Online journal ISSN - 1676-5680
GMR Genetics and Molecular Research
EVOLUTION AND TECHNOLOGY

Complete mitochondrial genome of Cabot's tragopan, *Tragopan caboti* (Galliformes: Phasianidae)

X.Z. Kan[1], X.F. Li[1], Z.P. Lei[2], M. Wang[1], L. Chen[1], H. Gao[3] and Z.Y. Yang[3]

[1]The Provincial Key Laboratory of the Conservation and Exploitation Research of Biological Resources in Anhui, College of Life Sciences, Anhui Normal University, Wuhu, China
[2]The Administration Bureau of Wuyanling National Nature Reserve, Taishun, China
[3]State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, China

Corresponding author: X.Z. Kan
E-mail: xianzhaokan@gmail.com

Genet. Mol. Res. 9 (2): 1204-1216 (2010)
Received February 25, 2010
Accepted March 29, 2010
Published June 22, 2010
DOI 10.4238/vol9-2gmr820

ABSTRACT. Cabot's tragopan, *Tragopan caboti*, is a globally threatened pheasant endemic to southeast China. The complete mitochondrial genome of Cabot's tragopan was sequenced. The circular genome contains 16,727 bp, encoding a standard set of 13 protein-coding genes, two ribosomal RNA genes, and 22 transfer RNA genes, plus the putative control region, a structure very similar to that of other Galliformes. As found in other vertebrates, most of these genes code on the H-strand, except for the NADH dehydrogenase subunit 6 (*nad6*) and eight tRNA genes (Gln, Ala, Asn, Cys, Tyr, Ser(UCN), Pro, Glu). All protein-coding genes initiated with ATG, except for cox1, which began with GTG, and had a strong skew of C *vs* G (GC skew = -0.29 to -0.73). One extra 'C' nucleotide was found in the NADH dehydrogenase subunit 3 (*nad3*). All the tRNA gene sequences have the potential to fold into typical cloverleaf secondary structures. Conserved sequences in three domains were identified within the control region

(D-loop). These results provide basic information for phylogenetic analyses among Galliform birds, and especially *Tragopan* species.

Key words: Mitochondrial genome; Cabot's tragopan; Phasianidae; *Tragopan caboti*; Galliformes

INTRODUCTION

The living birds classified within the order Galliformes form a large and cosmopolitan group comprising more than 250 species within some 70 genera (Monroe and Sibley, 1990). Galliformes are traditionally classified into seven families: Megapodiidae (mound builders, brush turkeys, and allies), Cracidae (curassows, guans, and chachalacas), Odontophoridae (New World quails), Numididae (guineafowl), Phasianidae (pheasants, partridges, Old World quails, and allies), Meleagrididae (turkeys), and Tetraonidae (grouse and allies) (del Hoyo et al., 1994). Based on molecular markers and comprehensive taxon sampling, current classification suggested that the Meleagrididae should be classified into the Phasianidae (Dyke et al., 2003; Crowe et al., 2006; He et al., 2009). *Tragopan* is a genus of bird in the family Phasianidae. These birds are commonly called "horned pheasants" because of two brightly colored, fleshy horns on their heads that they can erect during courtship displays. There are five recognized species in *Tragopan*, namely *T. melancephalus*, *T. satyra*, *T. temminckii*, *T. blythii*, and *T. caboti*. Cabot's tragopan, *T. caboti*, is a globally threatened pheasant endemic to southeast China, where it is known as the Yellow-bellied tragopan (Zhang and Zheng, 2007). The typical habitats of the tragopan have been seriously fragmented because of forest management for timber production and farmland reclamation in recent years (Deng and Zheng, 2004). As a result it was classified as Vulnerable (VU) in the International Union for the Conservation of Nature and Natural Resources (IUCN) Red List (IUCN 2009), and also listed in the first category of National Key Protected Wild Animals in China.

Mitochondrial genomes have great potential for resolving ancient patterns of evolutionary history and for serving as a model of genome evolution. With a few exceptions, metazoan mitochondrial genomes are double-stranded, circular molecules, 15-20 kb in size, containing 37 genes: 13 protein-coding genes, 22 transfer RNA genes (tRNAs) and two ribosomal RNA genes (rRNAs) (Wolstenholme, 1992; Boore, 1999). The genome is highly economized with few sections of noncoding DNA, intergenic regions or repetitive sequences, except for one major control region. The control region is believed to control the initiation of replication and transcription of animal mitochondrial DNA (mtDNA) (Shadel and Clayton, 1997). Because of its compactness, maternal inheritance, fast evolutionary rate compared to nuclear DNA, and short coalescence time, mtDNA is useful for population genetic studies such as the analysis of gene flow, hybridization and introgression (Moore, 1995). Thus, the mitochondrial genome can provide abundant information for evolutionary studies of many taxa, and can also be used as a source of molecular markers in the conservation studies of endangered species (He et al., 2009; Zhang et al., 2009). To date, the complete mtDNA sequences are available for 21 Galliformes species (Table 1); however, no complete mitochondrial sequence has been reported for members of the genus *Tragopan*. In this study, we present the complete mitochondrial genome of *T. caboti* (Galliformes: Phasianidae) and give a thorough description of its genome features in comparison to other Galliformes species.

Table 1. Genomic characteristics of Galliformes mitochondrial DNAs.

