Comparison of complete mitochondrial DNA control regions among five Asian freshwater turtle species and their phylogenetic relationships

Y. Jiang, L.W. Nie, Z.F. Huang, W.X. Jing, L. Wang, L. Liu and X.T. Dai

School of Life Sciences, Anhui Normal University,
Provincial Key Lab of the Conservation and Exploitation Research of
Biological Resources in Anhui, Wuhu, China

Corresponding author: L.W. Nie
E-mail: lwnie@mail.ahnu.edu.cn

ABSTRACT. The complete mitochondrial DNA (mtDNA) control regions (CR), cytochrome b (Cyt b), NADH dehydrogenase subunit 4 (ND4) and cytochrome coxidase subunit I (CO I) genes of four Asian freshwater turtles, *Mauremys japonica*, *Ocadia sinensis*, *M. mutica*, and *Annamemys annamensis*, were sequenced using universal PCR and long-PCR techniques. Combined with CR sequences of *Chinemys reevesii*, the composition and structure of CR of the five species were compared and analyzed. Three functional domains (TAS, CD and CSB) in CR and their conserved sequences (TAS, CSB-F, CSB-1, CSB-2, and CSB-3) were identified based on sequence similarity to those of other turtles. At the 3' end of CSB, six type motifs of variable number of tandem repeats (VNTRs) of five species were recognized, in which the TTATATTA motif may be the VNTR motif of the ancestral species of these five turtles. Comparison of nucleotide divergences among Cyt b, ND4, CO I, and CR of 11 turtle species using transitions + transversions and transversions-only methods supported the conclusion that CR evolved 2.6- to 5.7-fold faster than the other mtDNA genes.

After excluding VNTRs of CR, molecular phylogenetic trees were constructed with maximum parsimony, maximum likelihood and Bayesian inference methods. The results supported an expanded clade of *Mauremys*, which included species formerly in *Ocadia*, *Chinemys*, *Mauremys*, and *Annamemys*; this was also reflected in the results of VNTR analysis.

INTRODUCTION

The control region (CR), the longest non-coding region in mitochondrial DNA (mtDNA), has been regarded as the most variable region in the whole mtDNA sequence. Similar to other vertebrates, the following three parts have also been identified in turtles: central conserved sequence block (CD), termination associated sequence (TAS), and conserved sequence blocks (CSB). The CD domain comprises highly conserved CR DNA sequences. The CSB domain often contains a variable number of tandem repeat (VNTR) sequences. Up to now, many studies focused on the CR of mtDNA have been implemented in mammals, fishes and birds; however, studies on the CR structure of turtles are quite few, and only some species have been reported (Zardoya and Meyer, 1998; Zhang L et al., 2008; Yan et al., 2008; Zhang YY et al., 2009; Li et al., 2010).

The non-coding CR is the fastest evolving region of mtDNA (Vigilant et al., 1991). In comparison, the rRNA genes are the slowest, and the protein-coding genes and the tRNA genes possess moderate evolving rates. Some authors have estimated that CR evolves at a 5- to 10-fold higher substitution rate compared to the rest of the mtDNA sequences in the study of human mtDNA (Aquadro and Greenberg, 1983). However, CR has been found to have approximately the same transition rate as the cytochrome b gene (Cyt b) in research on the gnatcatcher (Zink and Blackwell, 1998). Evidence about lower evolving rate of mtDNA in turtles than in other vertebrates has been accumulated (Avise et al., 1992; Lamb et al., 1994), and it is worth considering whether CR evolves at a more rapid rate than other mtDNA sequences in turtle groups.

The turtle family Geoemydidae is the largest turtle family in the world, which is composed of 19 genera and more than 60 species (Rhodin et al., 2008). Five Asian freshwater turtles in this study, *Annamemys annamensis*, *Chinemys reevesii*, *Ocadia sinensis*, *Mauremys mutica*, and *M. japonica*, have been formerly distributed in several separate genera, *Annamemys*, *Chinemys*, *Ocadia*, and *Mauremys*, respectively, based on morphological studies (Zhao and Zhou, 2004). Hirayama et al. (2007) recommended splitting them into five genera (*Mauremys*, *Cathalemys*, *Chinemys*, *Ocadia*, and an unnamed new genus) based on the morphology of the palate. However, Honda et al. (2002) found a close affinity between *M. japonica* and *C. reevesii* by using both 12S and 16S rRNA data sets. Subsequent phylogenetic research based on molecular data also showed that the genera *Ocadia* and *Chinemys* rendered *Mauremys* paraphyletic (Barth et al., 2004; Feldman and Parham, 2004; Spinks et al., 2004). Furthermore, Barth et al. (2004) raised two resolving schemes, either unite all species into an expanded genus *Mauremys* or split *Mauremys* into four genera. There is still no concluding consensus on their phylogenetic relationships.

In this study, we determined the mitochondrial sequences of the CR and the Cyt b, NADH dehydrogenase subunit 4 (ND4) and cytochrome coxidase subunit I (CO I) genes of *M. japonica*, *O. sinensis*, *M. mutica*, and *A. annamensis*. Combined with the homologous sequences of several species, the CR structures of five turtles were characterized, the origin and evolution of tandem repeats of the mtDNA CR were elaborated, the evolution rates of protein-coding genes and non-coding genes were examined, and the phylogenetic relationships between species were revealed.

