ONLINE RESOURCES

Extremely high major histocompatibility complex class IIb gene intron 2 variation and population structure in Chinese alligator

CHUANPENG NIE[1,2], YANYAN LI[2], JUAN ZHAO[1] and XIAOBING WU[1]*

[1]*College of Life Sciences, Anhui Normal University, Wuhu 241000, People's Republic of China*
[2]*College of Life Sciences, Fuyang Teachers College, Fuyang 236041, People's Republic of China*

[Nie C., Li Y., Zhao J. and Wu X. 2012 Extremely high major histocompatibility complex class IIb gene intron 2 variation and population structure in Chinese alligator. *J. Genet*. **91**, e86–e91. Online only: http://www.ias.ac.in/jgenet/OnlineResources/91/e86.pdf]

Introduction

Identifying the processes that maintain genetic diversity within and among populations is a central goal of modern evolutionary genetics. So far, the studies on genetic diversity have mainly focussed on neutral DNA markers, such as mtDNA and microsatellites (Xu *et al.* 2010). While these markers are very informative in molecular clocks, to examine dispersal patterns of individuals (gene flow) and to classify them by relatedness and paternity analyses. The variation at neutral loci cannot provide direct information on selective processes involving the interaction of individuals with their environment or on the capacity for future adaptive changes (Sommer 2005). With the developments in molecular biology, researchers can directly examine selection at genes that underlie functional traits. Therefore, adaptive non-neutral markers have become especially valuable (Koskinen *et al.* 2002). One important example comes from the genes of the major histocompatibility complex (MHC). These genes are found in all jawed vertebrates and play a critical role in an organism's immune response (Karaiskou *et al.* 2010). High levels of variation in MHC alleles are believed to confer the ability to recognize a wide range of antigens and thus increase resistance to a greater number of environmental pathogens.

Chinese alligator (*Alligator sinensis*) is one of the 23 critically endangered crocodile species in the world. The wild population has been close to extinct during the past decades due to habitat loss and illegal hunting. The investigation indicated that the number of wild Chinese alligators has decreased from 500 individuals in the 1980s to currently less than 120 or 150 individuals (Thorbjarnarson *et al.* 2002). To prevent extinction of the species, Chinese government has performed a series of efforts including officially listing Chinese alligators as the first class protected animal, and established two breeding farms in Xuanzhou (Anhui Research Center of Chinese Alligator Reproduction (ARCCAR)) and Changxing (Zhejiang province) in the early 1980s.

Genetic variability in this relict species is obviously essential for the genetic management of the captive alligator; thus much research on this species has been performed. For neutral DNA, there are RAPD (Wu *et al.* 2002), mtDNA D-loop sequencing (Wang *et al.* 2003), AFLP (Wang *et al.* 2006) and microsatellite (Jing *et al.* 2009) studies. For nonneutral markers, there are two reports based on MHC genes. Shi *et al.* (2004) analysed 166 bp of the exon 2 fragments of MHC class IIb genes in three Chinese alligators from ARCCAR, and Liu *et al.* (2007) searched 260 bp of the exon 3 partial sequences in 14 Chinese alligators. They found that there was a higher polymorphism of MHC class IIb genes in Chinese alligators.

To understand better the polymorphism of MHC and population structure of this species, more individuals and MHC IIb gene intron 2 partial sequences were analysed in this study. Our specific goals were to (i) investigate the variation of the MHC class IIb gene; and (ii) to provide more detailed genetic information for conservation and management strategies for this endangered population.

Materials and methods

Samples and DNA extraction

A total of 21 individuals were sampled (table 1). Sample collection, transportation and storage are same as previously described (Wu *et al.* 2002). DNA extraction followed a conventional phenol/chloroform method (Sambrook and Russel

*For correspondence. E-mail: wuxb@mail.ahnu.edu.cn.

Keywords. MHC class IIb gene intron 2; variation; Chinese alligator; endangered.

Table 1. Samples for MHC sequence analysis.

Population	Source	Number	Type
Wild population (WP)	Xuanzhou, Anhui	7	Blood
Captive population in Xuanzhou (XZ)	ARCCAR	7	Blood
Captive population in Changxing (CX)	Changxing, Zhejiang	7	Blood

2001), and genomic DNA dissolved with double distilled water. The extracted DNA was examined on 1% agarose gels stained with 10 mg/mL ethidium bromide, stored at −20°C for further use.

