176, and Y188 corresponding to peptide-binding sites are strictly conserved among the alleles. Positions 242-248 in the α3 domain are also conserved in the 26 sequences, consistent with binding to CD8 (Devine et al., 2006).

Anda-β2M (GenBank Accession No: KF611890) encodes a mature protein of 98 amino acids with a 20-residue leader peptide (Fig. 1B). Two conserved cysteine residues (C^{45}/C^{100}) and a typical Ig and MHC motif (LDCKVAYH) are found in the mature protein. As in other β2M proteins, no putative glycosylation sites are observed in the Anda-β2M sequence.

Twenty-seven MHC IIA alleles were identified as Anda-DAA*0101–Anda-DAA*0503 (GenBank Accession Nos.: KF611846–KF611872) (Fig. 1C). Of these alleles, 25 possess the full-length sequences, which encode 249 amino acids, consisting of leader peptide, α1, α2, and CP/TM/CT domains. Two N-glycosylation sites (N^{99}NT/N^{139}VT) and two cysteine residues (C^{128}/C^{184}) are well conserved in the α1 and α2 domains of all sequences, in accordance with previous findings in other amphibians and mammals (Liu et al., 2002). Of the 18 amino acids involved in peptide binding, 14 are conserved across all alleles, and 4 show substitutions. The GxxxGxxGxxxG (x represents hydrophobic residues other than Gly) motif for interaction with MHC IIB chain (Xu et al., 2011), is observed in the transmembrane region.

Seventeen MHC IIB alleles were identified as Anda-DAB*0101–Anda-DAB*0602 (GenBank Accession Nos.: KF611873–KF611889) (Fig. 1D). Among them, 15 full-length sequences encode 270 amino acids, comprising leader peptide, β1, β2, and CP/TM/CT domains. One N-glycosylation site (N^{50}GC) is found in the β1 domain, as are four conserved cysteines, two in the β1 domain (C^{46}/C^{110}) and two in the β2 domain (C^{148}/C^{204}). As aligned, of 22 peptide-binding sites, 9 are conserved in the 17 sequences, and 8 show amino acid substitutions. Residues involved in CD4 binding (positions 136–180) (Yin et al., 2012), and those involved in MHC IIA chain binding (GxxGxxxGxxxxxxG) (Xu et al., 2011), also reside in the β2 domain and transmembrane region, respectively.

3.2. Molecular polymorphism and evolution mechanism

One to three alleles were identified from one single individual, indicating the presence of at least two loci in MHC genes. This finding is consistent with the tiger salamanders (Bos and DeWoody, 2005), but unlike the axolotl (Laurens et al., 2001). In general, MHC IA shows higher diversity compared to IIA and IIB, owing to the fact that class I genes undergo more rapid turnover than class II genes (Piontkivska and Nei, 2003). Pairwise comparison of the alleles reveals that 69.47% (66 of 95) variable amino acid sites occur in the α1 and α2 domains of MHC IA, 48.57% (17 of 35) variable sites in the α1 domain of MHC IIA, and 65.85% (27 of 41) variable sites in the β1 domain of MHC IIB (Fig. 1). The variability are concentrated in the α1/α2, α1, and β1 domains that correspond to the putative PBR, which allows MHC molecules to present a large array of antigen peptides to T cells. As the first terrestrial tetrapods, amphibians are exposed to a broader range of pathogens compared with ancestral aquatic fishes (Zhao et al., 2013). Therefore, the high level of MHC diversity could increase the ability of individual to combat pathogen infection.

Molecular evolution analysis using random sites models and LRTs reveals that the alternative models (M2a, M3, and M8) fit the data better than the nulls models (M1a, M0, and M7) ($\omega > 1$) (Table 1), indicating the presence of positive selection pressure