MATERIAL AND METHODS

Sample, DNA isolation, long-polymerase chain reaction (PCR) and sequencing

Five eastern Asian turtles were selected, including the formerly monophyletic genera: *Mauremys* (*M. mutica*, *M. japonica*), *Annamemys* (*A. annamensis*), *Ocadia* (*O. sinensis*), and *Chinemys* (sequences of *C. reevesii* were retrieved from GenBank). Total DNA was extracted using the proteinase K method, kept at -20°C, for PCR amplification.

Universal PCR protocol was used for the amplification of CO I, Cyt b and ND4 genes with primers listed in Table 1: 95°C pre-denaturation for 2 min, followed by 35 cycles at 94°C for 50 s, 52° to 55°C for 50 s, then 72°C for 1 min, and a final extension at 72°C for 10 min. Each reaction (50 µL) contained 200 ng template DNA, 5 µL 10X buffer, 4 µL MgCl$_2$ (2.5 mM), 3 µL of each dNTP, 2 µL of each primer (5 µM), and 0.5 U Taq DNA polymerase (5 U/µL; Fermentas). Long-PCR was carried out for the amplification of CR with a pair of long-PCR primers. PCR protocol was as follows: initial denaturation (94°C, 1 min), then 30 cycles of denaturation (98°C, 10 s), primer annealing and extension at the same temperature (58° to 62°C, 5 min), then a final extension (72°C, 10 min). Each reaction (50 µL) contained 5 µL 10X buffer II (Mg^{2+} Plus), 8 µL dNTP (2.5 µM), 2.5 µL template DNA, 1 µL of each primer (25 µM), 0.5 µL LA Taq polymerase (5 U/µL, TaKaRa, Dalian) and 32 µL super-purified water.

PCR products were checked for correctness size by electrophoresis on 1.0% agarose gel. DNA fragments of intended sizes were recovered using a DNA Gel Extraction Kit (Axygen, Hangzhou). Then, the purified PCR products were directly sequenced in both directions with ABI3730 automated sequencer (Invitrogen Biotechnology Co., Ltd., USA).

Table 1. Oligonucleotide primers used in this study.

Upper primer	Upper primer sequence (5'→3')	Lower primer	Lower primer sequence (5'→3')
L$_{CO\ I}$	THTTCTCYACYAACCAYAAAG	H$_{CO\ I}$	AAATCCTGCTATRATRGCGAA
L$_{Cyt\ b}$	AACCACCGTTGTATTCAACTA	H$_{Cyt\ b}$	CAATCTTTGGTTTACAAGACC
L$_{ND4}$	GAACCAGTTACACGAAAACG	H$_{ND4}$	GCTGTTTTTACGGCTGTTTTTG
L$_{CR}$	GGAGGACAACCAGTAGAAAACCCA	H$_{CR}$	ATTGGCTACACCTTGACCTGAC

H = A/C/T; Y = C/T; R = A/G.

Data analysis

Sequence data were analyzed with EditSeq (DNASTAR). tRNA genes were identified by comparison with the corresponding known sequences of other turtles using the tRNA Scan-SE1.21 software (http://lowelab.ucsc.edu/tRNAscan-SE). Besides, the VNTRs of the

five species were analyzed with SSRHunter 1.3 programs and eye-examined. MEGA v. 4.0 (Tamura et al., 2007) was used to calculate pairwise distances and sequence statistics. Tamura-Nei distances (Tamura and Nei, 1993) based on the CR, ND4, Cyt b, and CO1 genes were calculated with Tamura-Nei model at 875, 1342, 1119, and 1542 nucleotide sites, respectively, which account for the observed nucleotide bias. Mean divergences were obtained using gamma distance calculation [Tamura-Nei (TrN+G) model, $\alpha = 0.5$, transitions + transversions and transversions-only, respectively], which takes into account rate heterogeneity among sites.

Phylogenetic analyses

The complete mitochondrial CR of 11 Geoemydidae turtles (Table 2) were selected to construct phylogenetic trees. Two typical and uncontroversial tortoises, *Indotestudo elongate* (DQ080043) and *I. forstenii* (DQ080044), were used as outgroups based on the alliance between Geoemydidae and Testudinidae. Ambiguous alignments of VNTRs at the 3' end of CSB domain were removed from the final phylogenetic analysis in order to establish more accurate trees. The final phylogenetic analysis included 894 bp alignments of 13 turtles.

Table 2. Thirteen turtles analyzed in this study.

Suborder	Family	Species	Accession No.
Cryptodira	Geoemydidae	*Chinemys reevesii*	AY676201
		Ocadia sinensis	This study
		Mauremys mutica	This study
		Mauremys japonica	This study
		Annamemys annamensis	This study
		Cuora aurocapitata	AY874540
		Pyxidea mouhotii	DQ659152
		Cistoclemmys flavomarginata	EU708434
		Cistoclemmys galbinifrons	EU809939
		Sacalia quadriocellata	EF088646
		Cyclemys atripons	EF067858
	Testudinidae	*Indotestudo elongate*	DQ080043
		Indotestudo forstenii	DQ080044