PCR procedure

The MHC were amplified using the primers designed by the sequence of *Caiman crocodilus* (AF256651, AF256652 and AF277661) through Primer premier 5.0. Amplification reactions (30 μL) containing 30 ng genome DNA, 3 μL 10× PCR buffer, 2 μL 25 mM $MgCl_2$, 1 μL 25 mM dNTP (Sangon, Shanghai, China), 2 μL 10 mM primer (Genscript, Nanjing, China), 1 U *Taq* DNA polymerase (Sangon). PCR was performed in a Mastercycler gradient (Eppendorf, Hamburg, Germany). Initial denaturation of 95°C for 5 min was followed by 35 cycles of 94°C for 30 s, 47°C for 30 s, and 72°C for 1 min. A final extension of 72°C for 10 min was incorporated, followed by cooling to 4°C until recovery of the samples. PCR products were visualized using UV light and separated on a 1.5% agarose gel following staining with 10 mg/mL ethidium bromide.

Cloning and sequencing

Since MHC complex is a group of genes, one individual may include many different MHC sequences; in order to sequence accurately the PCR products, the PCR products should be cloned. PCR products purified by AxyPrep^{TM} PCR Cleanup kit (AXYGEN Biotechnology, Hangzhou, China) were cloned into pMD18-T vectors (Takara, Dalian, China) and transformed into DH5a competent cells following the manufacturer's instruction. After incubation at 37°C on LB agar–ampicillin plates overnight, at least 15 clones per individual were checked for an insert by PCR. Between 8 and 11 insert-positive clones per individual were sequenced using the M13 forward and reverse primer (Genscript, Nanjing, China).

Statistical analysis

After splicing with the ContigExpress software and correcting using peak Chromas, all sequences were compared (using the algorithm BLASTn) with those available via the NCBI database (http://www.ncbi.nlm.nih.gov) (Altschul *et al.* 1997). Nucleotide multiple alignments were performed with ClustalX (Thompson *et al.* 1997).

Nucleotide polymorphism and diversity were calculated using DnaSP v5 (Librado and Rozas 2009). Analysis of molecular variance (AMOVA) was performed using Arlequin v3.1 (Exeoffier *et al.* 2005) to assess genotypic variations across all the populations studied. Relationships of populations were estimated using the neighbour-joining (NJ) method on the basis of Nei's unbiased genetic diversity (Tamura *et al.* 2007).