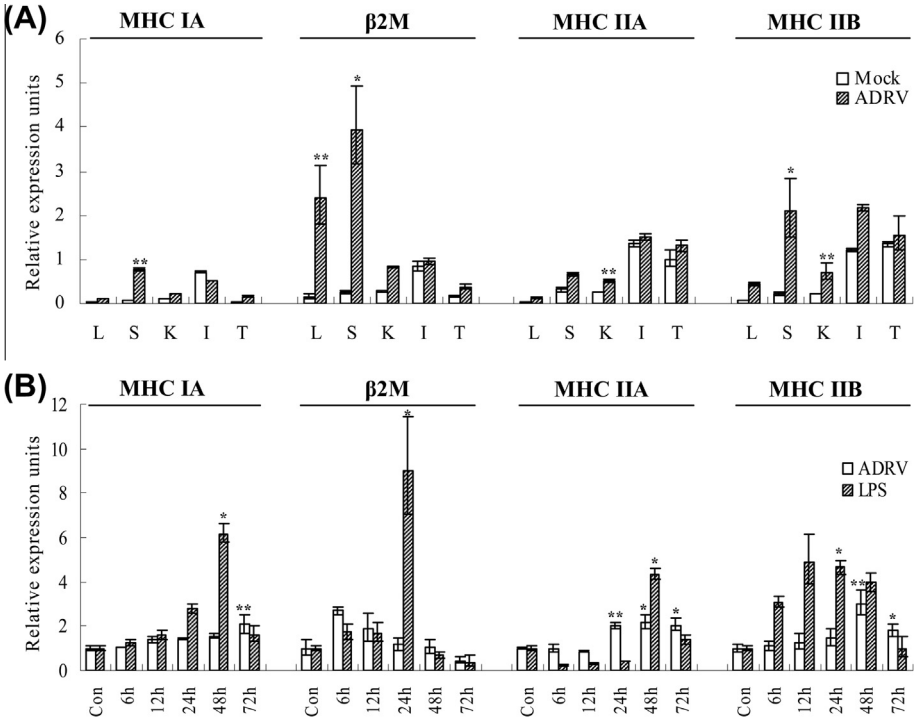

Fig. 4. Expression profiles of Anda-MHC after *in vivo* and *in vitro* stimulation. (A) Real-time PCR analysis of Anda-MHC transcriptional level in mock-infected and ADRV-infected tissues [liver (L), spleen (S), kidney (K), intestine (I), thymus (T)]. Values were expressed as a ratio relative to β-Actin levels in the same samples. (B) Real-time PCR analysis of Anda-MHC expression in CGST cells treated with ADRV or LPS for the indicated times. Values were normalized against that of the control cells (Con), and expressed as mean ± SD ($n = 3$). Statistical analysis was performed using Student' *t*-test. $^{*}P < 0.05$; $^{**}P < 0.005$.

on MHC sequences. Site-specific analysis in the case of model M2a and M8 identifies nine, four and ten condon sites that are under positive selection (posterior probability >0.95) in MHC IA, IIA, and IIB, respectively. Of these sites, three in MHC IA group, and eight in MHC IIB group are significantly selected, with the posterior probability >0.99. The majority of positively selected sites are found in the putative PBR domains: α2 domain (89%) of MHC IA, α1 domain (50%) of MHC IIA, and β1 domain (90%) in MHC IIB (Fig. 1). Moreover, most of the positively selected sites are identical with the highly variable sites. This result indicates that, as in mammalian MHC genes, the genetic diversity of Anda-MHC is maintained by positive Darwinian selection, probably as a result of pathogens.

3.3. Homology modeling and phylogenetic analysis

The predicted 3D homology model of MHC IA (Anda-UAA*0101) has classical organizations in two distinct domains. The N-terminal domain has an open antigen-binding groove, which consists of two adjacent α-helices on top of an antiparallel eight-stranded β-sheets. While the C-terminal domain has a typical Ig-like β-sandwich structure made of two large β-sheets, each containing three main strands and a few shorter strands (Fig. 2A). β2M is composed of two face-to-face β sheets, one of which has four strands, and the other has three (Fig. 2B). MHC IIA (Anda-DAA*0101) and IIB (Anda-DAB*0101) are also folded into two different domains: the N-terminal domains comprise one α-helix and four β-sheets, and the C-terminal domains comprise one six-stranded β-sandwich fold (Fig. 2C and D). This result shows that the overall structures of Anda-MHC genes are compatible with those of their mammalian counterparts.