DNA alignment was carried out with ClustalX v. 1.8 with default parameters. Gaps resulting from the alignments were treated as missing data. DAMBE (Xia and Lemey, 2009) was used to test substitution saturation. Phylogenetic trees were constructed using maximum parsimony (MP), maximum likelihood (ML) and Bayesian inference (BI). MP and ML trees were performed with PAUP*4.0b10 (Swofford, 2002) using the heuristic search. Of the 894 characters, 527 were constant, 128 variable characters were parsimony-uninformative and 239 were parsimony-informative. Bootstrap analyses (BP) were used to examine the robustness of the resulting bifurcations within trees. To find the best-known likelihood tree, the TrN+G model of evolution was selected by AIC in Modeltest 3.7 (Posada and Crandall, 1998). MP trees tested 1000 replicates while 100 bootstrap resamplings were carried out in ML analyzes. For BI analysis, the MrBayes v.3.1.2 program (Ronquist and Huelsenbeck, 2003) was used. Four chains ran in parallel for 100,000 generations. Trees were sampled every 1000 generations. The first 250 trees were discarded as burn-in, and then the Bayesian posterior probabilities (PP) were calculated according to the remaining set of trees.

RESULTS

The composition and structure of the control region of five freshwater turtles

Control region of all the five species was located between tRNA[pro] and tRNA[phe] (Figure 1), and their lengths ranged from 914 bp in *M. japonica* to 1175 bp in *A. annamensis*, which were related with the lengths of CSB, spanning from 286 bp in *M. japonica* to 547 bp in *A. annamensis*. The overall control region was AT-rich and GC-poor, which is consistent with observations in other turtles (Yan et al., 2008; Zhang et al., 2009). The average A+T base composition of CR in five freshwater turtles was 68.86% (Table 3). In the five species, we identified three domains reported in other turtles, which were TAS domain, CD domain and CSB domain in order. CD domain had one conserved sequence block (CSB-F), CSB domain had three characteristic boxes (CSB-1, CSB-2 and CSB-3) as well as several VNTRs located at the 3' end of CSB. The position and lengths of each domain are given in Figure 1 and Table 3, respectively.
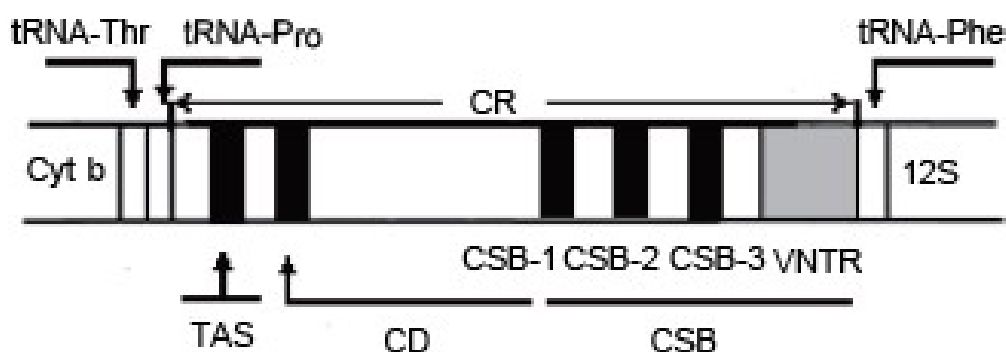
Figure 1. Similar structure of the control region (CR) of the five freshwater turtles. Cyt b = cytochrome b; VNTR = variable number of tandem repeats; TAS = termination associated sequence; CD = central conserved sequence block; CSB = conserved sequence blocks.

Table 3. Nucleotide compositions and length of the mtDNA control region of the five freshwater turtles.

	A (%)	C (%)	G (%)	T (%)	A+T (%)	Total length	TAS length	CD length	CSB length
M. mutica	32.5	18.9	11.2	37.4	69.9	1071	290	339	442
A. annamensis	32.5	16.6	10.5	40.4	72.9	1175	290	338	547
M. japonica	32.4	20.8	13.0	33.8	66.2	914	288	340	286
O. sinensis	31.7	21.2	12.4	34.8	66.5	935	288	339	308
C. reevesii	34.3	20.1	11.3	34.3	68.6	1072	305	339	428
Average	32.7	19.4	11.6	36.3	68.86	1033.4	292.2	339.0	402.2

For abbreviations, see legend to Figure 1.

The core sequences (ACAT) of TAS and its reverse complementary sequences (TGTA) were distinguished in five freshwater turtles. 5'-GACAT-3' was present in all five species and 5'-AACAT-3' appeared in all the species except *C. reevesii*, which had another sequence of 5'-TACAT-3' undiscovered in other four species. In five species, CSB-F was recognized, two transitions were found, which all occurred between A and G, and the general sequence of it (5'-YTCACGAGAYATAAGCAAC-3', Y = A or G) was concluded. Three short conserved sequences of CSB were also identified according to the sequences of reported relative species.

The characteristic motif was 5'-GACATA-3 in CSB-1, 5'-CTAAACCCCCCTACCCCC-3' in CSB-2 and 5'-TCGTCAAACCCCTAAATCC-3' in CSB-3. There was only one transition from A to G in CSB-1 and from T to C in CSB-3 of *O. sinensis*, and five turtles had the completely same sequences of CSB-2 (see Table 4).

Table 4. Common sequences of the termination associated sequence (TAS) and several conserved sequence blocks (CSB) of the five freshwater turtles.