Species	Family	Accession No.	Heavy-strand		Protein-coding genes					*trRNA gene*		*srRNA gene*		*tRNA genes*		Control region		References
			Length (bp)	AT%	Length (bp)	AT% (all)	AT% (1st)	AT% (2nd)	AT% (3rd)	Length (bp)	AT%	Length (bp)	AT%	Length (bp)	AT%	Length (bp)	AT%	
Arborophila rufipectus	Phasianidae	FJ194942	16,728	54.5	11,380	53.5	49.6	57.8	53.2	1,623	54.3	981	54.0	1,544	56.3	1,170	59.1	He et al., 2009
Bambusicola thoracica	Phasianidae	EU165706	16,726	54.4	11,392	53.8	49.7	55.6	56.0	1,622	54.0	974	52.1	1,545	57.2	1,146	59.3	Shen et al., 2009b
Coturnix chinensis	Phasianidae	AB073301	16,687	55.4	11,397	54.8	50.5	58.3	55.7	1,607	53.9	973	54.7	1,575	56.4	1,150	60.2	Nishibori et al., 2002
Coturnix japonica	Phasianidae	AP003195	16,697	55.6	11,403	55.0	50.7	58.1	56.1	1,615	55.6	974	54.4	1,544	57.5	1,155	59.5	Nishibori et al., 2001
Francolinus pintadeanus	Phasianidae	EU165707	16,694	54.8	11,388	54.2	50.0	58.1	54.6	1,606	55.0	973	52.1	1,544	56.4	1,169	59.7	Shen et al., 2009b
Gallus gallus	Phasianidae	AP003322	16,785	54.0	11,397	52.9	49.5	57.9	51.4	1,622	53.9	976	52.8	1,543	57.5	1,232	59.9	Nishibori et al., 2005
Gallus lafayetii	Phasianidae	AP003325	16,841	54.1	11,397	53.2	49.6	57.8	51.9	1,620	53.9	977	52.4	1,539	57.2	1,292	60.1	Nishibori et al., 2005
Gallus sonneratii	Phasianidae	AP006741	16,841	54.0	11,393	53.2	49.5	57.9	52.2	1,613	53.4	978	51.7	1,534	56.8	1,292	59.8	Nishibori et al., 2005
Gallus varius	Phasianidae	AP003324	16,783	54.1	11,397	53.1	49.8	57.9	51.7	1,622	54.0	973	52.4	1,545	57.3	1,228	59.9	Nishibori et al., 2005
Lophura ignita	Phasianidae	AB164627	16,688	55.1	11,393	54.4	49.7	58.1	55.3	1,610	55.5	965	53.0	1,552	58.2	1,147	58.5	Unpublished data[a]
Lophura nycthemera	Phasianidae	EU417810	16,680	54.8	11,391	53.9	50.0	58.0	53.8	1,609	55.6	967	52.2	1,543	57.9	1,148	59.0	Shen et al., 2009b
Meleagris gallopavo	Phasianidae	EF153719	16,717	56.5	11,394	55.9	51.1	58.5	58.2	1,618	56.7	973	55.2	1,541	58.0	1,164	59.8	Guan et al., 2009
Pavo muticus	Phasianidae	EU417811	16,698	55.4	11,391	55.0	50.3	58.0	53.9	1,609	54.9	979	52.2	1,544	57.8	1,156	58.4	Shen et al., 2009b
Phasianus versicolor	Phasianidae	AB164626	16,690	56.0	11,393	55.4	50.4	58.4	57.4	1,620	55.7	966	53.1	1,547	58.8	1,150	59.6	Unpublished data[a]
Polyplectron bicalcaratum	Phasianidae	EU417812	16,702	53.4	11,391	52.7	48.9	58.0	51.1	1,601	53.2	975	50.5	1,546	56.3	1,170	59.1	Shen et al., 2009b
Syrmaticus ellioti	Phasianidae	AB164624	16,688	55.6	11,393	55.2	49.9	58.5	57.1	1,604	55.8	973	52.2	1,553	58.2	1,153	59.2	Unpublished data[a]
Syrmaticus humiae	Phasianidae	AB164625	16,686	55.7	11,393	55.1	50.0	58.5	56.9	1,610	55.9	973	52.3	1,536	58.2	1,153	59.4	Unpublished data[a]
Syrmaticus reevesii	Phasianidae	AB164623	16,678	54.7	11,393	54.2	47.2	55.8	54.4	1,607	55.0	967	52.2	1,533	57.1	1,150	57.2	Unpublished data[a]
Syrmaticus soemmerringi	Phasianidae	AB164622	16,690	55.8	11,393	55.3	50.2	58.5	57.1	1,608	56.1	971	53.7	1,549	57.8	1,152	59.1	Unpublished data[a]
Tragopan caboti	Phasianidae	GU187969	16,727	54.7	11,390	54.0	49.8	58.2	54.0	1,614	54.9	978	52.9	1,546	57.3	1,177	58.9	This study
Alectura lathami	Megapodiidae	AY346091	16,698	52.7	11,395	52.0	48.5	58.0	49.4	1,615	52.7	980	51.0	1,559	56.8	1,120	55.3	Slack et al., 2007
Numida meleagris	Numididae	AP005595	16,726	53.7	11,378	52.7	49.2	57.9	51.1	1,624	53.3	980	52.0	1,551	58.1	1,169	59.2	Nishibori et al., 2004

[a]Unpublished data: Kato S, Nishihori M and Yasue H (2008).

MATERIAL AND METHODS

Sample collection and DNA extraction

A naturally dead individual Cabot's tragopan (male) was obtained from the Wuyanling National Nature Reserve (27°43'N, 119°39'E), Zhejiang Province, China. Total genomic DNA was extracted from the muscle tissue using the standard phenol/chloroform methods (Sambrook and Russell, 2001).