Phylogenetic trees were constructed for MHC IA, β2M, MHC IIA, and MHC IIB, respectively. As excepted, Anda-MHC IA (Anda-UAA*0101) clusters first with the axolotl, both of which belong to the urodele amphibian; and then with *Xenopus*, apart from fish, reptile, bird, and mammalian counterparts (Fig. 3A). Similarly, the β2M sequences are subdivided into two major clusters, one with sequences from amphibians, and the other with sequences from fish, birds, and mammals (Fig. 3B). Both Anda-MHC IIA (Anda-DAA*0101) and IIB (Anda-DAB*0101) are tightly clustered with

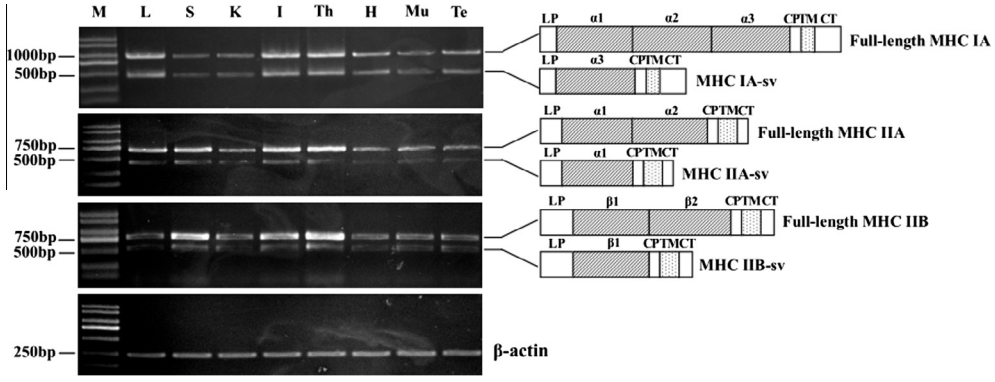

Fig. 5. Expression of different Anda-MHC isoforms in tissues. RT-PCR detection of MHC transcripts in eight tissues [liver (L), spleen (S), kidney (K), intestine (I), thymus (Th), heart (H), muscle (Mu), testis (Te)]. Two PCR products were observed: the bigger ones represent full-length MHC IA (1053 bp), MHC IIA (747 bp), and MHC IIB (810 bp), while the smaller ones represent MHC IA-sv (513 bp), MHC IIA-sv (465 bp), and MHC IIB-sv (468 bp), respectively. The small products are amplified by transcripts, in which α1/α2, α2, and β2 domains are spliced out, as shown in the diagrams. β-Actin was amplified as control. M, 2000-bp DNA markers.