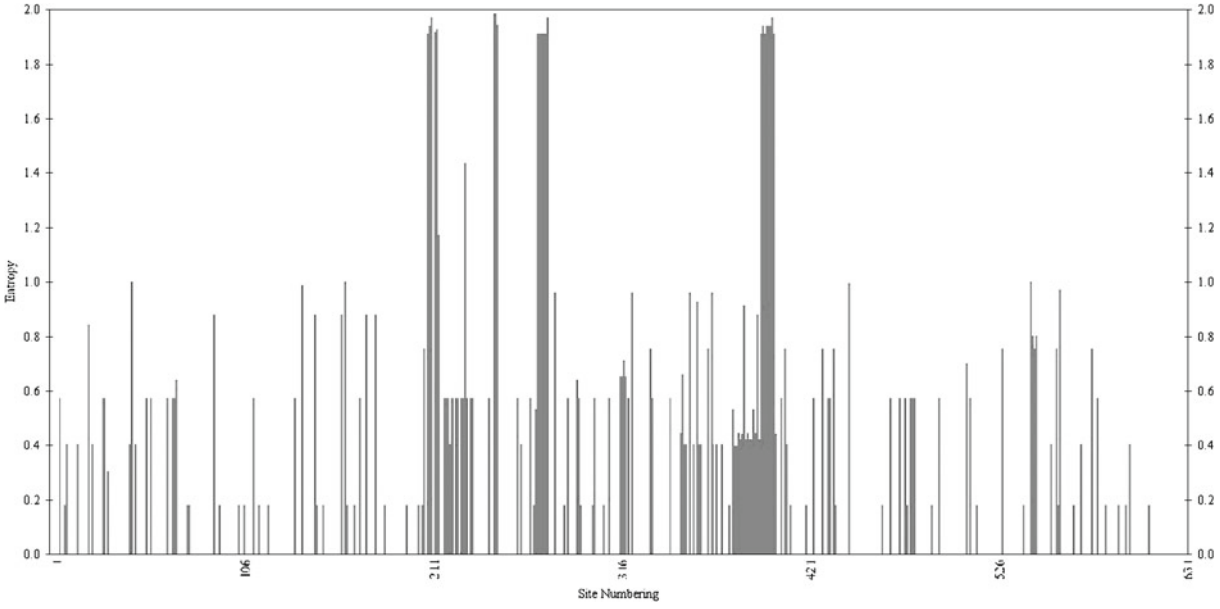

Figure 1. The variation index at per site numbering entropy measures the variation index at per base site. $H_i = -\text{sum}\,[P_i \cdot \log 2\,(P_i)]$; H_i, 0-2; P_i, the substitution frequencies of different nucleotides per site within the allelic sequences sampled.

Computation of DNA entropy was performed using the program DAMBE (Xia and Xie 2001). To see how the alleles were related, we reconstructed a nucleotide phylogeny by the NJ method with Kimura's two-parameter model, using MEGA v4 (Tamura *et al.* 2007). Bootstrap values were calculated by 1000 replicates, and branches corresponding to partitions reproduced in less than 50% of bootstrap replicates were collapsed. Network relationship of the alleles was estimated by TCS1.21 (Clement *et al.* 2000).

Results and discussion

Nucleotide and amino acid variation

An average of 10 clones per individual were sequenced, the same sequence which obtained from at least two clones or had different length in the same individual would be accepted in the subsequent data analysis. This strategy should allow for the removal of potential artificial polymorphisms because of the use of *Taq* DNA polymerase by which misincorporation could occasionally occur. By sequencing, sequences of three length types were obtained in Chinese alligators. Similar search of the sequences using the NCBI database showed 86%–94% similarity with *Caiman crocodilus*.

In total, 206 clones from 21 Chinese alligators were sequenced. The sequencing results of 191 clones were the target fragments. We obtained 20 sequences in seven individuals from wild population (WP), 15 in seven individuals from captive population in Xuanzhou (XZ), 18 in seven individuals from captive population in Changxing (CX): thus, there were 53 sequences for subsequent research. For simplicity, different sequences are tentatively referred to as alleles even though they may be from different loci (Miller and Lambert 2004). In 53 sequences, we indentified 43 alleles from 21 individuals. These 43 sequences have been deposited in GenBank (accession nos. JQ048623–JQ048665).

According to the intron–exon boundary GT–AG rules, and by aligning with the sequence of *C. crocodilus* (AF256651, AF256652 and AF277661), 37 alleles were identified and three length types in intron 2 partial sequences, namely 596 bp, 611 bp, 626 bp, suggesting that Chinese alligator has at least three class IIb loci (Kiemnec-Tyburczy *et al.* 2010).

In previous studies, 38 variable sites among 10 nucleotide sequences and 23 variable sites among amino acid sequences were detected in MHC IIb gene exon 2 partial sequences, which were 166-bp long (except one 160 bp) (Shi *et al.* 2004); 34 sequence haplotypes of exon 3 were detected in the sampled Chinese alligators, and 83 polymorphic (variable) sites were found within MHC IIb gene exon 3 partial sequences (260 bp) (Liu *et al.* 2007). About the genetic diversity of Chinese alligator, the relatively good molecular marker, is mtDNA control region and microsatellites. Huang and Wang (2004) amplified 5′end of control region, they distinguished 10 haplotypes from the obtained sequences. Studies based on microsatellite markers (Jing *et al.* 2009) revealed a limited allelic polymorphism (2–7 alleles per locus) in the Chinese alligator.

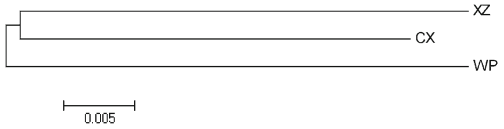

Figure 2. The cluster of three populations in Chinese alligator.

In this study, 183 polymorphic (gaps were included) sites in 634 sites were obtained. Relatively good genetic diversity was seen, suggesting that the research based on MHC IIb gene intron 2 was more powerful method for genetic diversity of Chinese alligator. The variation index at per site numbering are given (figure 1), this analysis computed an entropy-based measure of variability over sites and plotted the variability along the sequences. Nucleotide diversity in WP was higher than that in two captive populations; in two captive populations, nucleotide diversity in XZ was higher than that of CX. About haplotype diversity, WP was highest, followed by CX and XZ (table 2).