PCR amplification and sequencing

To minimize the possibility of obtaining nuclear copies of mitochondrial genes (NUMTs), two long overlapping fragments (~16 kb in length) were first amplified using the long and accurate-polymerase chain reaction (LA-PCR) kit (Takara, Dalian, China). The first LA-PCR primer set was LA16SF and LA16SR of Nishibori et al. (2001). The second primer set was CytbF605 and CytbR252, designed based on the sequence of Cytochrome b gene (*cob*) from Cabot's tragopan. The LA-PCR was conducted at 94°C for 1 min, followed by 35 cycles consisting of 10 s denaturation at 98°C and a 15-min annealing and extension at 68°C, with a final extension step of 10 min at 72°C, using a DNA Thermal Cycler TC-3000 (Techne, Barloworld Scientific Ltd., UK). The amplified fragments with a size of approximately 16 kb thus obtained were used as the templates for the short overlapping fragment (1.1-1.5 kb in length) amplification with 15 primer sets (Table 2). All the 15 sets of primers were designed based on the conserved sequences of mitochondrial genome among other species of Galliformes, which were aligned using CLUSTAL X, version 2.0.10 (Larkin et al., 2007). Each primer set amplified an mtDNA fragment containing an overlap of at least 100 bp with the adjacent amplified fragment at both termini. PCR was carried out in a volume of 25 μL containing 5-50 ng DNA template, 6.25 pmol of each primer, 0.2 mM of each dNTP, 2 mM $MgCl_2$, and 0.625 U *DreamTaq*™ DNA polymerase (Fermentas, Burlington, Canada). Amplification was conducted in the DNA Thermal Cycler TC-3000 as above. PCR cycles were as follows: one cycle of 4 min at 70°C, 4 cycles of 40 s at 94°C, 20 s at 52°C, and 2 min and 10 s at 72°C, followed by 36 cycles of 20 s at 94°C, 20 s at 50-55°C, and 2 min and 10 s at 72°C. The process was completed with a final elongation at 72°C for 10 min. The band with the right size was cut out and purified using an EZ Spin Column DNA Gel Extraction Kit (Bio Basic Inc.). The purified PCR products were sequenced directly on an ABI-PRISM 3730 sequencer using a BigDye Terminator, version 3.1, Cycle Sequencing Kit (Applied Biosystems) and the corresponding primer.

Gene identification and genome analyses

DNA sequences were analyzed using the DNASTAR Lasergene, version 5.0, software. Contig assembly was performed with the ContigExpress program (a component of Vector NTI Suite 6.0). The boundaries of protein-coding genes and rRNA genes were initially identified via DOGMA (Wyman et al., 2004) using the default setting, and refined by alignment with mitochondrial genomes of other species of Galliformes. Most tRNA genes were identified using tRNAscan-SE 1.21 (Lowe and Eddy, 1997) under the 'cove only' search mode, with the vertebrate mitochondrial genetic code and 'mito/chloroplast' source. Some tRNA genes,

Table 2. Primers used in amplifying and sequencing *Tragopan caboti*.

No. of primer pair	Name	Sequences (5'-3')	Size (bp)	Sources
1	LA16SF	CCTACGTGATCTGAGTTCAGACCGGAGCAATCCAG	35	Nishibori et al., 2001
	LA16SR	TGCACCATTAGGTTGTCCTGATCCAACATCGAGGT	35	Nishibori et al., 2001
2	CytbF605	ATGAATCAGGCTCTAACAACCCTCTGGGCATC	32	This study
	CytbR252	GATGCAGATGAAGAAGAATGAGGCGCCGTTTGC	33	This study
3	AVMT1F	GCCAAATAGCATCCTTCTCC	20	This study
	AVMT1R	GAGGTGGACGATCAATAAAT	20	This study
4	AVMT2F	AACCCATTATATGTATACGG	20	This study
	AVMT2R	TTACTGCTGAGTACCCGTGG	20	This study
5	AVMT3F	GCAAAAGACTTAGTCCTAACC	21	This study
	AVMT3R	CTTTTGCGACAGAGACGGGTT	21	This study
6	AVMT4F	AAGTCGTAACAAGGTAAGTGAC	22	This study
	AVMT4R	CGCCCCAACCGAAAAATGTC	20	This study
7	AVMT5F	AAGACGAGAAGACCCTGTGG	20	This study
	AVMT5R	AGCTCTGACTCTCCTTCTGT	20	This study
8	AVMT6F	TAAGCACCCTGGCCATCACC	20	This study
	AVMT6R	ATGAGATGAGTATTGTTGAT	20	This study
9	AVMT7F	ACACAGACACGAAAAATCCT	20	This study
	AVMT7R	GTGATAAAGTTGATGGCTCCT	21	This study
10	AVMT8F	CGCATAAATAACATAAGCTTC	21	This study
	AVMT8R	GAAGCATTAAGTGGTTTGAT	20	This study
11	AVMT9F	AAGCCTTCTCAGCAAAACGA	20	This study
	AVMT9R	GCTTAGGTTCATGGTCAGGT	20	This study
12	AVMT10F	ATGACATGCCCCAATTAAACC	21	This study
	AVMT10R	GATGGCTTGTTTTCGGTTTCC	21	This study
13	AVMT11F	CAAGCCTAGCCCCAACACCAG	21	This study
	AVMT11R	ATGGGGTTAGTCAGTGTAGGC	21	This study
14	AVMT12F	CTCTGACCACCTACACAACCT	21	This study
	AVMT12R	AGTAGTATGTAGAGGGTGT	19	This study
15	AVMT13F	ACTACGAACGGACACACAGCCG	22	This study
	AVMT13R	GAAGGCCAAATTGAGCGGAT	20	This study
16	AVMT14F	ATGACAAGGACGAGCTTAAG	20	This study
	AVMT14R	ATTATTTTTAGTAGGGGGTG	20	This study
17	AVMT15F	GCCAACCTTCATCTCACCATAA	22	This study
	AVMT15R	CTTGTGCGTGGGTTGTCTCGGG	22	This study