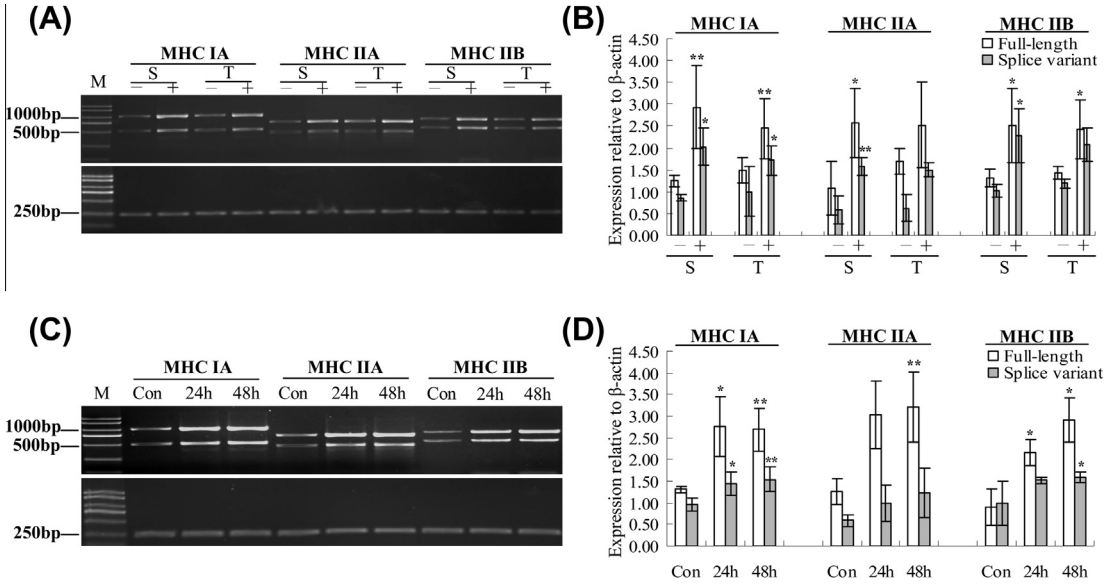

Fig. 6. Expression of different Anda-MHC isoforms upon virus stimulation. (A) RT-PCR analysis of MHC transcripts in mock-infected (−) and ADRV-infected (+) tissues [spleen (S), thymus (T)]. (B) Densitometric analysis of the above (A) data. Values were expressed as a ratio relative to β-Actin levels in the same samples. (C) RT-PCR analysis of MHC expression in mock-infected (Con) and ADRV-infected CGST cells for the indicated times. (D) Densitometric analysis of the above (C) data. Data represent means ± SD of triplicate experiments. M, 2000-bp DNA markers. $^{*}P < 0.05$; $^{**}P < 0.005$.

the sequences from other amphibians, as intermediate between those of fish and birds/reptiles/mammals (Fig. 3C and D). Together with the sequence analysis, these data further reinforce that Anda-MHC genes are orthologues of mammalian MHC.

3.4. Expression profiles after in vivo and in vitro stimulation

Challenge of the Chinese giant salamander with ADRV resulted in increases in the expression of MHC genes in spleen, kidney, liver, and thymus on 12 days post-infection (Fig. 4A). It was notable that the significant up-regulation was observed in spleen, the major peripheral lymphoid organ where B and T cells accumulate. The transcript levels of MHC IA, β2M, MHC IIA, and MHC IIB in spleen increased 11.4-fold, 14.9-fold, 2.1-fold, and 9.3-fold, respectively (Fig. 4A).

Infection of CGST cells with ADRV also induced up-regulations of MHC genes (Fig. 4B). MHC IA expression increased gradually from 6 to 72 h post-infection. β2M enhanced as early as 6 h, and