	TAS	CSB-F	CSB-1	CSB-2	CSB-3
C. reevesii	GACAT	ATCACGAGAGATAAGCAAC	TTAATGCTTGTAAGACATA	CTAAACCCCCCTACCCCC	TCGTCAAACCCCTAAATCC
M. mutica	.....	.........A.........	............G......	....	....
A. annamensis	.....	....	....	....	....
O. sinensis	.....	.........A.........	............G......	....	....
M. japonica	.....	G........A.........	....	....	...................T

Dots indicate the same nucleotides as the top sequences.

We identified six kinds of VNTRs at the 3' end of CR in five species. The repeated motifs and numbers are shown in Figure 2. Only one type of VNTR motifs was found in *O. sinensis* and *M. japonica*, which had shorter CR because of fewer repeat times or shorter repeat motifs. *C. reevesii*, *M. mutica* and *A. annamensis* had longer CR lengths due to more repeat types or repeat times. This phenomenon was more obvious in *A. annamensis*, which possessed 32 repeat times of (TTATATTA) motifs. The same motifs of (TTATATTA) were also found in *M. mutica* with 8 repeats. However, crossing-distributed motifs of R1 and R2 in *M. reevesii* were not turned up in other species, which may be caused by the species diversity. Moreover, we also found that the base composition of the VNTRs was primarily AT-rich with only a few C nucleotides in *C. reevesii*, *M. mutica* and *O. sinensis* and no G nucleotides (Figure 2).
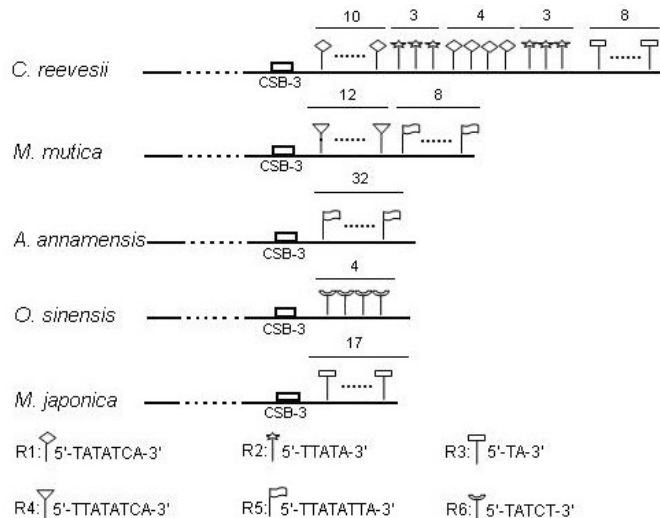
Figure 2. Comparison of the variable number of tandem repeats (VNTRs) of the control region of the five freshwater turtles. Numbers above the lines represent the repeat times of different VNTR motifs. CSB = conserved sequence blocks.

Among the three domains of CR, CD domain is regarded as the most conservative while TAS is the highest variable domain in the whole CR (Zhu et al., 2007). In our five freshwater species, there was almost no difference among their CD lengths, while CSB domain length varied the most. The average content of guanine (G) in CD domain (17.2%) was higher than that in TAS domain (13.5%) and CSB domain (5.5%), but the average content of adenine (A) (27.2%) was distinctly lower than that in TAS domain (32.0%) and CSB domain (37.8%). In addition, TAS domain and CSB domain showed higher nucleotide variation, 35.69 and 26.68%, respectively, while CD domain was the lowest (13.24%) (data not shown).

Comparative analysis of nucleotide divergences based on four mtDNA genes of 11 Geoemydidae species

Mean divergences were calculated with methods of transitions + transversions and transversions-only, displayed in Table 5. Two methods yielded the same results, which show the same trend change of four genes' mean divergence. CR was the most divergent compared to other three genes. The subsequent divergence variant was the ND4 gene, followed by the Cyt b gene, while the CO I was the least divergent gene. Through the analysis based on transitions + transversions, we concluded that the mean divergence was 1.0- to 1.5-fold higher for CR than for the other three genes. Nevertheless, distance comparisons based on transversions-only method evolved more times, through which CR were 2.6-, 3.2- and 5.7-fold faster than ND4, Cyt b and CO I gene, respectively. Since the species used for comparison of nucleotide divergences among each gene came from the same lineages, these dataset may provide information about the speed of evolution rate among these genes. All the inference above suggested that CR possesses the fastest evolutionary rate than the other three protein-coding genes.

Table 5. Nucleotide divergences among 11 Geoemydidae turtles based on four mtDNA gene sequence data.

	Mean divergence ± standard error[a]	Mean divergence ± standard error[b]
CR	0.174 ± 0.021	0.057 ± 0.007
ND4	0.172 ± 0.019	0.022 ± 0.004
Cyt b	0.151 ± 0.019	0.018 ± 0.004
CO I	0.114 ± 0.012	0.010 ± 0.002

[a]Tamura-Nei distances based on transitions + transversions; [b] Tamura-Nei distances based on transversions-only. CR = control region; ND4 = NADH dehydrogenase subunit 4; Cyt b = cytochrome b; CO I = cytochrome coxidase subunit I.