Population structure

The genetic distances between the WP and the two captive populations were higher than that between the two captive populations. The NJ tree (figure 2) revealed that populations XZ and CX were genetically closely related, while population WP was the most distant from the others.

The difference among three populations was not significant ($P > 0.05$) in AMOVA. The degree of differentiation within a population accounted for 97.10% of variation, while only 2.90% of the variation was among populations (table 3), suggesting that the genetic differences mainly

Table 2. Analysis of genetic diversity of different populations.

Population	Number of specimens	Average number of nucleotide differences (k)	Haplotype diversity (H_d)	Nucleotide diversity (π)	Number of haplotypes (h)
XZ	7	36.162	0.9905	0.06007	14
WP	7	37.811	0.9789	0.06281	17
CX	7	29.438	0.9869	0.04890	16
Total	21	33.845	0.9746	0.05622	37

Table 3. Hierarchical analysis of molecular variance (AMOVA) within/among populations.

Source of variation	Degrees of freedom (d.f.)	Sum of squares	Variance component	Percentage of variation	Fixation index (F_{ST})	P value
Among populations	2	19.498	0.54416 V_a	2.90		
Within populations	50	964.882	18.20933 V_b	97.10		
Total	52	984.381	18.75349		0.02902	0.80743 ± 0.01221

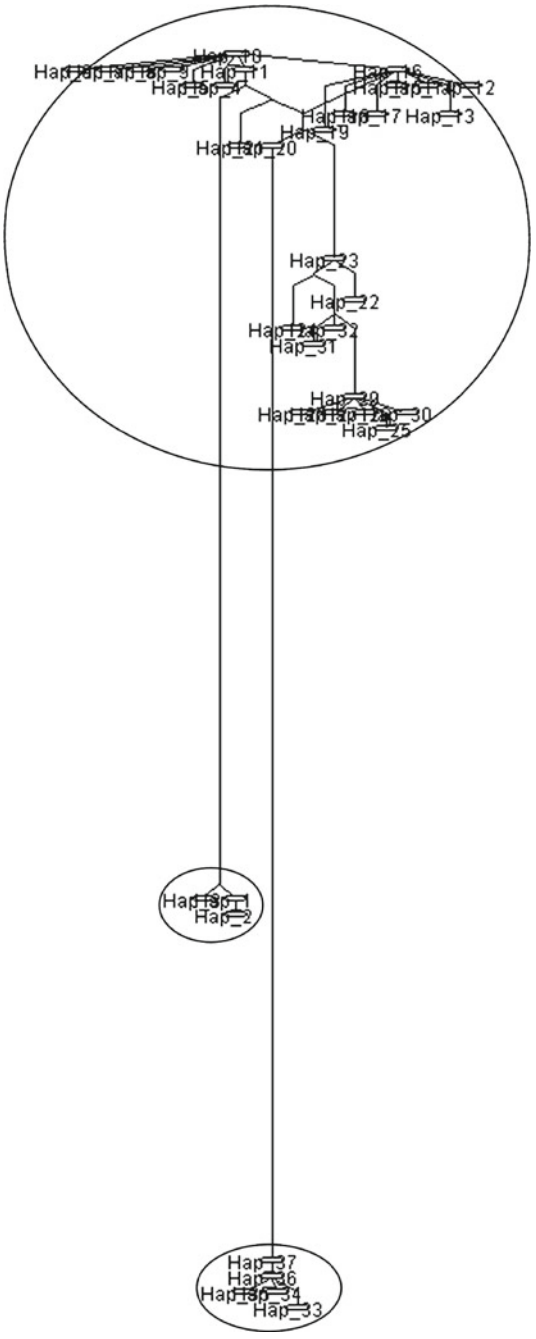

Figure 3. Network relationship estimated by TCS 1.21.