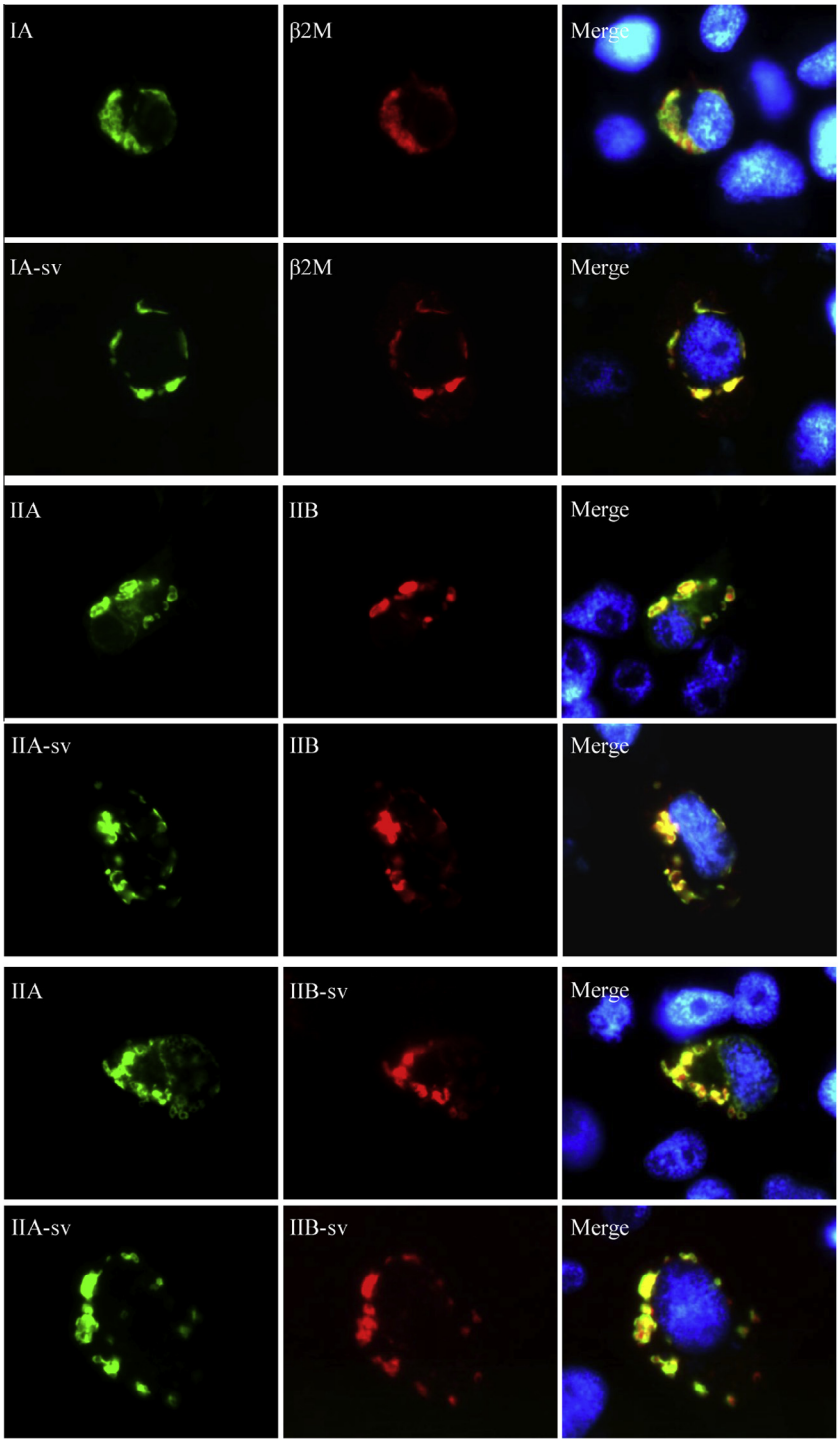

Fig. 7. Localization of the full-length and truncated Anda-MHC isoforms. The full-length MHC IA (green) was present on the cell surface and colocalized (Merge, yellow) with β2M (red) (row 1). Similarly, MHC IA-sv (green) could also colocalized with β2M (red) (row 2). MHC IIA (green) and MHC IIA-sv (green) were found on the plasma membrane, which showed extensive colocalization with MHC IIB (red) and MHC IIB-sv (red) (row 3-6). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

then sustained decreased from 12 to 72 h. MHC IIA and IIB remained at a low level during the first 12 h, and then increased to the peak at 48 h, followed by a slight decrease at 72 h. Moreover, similar inductions of MHC genes were observed in the cells treated with LPS. The expression of MHC IA and β2M was stable from 0 to 12 h, and reached the highest levels at 48 and 24 h, respectively. MHC IIA decreased from 6 to 24 h, but significantly enhanced at 48 h. MHC IIB increased as early as 6 h, reached a peak at 12 h, and then returned to a level similar to control at 72 h (Fig. 4B). These data show the potential roles of Anda-MHC genes in antiviral defense, and further studies are required to elucidate their precise function in the immune system.

3.5. Alternatively splice isoforms of Anda-MHC genes

Sequencing MHC alleles revealed that there were different isoforms: full-length transcripts and truncated splice variants, termed as MHC IA-sv, MHC IIA-sv, and MHC IIB-sv, respectively. Despite the fact that the extracellular domains (α1/α2, α2, and β2) were completely spliced out, the ORFs of these splice variants were not changed. Aside from one MHC IIB alternative splicing transcript missing the β1 domain (Bulut et al., 2008), to our knowledge, such transcripts have rarely been described in amphibians.

The transcriptional pattern assessment revealed that two transcripts were present for MHC IA, IIA, and IIB, corresponding to the full-length and truncated ones. The two transcripts were generally expressed in various tissues, and the full-length isoform was predominant (Fig. 5). Moreover, they both showed increase in mRNA expression upon stimulation with ADRV (Fig. 6), suggesting the coexistence of different MHC isoforms with similar expression profiles in the Chinese giant salamander.

The subcellular localization of MHC isoforms was further characterized. As shown in Fig. 7, MHC IA was present on the cell surface, which generated a characteristic rim pattern. Although the α1 and α2 domains were absent, MHC IA-sv was expressed on the cell surface and colocalized with β2M, owing to it possesses α3 and transmembrane sequences. Similarly, MHC IIA, IIB, and their variants were found on the plasma membrane in a punctuate vesicular pattern. A high colocalization was evident between the full-length forms of MHC IIA and IIB, as well as between the truncated transcripts. MHC IIA-sv and MHC IIB-sv, having α1/β1 and transmembrane domains, which could explain this phenomenon.