Phylogenetic results

Results of substitution saturation test based on 13 CR sequences indicated that the base substitutions did not reach saturation, and these data were adapted to analyze phylogeny of related species (data not shown). Phylogenetic trees of 13 species constructed by MP, ML and BI methods recovered the same topological structure (Figure 3A and B). In all trees, the five species, *C. reevesii*, *O. sinensis*, *M. japonica*, *M. mutica*, and *A. annamensis*, as

well as four *Cuora* species, always congregated together with high supports (≥87%). The five turtles in this study were partitioned into two clusters: *M. mutica* and *A. annamensis* assembled together with strong supports, while *O. sinensis*, *C. reevesii* and *M. japonica* were collected together with relatively high BP and PP values, in which *M. japonica* showed a closer affinity with *M. reevesii* than with *O. sinensis*, but all the supports were low. The five assembled clade was sister to the *Cuora*, which formerly included *Cistoclemmys* and *Pyxidea* genus. In all selected Geoemydidae species, *Cyclemys atripons* was sister to *Sacalia quadriocellata* and resided at the basis of the Geoemydidae ML tree, but MP and BI methods recovered that *S. quadriocellata* diverged first than *C. atripons*.
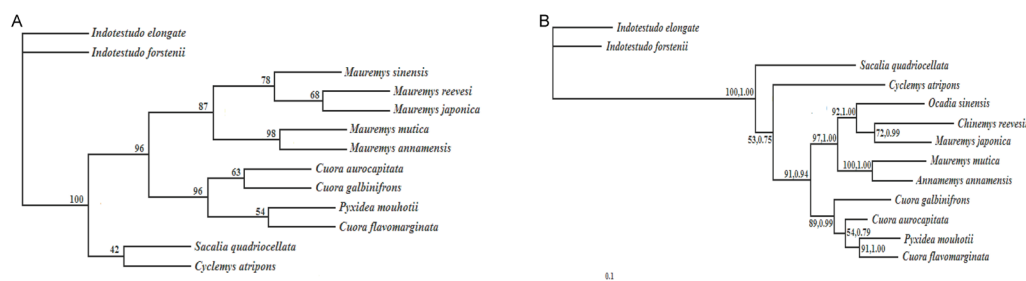
Figure 3. A. Maximum likelihood cladogram derived from 894 nucleotides of the control region in 13 turtle species. Numbers above branches indicate percentages of 1000 bootstrap replicates for maximal parsimony. *Indotestudo elongate* and *I. forstenii* are used as outgroups. **B.** Maximum parsimony and Bayesian analysis cladogram of 13 turtle species. Numbers near branches indicate percentages of 100 bootstrap replicates for maximum likelihood followed by posterior probabilities for Bayesian inference.

DISCUSSION

Characteristics of control region in five freshwater species

One control region located between tRNA^{pro} and tRNA^{phe} was observed in each of these five species, which is similar to other turtles (Zhang L et al., 2008; Yan et al., 2008; Zhang YY et al., 2009). Similar with other vertebrates, CR was divided into three domains (TAS, CD, CSB) despite its high variability. They showed identical content of adenine (A) + thymine (T) and guanine (G), especially in CSB, which was the highest for A+T (74.2%) and the lowest for G (5.5%) (Zhang et al., 2009), but the content of G in CD domain was higher than that in TAS domain and CSB domain. This coincides with the viewpoint that TAS domain is cytosine (C)-rich, CD domain is guanine (G)-rich, and CSB domain is AT-rich but low in light strand G in birds (Crochet and Desmarais, 2000); however, the C-rich content of TAS was not obvious in our results.
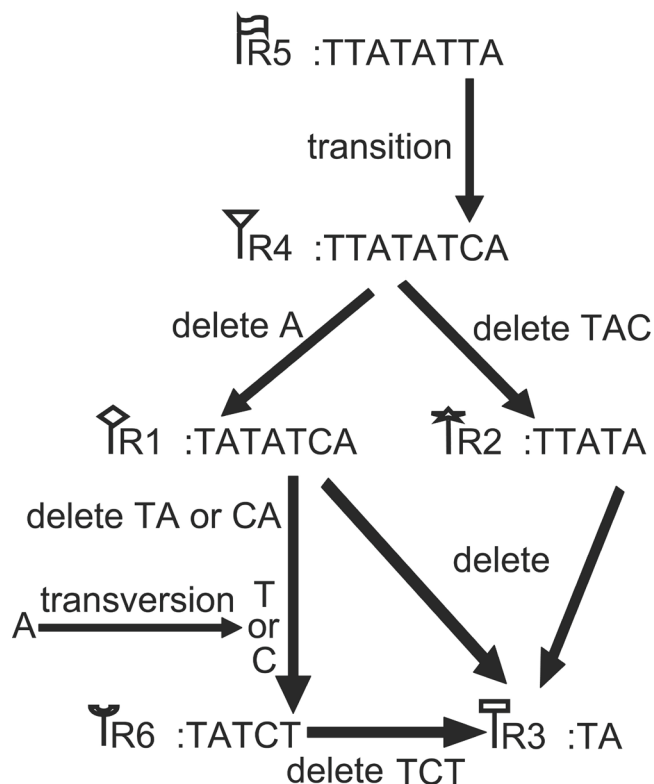
We have discerned the homologous sequences (RACAT, R = G, T, or A) of TAS and its complementary sequences (MTGTA, M = C, A or T) in five turtles. TAS is supposed to be involved in the termination of DNA synthesis (Doda et al., 1981) through forming stable stem-loop secondary structures. As two TASs appeared in the same species, it is likely that the other is not functional (Zhang et al., 2009). Up to now, only CSB-F can be found within the CD domain in turtles (Peng et al., 2006; Yan et al., 2008; Zhang et al., 2009), which is considered