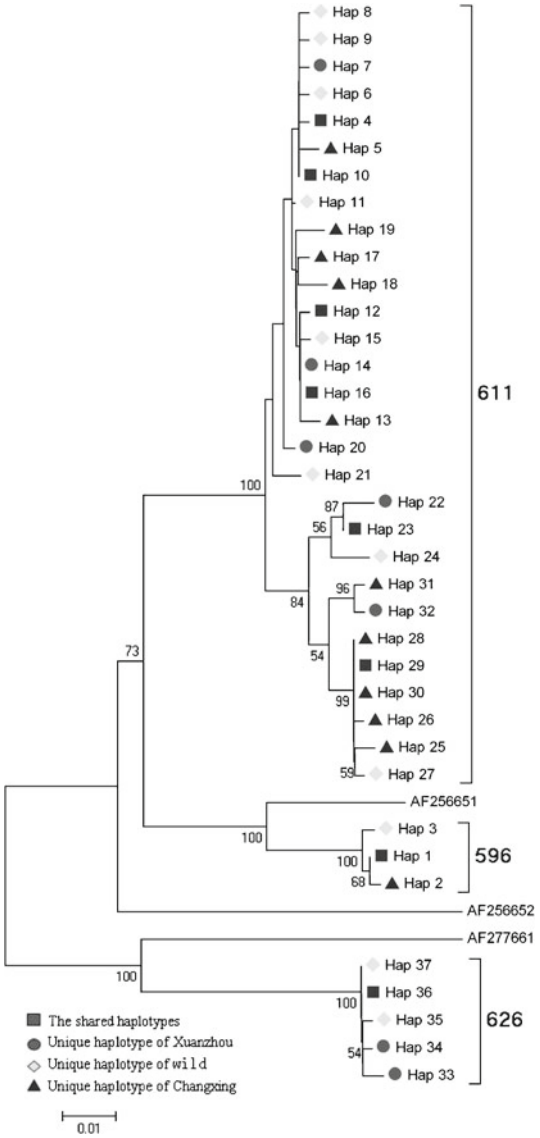

Figure 4. NJ phylogeny of all alleles. A neighbour-joining phylogeny based on genetic distance of the Kimura two-parameter model and bootstrap resampling 1000 times. The tree also includes three alleles from *Caiman crocodilus* (AF256651, AF256652 and AF277661) (bootstrap values >50% are indicated above the branches).

occurred within populations. Population fixation index (F_{ST}) represents the genetic differentiation among populations; the larger the F_{ST} values, the higher the degree of differentiation among populations (López-Fanjul *et al.* 2007). In this study, the result indicated a lack of isolation among three populations and there were no differentiation, the same as found by Wu *et al.* (2002) and Liu *et al.* (2007).

Usually, gene flow values (*Nm*) less than 1 indicate a limited group of gene flow, *Nm* values higher than 1 may represent large levels of gene flow and genetic exchange now or in the past. In this study, *Nm* based on intron 2 partial sequence was 13.31, suggesting that intergroup gene flow may have occurred.

Alleles phylogenetic analysis

Evolution of MHC genes in Chinese alligator is still poorly understood, Shi *et al.* (2004) and Liu *et al.* (2007) had conducted a preliminary study on MHC IIb genes. Their study was based on partial sequences of exons.

Two-component system (TCS) network are constructed from entire sequences. Three parts are shown in the network relationship diagram (figure 3), as exactly corresponded to the three length types. In the nucleotide phylogenetic NJ tree (figure 4), the 596 bp and 626 bp alleles groups were separated from 611 bp alleles group. We also found trans-species polymorphism within the exon 3 regions, three alleles of *C. crocodilus* were scattered in the alleles of Chinese alligator, suggesting that exon 3 had high variation in or between species. We think the same study can be carried out in other endangered animals.

Suggestion for conservation

In this study, altogether three types of MHC sequences were found in all the individuals, however, only one or two types of different sequences could be obtained in one individual. In addition, the sequences from the same and individual were mostly same and the sequences from the different individual were mostly different. This result showed that there was large difference between individuals but little difference within individuals. To increase the genetic diversity of Chinese alligator, we should increase the number of the individual, especially wild individuals.

So we suggest that exchange between the two captive populations (XZ and CX) would be an important step to increase the level of genetic diversity in both populations. Secondly, the reintroduction of breeding population will be an important part of protecting Chinese alligator. To increase the number of individuals of wild populations and genetic polymorphisms, we can let some individuals with relatively high diversity reintroduced into the wild. Lastly, the results showed the wild population provided major valuable information for population rejuvenation, so it should be strictly protected; at the same time, we must strengthen the education about alligator conservation.

Acknowledgements

The authors are grateful to several anonymous reviewers for their helpful suggestions on revising the manuscript. This work was supported by the National Natural Science Foundation of China (NSFC, no. 30770312), the Excellent Creative Research Team of Animal Biology in Anhui Normal University and the Key Laboratory of Biotic Environment and Ecological Safety in Anhui Province, Zoology Practical Skills Training Center in Fuyang Teachers College. We are also grateful to the assistance of Anhui Research Center for Chinese Alligator Reproduction in this study.