The α1 and α2 domains of MHC IA are responsible for the binding and presentation of foreign peptides to the TCR. The β2 domain of MHC IIB is associated with α2 domain of MHC IIA and interacts with CD4 molecules on T cells. It thus seems unlikely that these splice variants directly participate in the antigen presentation. Alternative splicing of mRNA transcripts is increasingly regarded as an extra mechanism to increase MHC diversity (Belicha-Villanueva et al., 2010). Up to 70% of MHC transcripts have been reported to undergo alternative splicing in different species. Alternative splicing can generate truncated non-functional proteins or those functioning as dominant negative mutant (Bulut et al., 2008). Indeed, a recent study has provided evidence that MHC IA splice variant acts as a negative modulator for antigen presentation (Dai et al., 2012). Therefore, further work that explores the functional significance of MHC splice variants is needed to assess whether they are translated into proteins *in vivo*, and whether they are involved in the regulation of immune responses.

4. Conclusion

We report the first characterization of MHC genes in the ancient and endangered Chinese giant salamander. Anda-MHC is polygenic, containing different class I and class II genes, and polymorphic, having multiple alleles of each gene. And the presence of distinct splice variants with identical distribution further highlights the genetic diversity of Anda-MHC. Enhanced expression of these genes in response to ranavirus suggests their roles in immune defense against amphibian pathogens. Thus, the data will be useful in further characterizing the adaptive genetic variation, splice variants and immunity of the MHC among amphibian lineages.

Acknowledgements

This work was supported by grants from the National Basic Research Program of China (2010CB126303), and the National Natural Science Foundation of China (31372551, 31202028, 31072231, 31072239).

References

Belicha-Villanueva, A., Blickwedehl, J., McEvoy, S., Golding, M., Gollnick, S.O., Bangia, N., 2010. What is the role of alternate splicing in antigen presentation by major histocompatibility complex class I molecules? Immunol. Res. 46, 32–44.
Bulut, Z., McCormick, C.R., Bos, D.H., DeWoody, J.A., 2008. Polymorphism of alternative splicing of major histocompatibility complex transcripts in wild tiger salamanders. J. Mol. Evol. 67, 68–75.
Bos, D.H., DeWoody, J.A., 2005. Molecular characterization of major histocompatibility complex class II alleles in wild tiger salamanders (*Ambystoma tigrinum*). Immunogenetics 57, 775–781.
Chen, Z.Y., Gui, J.F., Gao, X.C., Pei, C., Hong, Y.J., Zhang, Q.Y., 2013. Genome architecture changes and major gene variations of *Andrias davidianus* ranavirus (ADRV). Vet. Res. 44, 101.
Chen, Z.Y., Lei, X.Y., Zhang, Q.Y., 2012. The antiviral defense mechanisms in mandarin fish induced by DNA vaccination against a rhabdovirus. Vet. Microbiol. 157, 264–275.
CITES (The Convention on International Trade in Endangered Species of Wild Fauna and Flora), 2008. Appendices I, II and III. <http://www.cites.org/eng/app/appendices.shtml/> valid from 12 February 2008.
Dai, Z.X., Zhang, G.H., Zhang, X.H., Xia, H.J., Li, S.Y., Zheng, Y.T., 2012. The beta2-microglobulin-free heterodimerization of rhesus monkey MHC class I A with its normally spliced variant reduces the ubiquitin-dependent degradation of MHC class I A. J. Immunol. 188, 2285–2296.
Devine, L., Thakral, D., Nag, S., Dobbins, J., Hodsdon, M.E., Kavathas, P.B., 2006. Mapping the binding site on CD8 beta for MHC class I reveals mutants with enhanced binding. J. Immunol. 177, 3930–3938.
Dong, W., Zhang, X., Yang, C., An, J., Qin, J., Song, F., Zeng, W., 2011. Iridovirus infection in Chinese giant salamanders, China. Emerg. Infect. Dis. 17, 2388–2389.
Gao, K.Q., Shubin, N.H., 2003. Earliest known crown-group salamanders. Nature 422, 424–428.
Geng, Y., Wang, K.Y., Zhou, Z.Y., Li, C.W., Wang, J., He, M., Yin, Z.Q., Lai, W.M., 2011. First report of a ranavirus associated with morbidity and mortality in farmed Chinese giant salamanders (*Andrias davidianus*). J. Comp. Path. 145, 95–102.
Gui, J.F., Zhu, Z.Y., 2012. Molecular basis and genetic improvement of economically important traits in aquaculture animals. Chin. Sci. Bull. 57, 1751–1760.
Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., Schwede, T., 2008. The SWISS-MODEL repository and associated resources. Nucleic Acids Res. 37, 387–392.
Laurens, V., Chapusot, C., del Rosario Ordonez, M., Bentrari, F., Padros, M.R., Tournefier, A., 2001. Axolotl MHC class II β chain: predominance of one allele and alternative splicing of the β1 domain. Eur. J. Immunol. 31, 506–515.
Liu, Y., Kasahara, M., Rumfelt, L.L., Flajnik, M.F., 2002. Xenopus class II A genes: studies of genetics, polymorphism, and expression. Dev. Comp. Immunol. 26, 735–750.
Neefjes, J., Jongsma, M.L., Paul, P., Bakke, O., 2011. Towards a systems understanding of MHC class I and MHC class II antigen presentation. Nat. Rev. Immunol. 11, 823–836.
Ohta, Y., Goetz, W., Hossain, M.Z., Nonaka, M., Flajnik, M.F., 2006. Ancestral organization of the MHC revealed in the amphibian Xenopus. J. Immunol. 176, 3674–3685.
Piertney, S.B., Oliver, M.K., 2006. The evolutionary ecology of the major histocompatibility complex. Heredity 96, 7–21.
Piontkivska, H., Nei, M., 2003. Birth-and-death evolution in primate MHC class I genes: divergence time estimates. Mol. Biol. Evol. 20, 601–609.
Robert, J., Cohen, N., 2011. The genus *Xenopus* as a multispecies model for evolutionary and comparative immunobiology of the 21st century. Dev. Comp. Immunol. 35, 916–923.
Savage, A.E., Zamudio, K.R., 2011. MHC genotypes associate with resistance to a frog-killing fungus. Proc. Natl. Acad. Sci. USA 108, 16705–16710.
Sommer, S., 2005. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Front. Zool. 2, 16.
Spurgin, L.G., Richardson, D.S., 2010. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. Proc. Biol. Sci. 277, 979–988.