to be correlative with regulation of H-strand synthesis, but other sequences such as CSB-B, -C, -D, and -E in CR found in other vertebrates (Randi and Lucchini, 1998; Zhu et al., 2007) were not found in turtles. CSB-F is present in most vertebrates, which implies that CSB-F might be more indispensable than CSB-B, -C, -D, and -E, in H-strand replication. CSB is the most important functional domain for the presence of O_H, the two promoters (HSP and LSP) and three short conserved sequence blocks (CSB-1, CSB-2 and CSB-3), which are usually considered to regulate the initiation of replication and the transcription of the mitochondrial genome. CSB-1 is the most conserved component of CSB and it exists in the CR of almost all vertebrates. CSB-2 or CSB-3 was only found in a few of fishes, combined CSB-2 and CSB-3 was found in avain, and complete deletion of CSB-3 was found in cattle, dolphin, sheep, and whale (Su, 2005). These seemed to indicate that CSB-1 plays a more dominant role than CSB-2 and CSB-3, but distant difference and insertion of repeat sequences cannot be found in our five species, showing a discrepancy of CSB1-3 of mtDNA CR between mammalians and turtles, which may be associated with low evolution and high conservation of mtDNA in turtles.

Several VNTR motifs were found at the 3' end of CSB adjoining to tRNA[phe]. Results suggested that the length of the mitochondrial CR of the five species may be attributed to repeat motifs and repeat times at the 3' end of their control region. It was because of the existence of 32 repeat times of TTATATTA motifs that *A. annamensis* had the longest length in these species, while *O. sinensis* and *M. japonica* had shorter CR sequence lengths owing to only one type of VNTR motifs or shorter repeats. Extensive VNTRs have been found in the mtDNA CR of many vertebrate species. Compared with the *Cuora* turtles (Zhang et al., 2009), the five species in this paper exhibited less repeat motifs and repeat times. Some molecular mechanisms have been recommended for the origin of tandem repeated sequences and subsequent generation of mtDNA length variants, such as recombination and transposition (Rand and Harrison, 1989), unequal crossing-over (gene conversion; Hoelzel et al., 1993), and slipped-strand mispairing (Fumagalli et al., 1996). Our results suggested that the origin of tandem repeats of mtDNA CR may result from slipped-strand mispairing during mtDNA replication. It is proposed that a new repeated sequence may be generated through deletions of extraneous and mutational sequences, looping out and expansion of the sequence (Xiao et al., 2006). The sequence polymorphism within the repeated arrays results in an intermediate stage in the replication process, and its presence requires a balance between two opposing evolutionary forces: point mutations and insertion/deletion mutations (Rand, 1994). The former causes divergence, whereas the latter causes homology.

In this paper, TTATATTA was found in two species, *M. mutica* and *A. annamensis*, and TTATATCA motifs that specially existed in *M. mutica* may have undergone a transition of TTATATTA between C and T (Figure 4). This may cause a closer relationship between the two species. In *C. revesii*, TATATCA probably have undergone the deletion of T from the TTATATCA unit, while TTATA motifs derived from deletions of TTATATCA or TTATATTA, and TATCT in *O. sinensis* may have experienced a deletion of TA or CA from TATATCA of *C. revesii*, then a transversion from A to T or A to C, respectively. TA motifs found in *C. revesii* and *M. japonica* may undergo the deletion of TCT or TTA from TATCT or TTATA presented in *O. sinensis* and *C. revesii*, respectively. These may reveal that there are some correlations among VNTRs of turtle lineages.

R5 :TTATATTA

↓ transition

R4 :TTATATCA

delete A ↙   ↘ delete TAC

R1 :TATATCA   R2 :TTATA

delete TA or CA

A → transversion → T or C

delete

R6 :TATCT → delete TCT → R3 :TA

Figure 4. Tentative conversion processes among different VNTR motifs.

Considerations of evolution rate

The CR is the fastest evolving region of mtDNA and our results also agree on this in turtle groups through the compative analysis of nucleotide divergences based on the four mtDNA genes of the same species. Mean divergences based on transversions-only showed 2.6- to 5.7-fold of CR compared to the other three protein-coding genes, which may imply that CR is evolving 2.6 to 5.7 times as fast as other genes. There is a high A+T content in CR (68.86%), which increases the possibility of transversion from A to T, causing a decrease of transition/transversion values (Desalle et al., 1987), although transitions/transversions did not reach saturation in these five species, and a more accurate deduction of evolving rate should be based on transversions-only. Previously, a rapid evolution rate (5- to 10-fold or 2.8 to 5 times) of CR was supported by some authors (Aquadro and Greenburg, 1983; Cann et al., 1984), while other studies suggested that CR evolves at approximately the same rate as Cyt b transitions (Zink and Blackwell, 1998). Ruokonen and Kvist (2002), stydying the avian CR, revealed that the control region was not always the most variable region of the mtDNA; there were differences in the rate of divergence among the lineages. Studying on Geoemydidae turtles CR, we supported that CR evolved 2.6- to 5.7-fold compared to other genes of mtDNA.