which were not found by the tRNA-SE1.21, were identified by proposed secondary structures and anti-codons (Kumazawa and Nishida, 1993; He et al., 2009; Shen et al., 2009a). Gene map of the mitochondrial genome of *T. caboti* was initially generated with OGDRAW (Lohse et al., 2007) and modified manually

RESULTS AND DISCUSSION

Genome composition

The complete sequence of the mtDNA of *T. caboti* is 16,727 bp in length (Figure 1 and Table 1) and was deposited in GenBank with the accession number (GU187969). The size of the Galliformes mtDNA ranges from 16,678 bp (*Syrmaticus reevesii*) to 16,841 bp (*Gallus lafayettii* and *G. sonneratii*). Similar to the typical mitochondrial genomes of vertebrates, the Cabot's tragopan mtDNA consists of 13 typical protein-coding genes, 22 transfer RNA genes, 2 ribosomal RNA genes (*srRNA* and *lrRNA*), and one putative control region (D-loop) (Table 3). As found in other vertebrates, most of these genes were coded on the H-strand, except for one protein-coding gene *(nad6*) and eight tRNA genes ($tRNA^{Gln}$, $tRNA^{Ala}$, $tRNA^{Asn}$, $tRNA^{Cys}$, $tRNA^{Tyr}$, $tRNA^{Ser(UCN)}$, $tRNA^{Pro}$, $tRNA^{Glu}$).

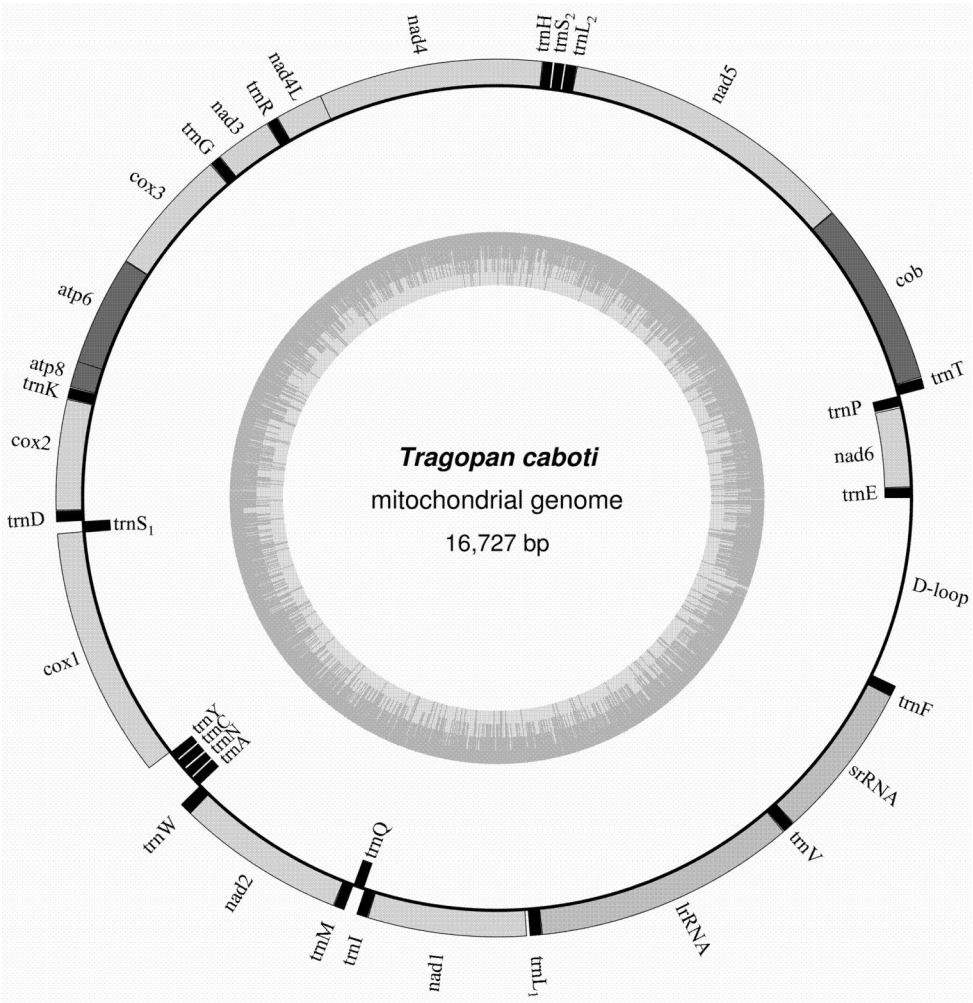

Figure 1. Gene map of the mitochondrial genome of *Tragopan caboti*. Genes encoded on the heavy or light strands are shown outside or inside the circular gene map, respectively. The inner ring displays the GC content. Twenty-two tRNA genes are designated by single-letter amino acid codes. The figure was initially generated with OGDRAW and modified manually.