Sutton, J.T., Nakagawa, S., Robertson, B.C., Jamieson, I.G., 2011. Disentangling the roles of natural selection and genetic drift in shaping variation at MHC immunity genes. Mol. Ecol. 20, 4408–4420.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.

Teacher, A.G., Garner, T.W., 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.

Wang, L., Yang, H., Li, F., Zhang, Y., Yang, Z., Li, Y., Liu, X., 2013. Molecular characterization, tissue distribution and functional analysis of macrophage migration inhibitory factor protein (MIF) in Chinese giant salamanders *Andrias davidianus*. Dev. Comp. Immunol. 39, 161–168.

Xu, T., Sun, Y., Shi, G., Cheng, Y., Wang, R., 2011. Characterization of the major histocompatibility complex class II genes in miiuy croaker. PLoS ONE 6, e23823.

Yang, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591.

Yin, Y., Wang, X.X., Mariuzza, R.A., 2012. Crystal structure of a complete ternary complex of T-cell receptor, peptide-MHC, and CD4. Proc. Natl. Acad. Sci. USA 109, 5405–5410.

Yang, Z., Wong, W.S., Nielsen, R., 2005. Bayes empirical bayes inference of amino acid sites under positive selection. Mol. Biol. Evol. 22, 1107–1118.

Zhang, P., Chen, Y.Q., Liu, Y.F., Zhou, H., Qu, L.H., 2003. The complete mitochondrial genome of the Chinese giant salamander, *Andrias davidianus* (Amphibia: Caudata). Gene 311, 93–98.

Zhang, Q.Y., Gui, J.F., 2012. Atlas of Aquatic Viruses and Viral Diseases. Science Press, Beijing, pp. 251–253.

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

Zhu, R., Wang, J., Lei, X.Y., Gui, J.F., Zhang, Q.Y., 2013. Evidence for *Paralichthys olivaceus* IFITM1 antiviral effect by impeding viral entry into target cells. Fish Shellfish Immunol. 35, 918–926.

Zhu, R., Zhang, Y.B., Zhang, Q.Y., Gui, J.F., 2008. Functional domains and the antiviral effect of the dsRNA-dependent protein kinase PKR from *Paralichthys olivaceus*. J. Virol. 82, 6889–6901.