Phylogenetic analysis

Control region provided an effective information (Figure 3). Our phylogenetic trees suggested an expanded genus *Mauremys*, including the former genera of *Chinemys*, *Ocadia*, *Annamemys*, and *Mauremys*. The same point has also been advanced by others (Barth et al., 2004), who also raised another viewpoint that species of these genera could be split into four genera. Hirayama et al. (2007) recommend that *Mauremys* should be split into five genera (*Mauremys*, *Cathalemys*, *Chinemys*, *Ocadia*, and an unnamed new genus) based on the morphology of the palate. In our expanded genus *Mauremys*, *C. reevesii*, *M. japonica* and *O. sinensis* clustered together while *A. annamensis* was sister to *M. mutica* with high supports. This closer relationship between *M. mutica* and *M. annamensis* has also been suggested by others (Barth et al., 2004; Spinks et al., 2004). Interestingly, Fong et al. (2007), on the basis of both nuDNA and mtDNA data, suggested that *A. annamensis* may be reckoned as the same species as *M. mutica* due to their paraphyletic relationships, but most of their samples sourced from farms or trade markets. This was not mirrored in our results. The closest affinity of *M. japonica* and *C. reevesii* has been revealed (Honda et al., 2002), others (Barth et al., 2004; Spinks et al., 2004) have also displayed that *M. japonica* is sister to *O. sinensis*, then paraphyletic with respect to *Chinemys*, but supports are quite low. Congruent with this, our results supported a clade composing these three species, but their relationships were still obscured due to low supports. Besides, coincident viewpoint of sister relationship of expanded *Mauremys* and *Cuora* has been confirmed (Barth et al., 2004; Spinks et al., 2004; Zhang et al., 2009). Closer relationships between *Cuora* and *Mauremys* could explain the frequently reported hybrids between these two genera (Stuart and Parham, 2007). Combined with phylogenetic analysis, we suggested that morphological characters of these five species in this study can be considered as evidence for distinguishing among species rather than genus, such as the difference of secondary palate between clades of *O. sinensis* + *M. japonica* + *Chinemys* and *M. mutica* and *A. annamensis*; straight yellow and black rays covering limbs and the sides of the head of *O. sinensis*, and yellow-green longitudinal grain at the sides of the head of *A. annamensis*.

Correlations between VNTR structures and phylogeny results of five species in this study

A positive relevance between VNTRs and phylogeny was detected in our results although phylogenetic trees were constructed after removed VNTRs. TTATATTA may stand for VNTR motifs of ancestral species of these five turtles, while our phylogenetic results implied unification of the five turtles that formerly belonged to four genera into an expanded genus *Mauremys*. Compared with each other, the same motifs were found in *M. mutica* and *A. annamensis*, and another motif TTATATCA can only pass through one step mutation from TTATATTA; this also suggested a closer phylogenetic relationship between them. Other three species undergo more changes of ancestral motifs. Motifs in *O. sinensis* and *M. japonica* may have undergone different changes from *C. revesii*, so they exhibited closer relationships among five species in our phylogenetic analyzes. Three types of VNTRs were found in *C. revesii*, in which there were two crossing-distributed repeated sequences, and only one type of VNTR motif appeared in *O. sinensis* and *M. japonica*; this may indicate that *C. revesii* evolved at more rapid rate. This structure characteristic of VNTRs was reflected by the lon-

gest branch length of *C. revesii* in the phylogenetic trees (Figure 3B). Just as *M. japonica* and *C. revesii* had the same motifs of TA, they did show closer affinity in our three trees.

ACKNOWLEDGMENTS

REFERENCES

Aquadro CF and Greenberg BD (1983). Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics* 103: 287-312.

Avise JC, Bowen BW, Lamb T, Meylan AB, et al. (1992). Mitochondrial DNA evolution at a turtle's pace: evidence for low genetic variability and reduced microevolutionary rate in the Testudines. *Mol. Biol. Evol.* 9: 457-473.

Barth D, Bernhard D, Fritzsch G and Fritz U (2004). The freshwater turtle genus *Mauremys* (Testudines, Geoemydidae) - a textbook example of an east-west disjunction or a taxonomic misconcept? *Zoolog. Scripta* 33: 213-221.

Cann RL, Brown WM and Wilson AC (1984). Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. *Genetics* 106: 479-499.

Crochet PA and Desmarais E (2000). Slow rate of evolution in the mitochondrial control region of gulls (Aves: Laridae). *Mol. Biol. Evol.* 17: 1797-1806.

Desalle R, Freedman T, Prager EM and Wilson AC (1987). Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. *J. Mol. Evol.* 26: 157-164.

Doda JN, Wright CT and Clayton DA (1981). Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. *Proc. Natl. Acad. Sci. U. S. A.* 78: 6116-6120.

Feldman CR and Parham JF (2004). Molecular systematics of old world stripe-necked turtles (Testudines: *Mauremys*). *Asiat. Herpetol. Res.* 10: 28-37.

Fong JJ, Parham JF, Shi H and Stuart BL (2007). A genetic survey of heavily exploited, endangered turtles: caveats on the conservation value of trade animals. *Anim. Conserv.* 10: 452-460.

Fumagalli L, Taberlet P, Favre L and Hausser J (1996). Origin and evolution of homologous repeated sequences in the mitochondrial DNA control region of shrews. *Mol. Biol. Evol.* 13: 31-46.