The overall base composition of H-strand is as follows: A (30.4%), T (24.3%), G (13.7%), C (31.6%), and the A+T content of *T. caboti* (54.7%) is similar to those of other Galliformes (ranging from 52.7 to 56.5%) (Table 1). As in most vertebrates, the overall base composition is skewed against guanine in the *T. caboti* mt genome, which is due to a strong bias against the use of guanine at the third codon position (San Mauro et al., 2004).

Furthermore, one extra nucleotide 'C' is present in *nad3* of *T. caboti*, which is consistently observed in Galliformes except for *Arborophila rufipectus* (He et al., 2009). This extra nucleotide is found in many other birds and some turtles and is thought not to be translated (Mindell et al., 1998; Slack et al., 2003). We found 82 species present in the extra nucleotide of *nad3* within all currently available 107 avian mitochondrial genome sequences retrieved from

Table 3. Localization and features of genes in the mitochondrial genome of *Tragopan caboti*.

Gene/region	Strand	Position		Size (bp)		Codon		Anticodon	Intergenic
		From	To	Nucleotide	Amino acid	Start	Stop[a]		nucleotides[b]
D-loop	H	1	1,177	1,177					0
tRNA^{Phe}	H	1,178	1,245	68				GAA	-1
srRNA	H	1,245	2,222	978					0
tRNA^{Val}	H	2,223	2,295	73				TAC	1
lrRNA	H	2,297	3,910	1,614					0
tRNA$^{Leu(UUR)}$	H	3,911	3,984	74				TAA	16
nad1	H	4,001	4,975	975	324	ATG	TAA		0
tRNA^{Ile}	H	4,976	5,046	71				GAT	6
tRNA^{Gln}	L	5,053	5,123	71				TTG	-1
tRNA^{Met}	H	5,123	5,191	69				CAT	0
nad2	H	5,192	6,230	1,039	346	ATG	T-		0
tRNA^{Trp}	H	6,231	6,308	78				TCA	5
tRNA^{Ala}	L	6,314	6,382	69				TGC	3
tRNA^{Asn}	L	6,386	6,458	75				GTT	2
tRNA^{Cys}	L	6,461	6,526	67				GCA	-1
tRNA^{Tyr}	L	6,526	6,596	71				GTA	1
cox1	H	6,598	8,148	1,551	516	GTG	AGG		-9
tRNA$^{Ser(UCN)}$	L	8,140	8,214	75				TGA	2
tRNA^{Asp}	H	8,217	8,285	69				GTC	1
cox2	H	8,287	8,970	684	227	ATG	TAA		1
tRNA^{Lys}	H	8,972	9,041	70				TTT	2
atp8	H	9,044	9,208	165	54	ATG	TAA		-10
atp6	H	9,199	9,882	684	227	ATG	TAA		-1
cox3	H	9,882	10,665	784	261	ATG	T-		0
tRNA^{Gly}	H	10,666	10,734	69				TCC	0
nad3	H	10,735	11,085	351		ATG	TA-		0
tRNA^{Arg}	H	11,086	11,154	69				TCG	0
nad4L	H	11,155	11,451	297	98	ATG	TAA		-7
nad4	H	11,445	12,822	1,378	459	ATG	T-		0
tRNA^{His}	H	12,823	12,891	69				GTG	1
tRNA$^{Ser(AGY)}$	H	12,893	12,957	65				GCT	1
tRNA$^{Leu(CUN)}$	H	12,959	13,029	71				TAG	0
nad5	H	13,030	14,844	1,815	604	ATG	TAA		-1
cob	H	14,844	15,986	1,143	380	ATG	TAG		2
tRNA^{Thr}	H	15,989	16,057	69				TGT	2
tRNA^{Pro}	L	16,060	16,129	70				TGG	6
nad6	L	16,136	16,657	522	173	ATG	TAG		1
tRNA^{Glu}	L	16,659	16,727	69				TTC	0

[a] "-" Indicates termination codons completed via polyadenylation. [b]Negative values represent overlapping nucleotides.

GenBank (data not shown). However, the extra 'C' of *nad3* was not observed in all available nine mtDNA sequences from Passeriformes. Although Russell and Beckenbach (2008) put forward a hypotheses that certain mitochondrial translation systems have the ability to tolerate frameshift insertions using programmed translational frameshifting, the function of the extra "C" in *nad3* and its phylogenetic implications are still unclear.

Protein-coding genes: nucleotide composition and codon usage

As shown in Table 1, the total length of the 13 protein-coding genes (PCGs) in *T. caboti* mtDNA is 11,390 bp, accounting for 68.1% of the complete mitogenome. The length of 13 PCGs found in the Galliformes species varies from 11,378 bp (*Numida meleagris*) to 11,403 bp (*Coturnix japonica*). The 13 PCGs found in the *T. caboti* mtDNA are similar in length to most other Galliformes species. The longest PCG of *T. caboti* mtDNA is the *nad5* gene (1815 bp), whereas the shortest is *atp*8 gene (165 bp) (Table 4).

Table 4. Base composition for protein-coding genes found in mitochondrial DNA of *Tragopan caboti*.