References

Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Zhang Z., Miller W. *et al.* 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. **25**, 3389–3402.

Clement M., Posada D. and Crandall K. A. 2000 TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9**, 1657–1660.

Exeoffier L., Laval G. and Sehneider S. 2005 Arlequin ver 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **1**, 47–50.

Huang L. and Wang Y. Q. 2004 SSR polymorphism of *Alligator sinensis* and conservation strategy of genetic diversity. *Acta Genet. Sin*. **31**, 143–150.

Jing W., Wang X. L., Lan H. and Fang S. G. 2009 Eleven novel microsatellite markers for the Chinese alligator (*Alligator sinensis*). *Conserv. Genet.* **10**, 543–546.

Karaiskou N., Moran P., Georgitsakis G., Abatzopoulos T. J. and Triantafyllidis A. 2010 High allelic variation of MHC class II alpha antigen and the role of selection in wild and cultured *Sparus aurata* populations. *Hydrobiologia* **638**, 11–20.

Kiemnec-Tyburczy K. M., Richmond J. Q., Savage A. E. and Zamudio K. R. 2010 Selection., trans-species polymorphism., and locus identification of major histocompatibility complex class IIβ alleles of New World ranid frogs. *Immunogenetics* **62**, 741–751.

Koskinen M. T., Haugen T. O. and Primmer C. R. 2002 Contemporary fisherian life-history evolution in small salmonid populations. *Nature* **419**, 826–830.

Librado P. and Rozas J. 2009 DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452.

Liu H., Wu X. B., Yan P. and Jiang Z. G. 2007 Polymorphism of exon 3 of MHC class IIB gene in Chinese alligator (*Alligator sinensis*). *J. Genet. Genomics* **34**, 918–929.

López-Fanjul C., Fernández A. and Toro M. A. 2007 The effect of dominance on the use of the QST-FST contrast to detect natural selection on quantitative traits. *Genetics*. **176**, 725–727.

Miller H. C. and Lambert D. M. 2004 Genetic drift outweighs balancing selection in shaping post-bottleneck major histocompatibility complex variation in New Zealand robins (*Petroicidae*). *Mol. Ecol.* **13**, 3709–3721.

Sambrook J. and Russel D. W. 2001 *Molecular cloning: a laboratory manual*, vol. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.

Shi Y., Wu X. B., Yan P. and Chen B. H. 2004 Cloning and sequences analysis of the second exon of MHC class IIB genes in Chinese alligator (*Alligator sinensis*). *Zool. Res.* **25**, 415–421.

Sommer S. 2005 The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Front Zool.* **2**, 16–34.

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

Thompson J. D., Gibson T. J., Plewniak F., Jeanmougin F. and Higgins G. D. 1997 The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*. **24**, 4876–4882.

Thorbjarnarson J., Wang X. M. and Ming S. 2002 Wild populations of the Chinese alligator approach extinction. *Biol. Conserv.* **103**, 93–102.

Wang Y. Q., Zhu W. Q. and Wang C. L. 2003 D-loop sequence variation of mitochondrial DNA in captive Chinese alligator. *Acta Genet. Sin.* **30**, 425–430.

Wang Y. Q., Zhu W. Q., Huang L., Zhou K. Y. and Wang R. P. 2006 Genetic diversity of Chinese alligator (*Alligator sinensis*) revealed by AFLP analysis: an implication on the management of captive conservation. *Biodiversity Conserv.* **15**, 2945–2955.

Wu X. B., Wang Y. Q., Zhou K. Y., Zhu W. Q., Tong Z. Z., Nie J. S. *et al*. 2002 Genetic variation in captive population of Chinese alligator, *Alligator sinensis*, revealed by random amplified polymorphic DNA (RAPD). *Biol. Conserv.* **106**, 435–441.

Xia X. and Xie Z. 2001 DAMBE: software package for data analysis in molecular biology and evolution. *J. Hered.* **92**, 371–373.

Xu S. X., Ren W. H., Zhou X. M., Zhou K. Y. and Yang G. 2010 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.* **71**, 6–22.

Received 2 February 2012, in revised form 1 May 2012; accepted 23 May 2012
Published on the Web: 10 August 2012