Hirayama R, Kaneko N and Okazaki H (2007). *Ocadia nipponica*, a new species of aquatic turtle (Testudines: Testudinoidea: Geoemydidae) from the Middle Pleistocene of Chiba Prefecture, central Japan. *Paleontol. Res.* 11: 1-19.

Hoelzel AR, Hancock JM and Dover GA (1993). Generation of VNTRs and heteroplasmy by sequence turnover in the mitochondrial control region of two elephant seal species. *J. Mol. Evol.* 37: 190-197.

Honda M, Yasukawa Y and Ota H (2002). Phylogeny of Eurasian freshwater turtles of the genus *Mauremys* Gray 1869 (Testudines), with special reference to a close affinity of *Mauremys japonica* with *Chinemys reevesii*. *J. Zool. Systemat. Evol. Res.* 40: 195-200.

Lamb T, Lydeard C, Walker RB and Gibbons JW (1994). Molecular systematics of map turtles (*Graptemys*): a comparison of mitochondrial restriction site versus sequence data. *Syst. Biol.* 43: 543-559.

Li XS, Nie LW, Wang L, Xiong L, et al. (2010). The mitochondrial genome complete sequence and organization of the pig-nosed turtle *Carettochelys insculpta* (Testudines, Carettochelyidae) and its phylogeny position in Testudines. *Amphibia-Reptilia* 31: 541-551.

Peng QL, Nie LW and Pu YG (2006). Complete mitochondrial genome of Chinese big-headed turtle, *Platysternon megacephalum*, with a novel gene organization in vertebrate mtDNA. *Gene* 380: 14-20.

Posada D and Crandall KA (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817-818.

Rand DM (1994). Concerted evolution and RAPping in mitochondrial VNTRs and the molecular geography of cricket populations. *EXS* 69: 227-245.

Rand DM and Harrison RG (1989). Molecular population genetics of mtDNA size variation in crickets. *Genetics* 121: 551-569.

Randi E and Lucchini V (1998). Organization and evolution of the mitochondrial DNA control region in the avian genus *Alectoris*. *J. Mol. Evol.* 47: 449-462.

Rhodin AGJ, van Dijk PP and Parham JF (2008). Turtles of the World: Annotated Checklist of Taxonomy and Synonymy.

Chelonian Research Foundation, Lunenburg.

Ronquist F and Huelsenbeck JP (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572-1574.

Ruokonen M and Kvist L (2002). Structure and evolution of the avian mitochondrial control region. *Mol. Phylogenet. Evol.* 23: 422-432.

Spinks PQ, Bradley SH, Iverson JB and McCord WP (2004). Phylogenetic hypotheses for the turtle family Geoemydidae. *Mol. Phylogenet. Evol.* 32: 164-182.

Stuart BL and Parham JF (2007). Recent hybrid origin of three rare Chinese turtles. *Conserv. Genet.* 8: 169-175.

Su Y (2005). Conserved and heteroplasmy on mitochondrial DNA control region in animal. *Sichuan J. Zoo.* 24: 669-672.

Swofford DL (2002). PAUP*: Phylogenenc Analysis using Parsimony (*and Other Methods), Version 4.0b l0. Sinauer Associates, Sunderland.

Tamura K and Nei M (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10: 512-526.

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.

Vigilant L, Stoneking M, Harpending H, Hawkes K, et al. (1991). African populations and the evolution of human mitochondrial DNA. *Science* 253: 1503-1507.

Xia X and Lemey P (2009). Assessing Substitution Saturation with DAMBE. 2nd edn. In: The Phylogenetic Handbook: A Practical Approach to DNA and Protein Phylogeny (Lemey P, Salemi M and Vandamme A-M, eds.). Cambridge University Press, Cambridge, 615-630.

Xiao B, Ma F, Sun Y and Li QW (2006). Comparative analysis of complete mitochondrial DNA control region of four species of Strigiformes. *Yi. Chuan Xue. Bao.* 33: 965-974.

Yan L, Zhang Y, Wang N and Zhang L (2008). Comparison of mitochondrial control region sequences between *Chelydridae* and *Platysternidae*. *Zool. Res.* 29: 127-133.

Zardoya R and Meyer A (1998). Cloning and characterization of a microsatellite in the mitochondrial control region of the African side-necked turtle, *Pelomedusa subrufa*. *Gene* 216: 149-153.

Zhang L, Nie L, Cao C and Zhan Y (2008). The complete mitochondrial genome of the Keeled box turtle *Pyxidea mouhotii* and phylogenetic analysis of major turtle groups. *J. Genet. Genom.* 35: 33-40.

Zhang YY, Nie LW, Huang YQ and Pu YG (2009). The mitochondrial DNA control region comparison studies of four hinged turtles and its phylogentic significance of the genus *Cuora sensu lato* (Testudinata: Geoemydidae). *Genes Genom.* 31: 349-359.

Zhao EM and Zhou T (2004). Atlas for Identification of Turtles and Tortoises. China Agriculture Press, Beijing.

Zhu SH, Zheng WJ, Zou JX and Yang YC (2007). Mitochondrial DNA control region structure and molecular phylogenetic relationship of Carangidae. *Zool. Res.* 28: 606-614.

Zink RM and Blackwell RC (1998). Molecular systematics and biogeography of aridland gnatcatchers (genus *Polioptila*) and evidence supporting species status of the California gnatcatcher (*Polioptila california*). *Mol. Phylogenet. Evol.* 9: 26-32.