Gene	Length (bp)	Proportion of nucleotides (%)					AT skew	GC skew
		A	C	G	T	A+T		
*nad*1	975	27.4	33.6	12.3	26.7	54.1	0.01	-0.46
nad2	1041	32.3	34.6	8.7	24.4	56.7	0.14	-0.60
cox1	1551	26.9	30.2	16.7	26.2	53.1	0.01	-0.29
cox2	684	29.5	29.7	14.6	26.1	55.7	0.06	-0.34
atp8	165	36.4	34.6	5.5	23.6	60.0	0.21	-0.73
atp6	684	27.9	35.7	11.0	25.4	53.4	0.05	-0.53
cox3	784	27.8	32.0	15.6	24.6	52.4	0.05	-0.34
nad3	352	28.7	34.4	12.8	24.2	52.8	0.06	-0.46
nad4L	297	27.6	32.7	13.1	26.6	54.2	0.09	-0.43
nad4	1378	31.7	34.3	10.5	23.5	55.2	0.02	-0.53
nad5	1815	31.0	34.6	11.5	23.1	54.0	0.15	-0.50
cob	1143	27.4	35.1	12.3	25.3	52.7	0.04	-0.48
nad6	522	11.9	10.0	37.4	40.8	52.7	-0.55	-0.58
Average		30.4	33.8	11.9	24.0	54.4	0.03	-0.48

The AT composition at the first codon position is 49.8%. The values of the second and third codon positions are 58.2 and 54.0%, respectively (Table 1). As with most Galliformes species (except for *Bambusicola thoracica*), the AT composition of the second codon position is the highest (Table 1). The bias of the base composition in each protein-coding gene can be described by skewness (Perna and Kocher, 1995), which measures the relative numbers of As to Ts and Gs to Cs, and is calculated as (A% - T%) / (A% + T%) and (G% - C%) / (G% + C%), respectively. Twelve PCGs of *T. caboti* mitochondrial genome (with the exception of *nad6*) have a slight skew of A *vs* T (AT skew = 0.01 to 0.55), and a strong skew of C *vs* G (GC skew = -0.29 to -0.73) (Table 4).

Within the mitochondrial genome of *T. caboti*, there are 3 reading frame overlaps (*cox1* and $tRNA^{ser(UCN)}$ share 9 nucleotides; *atp8* and *atp6* share 10 nucleotides; *nad4L* and *nad4* share 7 nucleotides). Other overlaps are shown in Table 3.

Of the 13 typical PCGs (*nad1-6* and *4L*, *cox1-3*, *atp6* and *atp8*, *cob*), 12 genes (with the exception of *nad6*) are encoded on the heavy strand. All PCGs initiate with ATG, except for *cox*1, which begins with GTG. Four types of stop codons were used by the coding genes, including TAA for *nad1*, *cox2*, *atp8*, *atp6*, *nad4L*, and *nad5*; TAG for *cob* and *nad6*; AGG for *cox1*, and incomplete stop codon T- or TA- for *nad2*, *nad3*, *cox3*, and *nad4*, respectively. The use of an incomplete stop codon T- as a common mechanism for stopping protein translations was also observed in other avian species.

The pattern of codon usage in the *T. caboti* mtDNA was also studied (Table 5). There are 3785 codons for all the 13 protein-coding genes after stop codons are excluded. The most frequently used amino acid was Leu (17.39%), followed by Thr (9.36%), Ile (8.09%), Ser (7.77%), and Ala (7.56%).

Non-coding regions

A total of 17 intergenic sequences, ranging in size from 1 to 1177 bp, were found in the mitochondrial genome of *T. caboti* (Table 1). Among these, the longest non-coding region (1177 bp) was found between $tRNA^{Glu}$ and $tRNA^{Phe}$ genes. The length of control region of Galliformes species varied between 1120 bp (*Alectura lathami*) and 1292 bp (*Gallus lafayettei* and *G. sonneratii*), and ranging in AT content from 55.3% (*A. lathami*) to 60.2% (*C. chinensis*) (Table 1). Based on the distribution of the variable nucleotide positions and differential frequencies

Table 5. Codon usage of 13 protein-coding genes in the mitochondrial genome of *Tragopan caboti*.

Amino acid	Codon	Number	Frequency (%)	Amino acid	Codon	Number	Frequency (%)
Phe	TTT	74	1.95	Tyr	TAT	41	1.08
	TTC	141	3.71		TAC	64	1.69
Leu	TTA	69	1.82	Stop	TAA	7	0.18
	TTG	23	0.61		TAG	3	0.08
	CTT	72	1.90	His	CAT	27	0.71
	CTC	191	5.03		CAC	87	2.29
	CTA	257	6.77	Gln	CAA	84	2.21
	CTG	48	1.26		CAG	10	0.26
Ile	ATT	86	2.27	Asn	AAT	28	0.74
	ATC	221	5.82		AAC	100	2.63
Met	ATA	131	3.45	Lys	AAA	80	2.11
	ATG	35	0.92		AAG	9	0.24
Val	GTT	33	0.87	Asp	GAT	18	0.47
	GTC	43	1.13		GAC	45	1.19
	GTA	67	1.77	Glu	GAA	83	2.19
	GTG	23	0.61		GAG	12	0.32
Ser	TCT	43	1.13	Cys	TGT	4	0.11
	TCC	79	2.08		TGC	21	0.55
	TCA	110	2.90	Trp	TGA	93	2.45
	TCG	6	0.16		TGG	16	0.42
Pro	CCT	31	0.82	Arg	CGT	5	0.13
	CCC	84	2.21		CGC	16	0.42
	CCA	104	2.74		CGA	41	1.08
	CCG	9	0.24		CGG	8	0.21
Thr	ACT	60	1.58	Ser	AGT	14	0.37
	ACC	154	4.06		AGC	42	1.10
	ACA	138	3.64		AGA	0	0
	ACG	3	0.08		AGG	1	0.03
Ala	GCT	49	1.29	Gly	GGT	19	0.50
	GCC	132	3.48		GGC	81	2.13
	GCA	95	2.50		GGA	87	2.29
	GCG	11	0.29		GGG	28	0.74

of the nucleotides, the mitochondrial control region is divided into three domains (Brown et al., 1986; Saccone et al., 1991; Randi and Lucchini, 1998). The nucleotide composition of the *T. caboti* control region was A = 27.36%, T = 31.52%, C = 27.02%, and G = 14.10%, with a bias against G, which is usual for the mtDNA sense strand of vertebrates (Wolstenholme, 1992) (Table 6). The domain I (ETAS, extended termination-associated sequences) contains part A (nt 1-163 in Figure 2) and part B (nt 164-315 in Figure 2). The first conserved block (5'-TACCCCCCCTTTCCCCCCCAGGGGGGGTA-3') in part A has sequence similarity to the "goose hairpin" as described in *Anas caerulescens* by Quinn (1992) (Figure 2). Furthermore, in part A, ETAS1 and ETAS2 are found in positions 64-126 and 124-163 nt, respectively, and overlapped one another by 3 bp, with 67.5 and 44.3% similarity to the consensus mammalian

Table 6. The nucleotide composition of the mitochondrial DNA control region of *Tragopan caboti*.

Region	Position		Nucleotide frequency (%)[a]			
	From	To	A	C	G	T
Domain I	1	315	31.43 (29.52)	28.25 (31.39)	12.70 (14.78)	27.62 (24.30)
Domain II	316	783	16.45 (20.38)	29.27 (29.07)	20.94 (19.34)	33.33 (31.22)
Domain III	784	1177	37.06 (35.49)	23.35 (26.18)	7.11 (7.71)	32.49 (30.62)
Complete CR	1	1177	27.36 (28.46)	27.02 (28.88)	14.10 (13.94)	31.52 (28.71)

[a]The numbers in parentheses indicate the nucleotide frequency of the average avian control region (CR) sequences (Ruokonen and Kvist, 2002).

ETAS1 and ETAS2 (Sbisa et al., 1997). In part B (nt 164-315), a conserved sequence block 1 (CSB1)-like block (5'-TAACTATGAATGGTTACAGGACATA-3') has 73.1% similarity to the CSB1 in domain III (Figure 2). Four conserved sequence boxes in the central domain II (nt 316-783) were identified as boxes C, D, E, F (Figure 2). In domain III (nt 784-1177), a poly(C) sequence (nt 783-796), similar to the O_H (origin of H-strand replication) of mammals, maps just a few nucleotides downstream from the putative CSB1 (nt 807-832) (Figure 2). The CSB domain of *T. caboti* has no obvious tracks of CSB2 and CSB3 (Walberg and Clayton, 1981). The bidirectional light- and heavy-strand transcription promoters (LSP/HSP) (L'Abbe et al., 1991) are found in *T. caboti* (Figure 2).

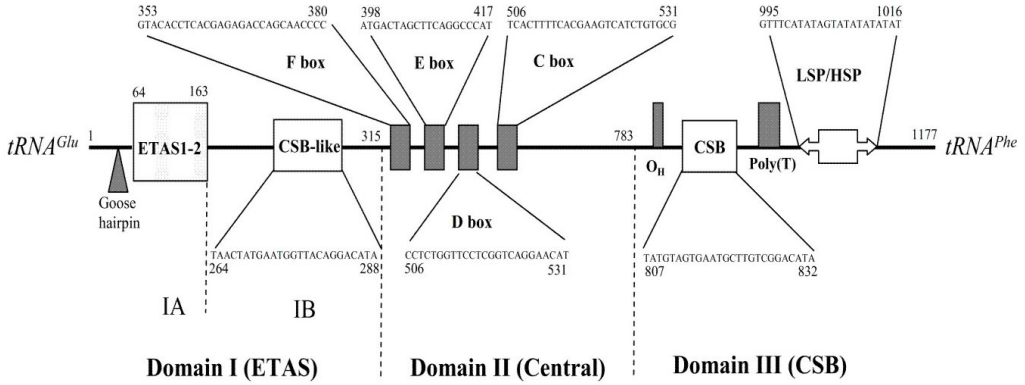

Figure 2. Schematic representation of the organization of the *Tragopan caboti* control region. ETAS = extended termination-associated sequences; F through C boxes = conserved sequence boxes in the central domain; O_H = origin of H-strand replication; CSB = conserved sequence block; CSB-like = a sequence similar to the CSB; LSP = light-strand transcription promoter; HSP = heavy-strand transcription promoter.

Ribosomal and transfer RNA genes

As in all other mt genomes sequenced so far, rRNA genes of *T. caboti* include small subunit rRNA (*srRNA*) and large subunit rRNA (*lrRNA*). The *srRNA* gene is located between $tRNA^{Phe}$ and $tRNA^{Val}$ genes, and the *lrRNA* gene is located between $tRNA^{Val}$ and $tRNA^{Leu(UUR)}$ genes. The lengths of srRNA and lrRNA are 978 and 1546 bp, and A+T content are 52.9 and 54.9%, respectively, which are within the range observed in other Galliformes species (Table 1).

The complete mitochondrial sequence contains 22 tRNA genes, which are interspersed in the genome and range in size from 65 ($tRNA^{Ser(AGY)}$) to 78 ($tRNA^{Trp}$) nucleotides (Table 3). $tRNA^{Lys}$ and $tRNA^{Ser(AGY)}$, which were not found by the tRNAscan-SE, were identified by comparison with *G. gallus* counterparts. All the tRNA gene sequences have the potential to fold into typical cloverleaf secondary structures (Figure 3). The DHU arm contains 3-4 nucleotide pairs, and the TΨC arm contains 3-5 nucleotide pairs.

This is the first complete nucleotide sequences for the mitochondrial genome of Cabot's tragopan *Tragopan caboti* distributed in the Wuyanling National Nature Reserve, China. We also reported the genome organization and codon usage of *T. caboti* mitochondrial DNA. These results will provide basic information for phylogenetic analyses among the Galliformes birds, and especially the *Tragopan* species.

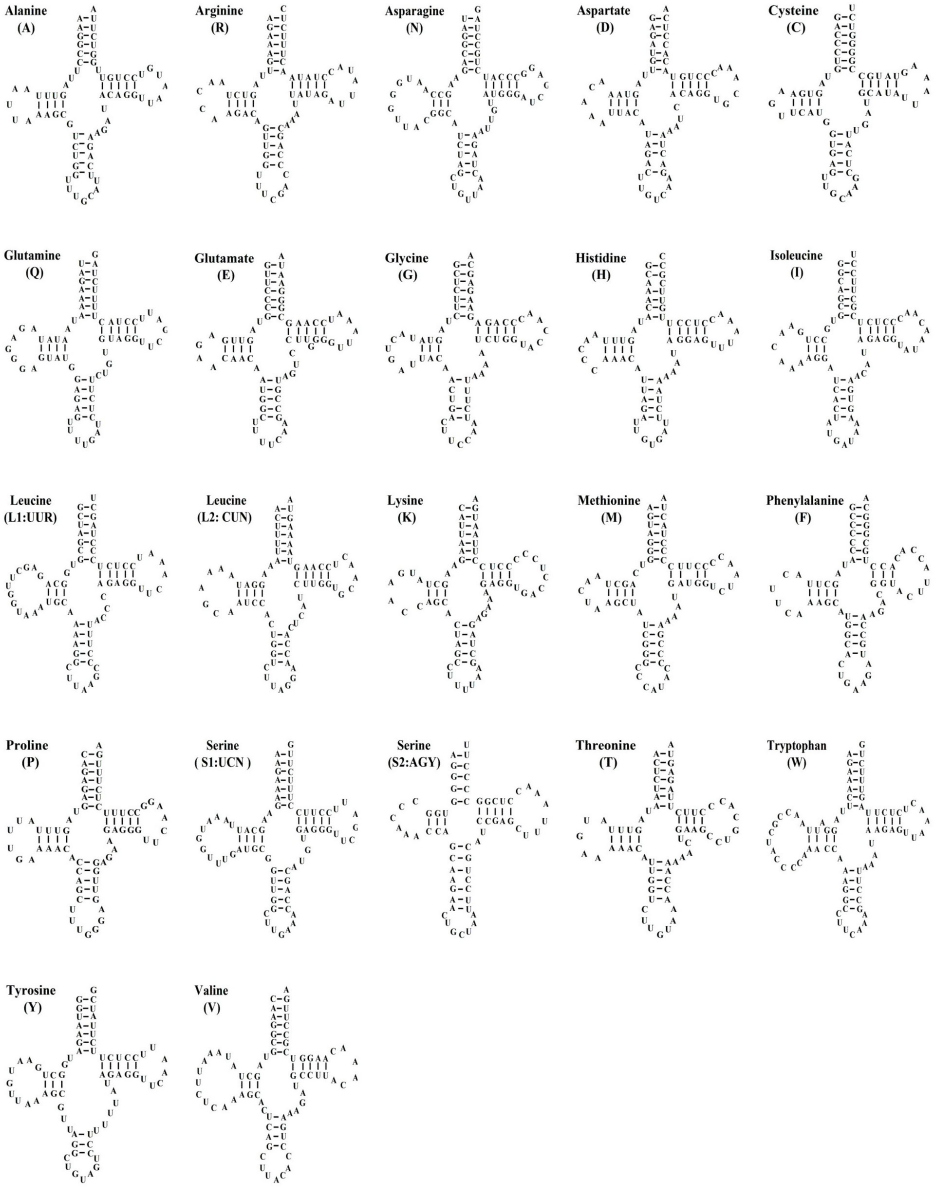

Figure 3. Inferred secondary structures of 22 tRNAs found in *Tragopan caboti* mtDNA.

ACKNOWLEDGMENTS

Research supported by the National Natural Science Foundation of China (#30870172), the State Key Laboratory of Systematic and Evolutionary Botany, China (2009), the Key Program of Natural Science Foundation of the Anhui Higher Education Institutions

(#KJ2008A13ZC), the Natural Science Foundation of Anhui Normal University (#2008xzx14), the Key Lab. of Biotic Environment and Ecology Safety in Anhui Province, and the Program for Innovative Research Team in Anhui Normal University.

REFERENCES

Boore JL (1999). Animal mitochondrial genomes. *Nucleic Acids Res.* 27: 1767-1780.

Brown GG, Gadaleta G, Pepe G, Saccone C, et al. (1986). Structural conservation and variation in the D-loop-containing region of vertebrate mitochondrial DNA. *J. Mol. Biol.* 192: 503-511.

Crowe TM, Bowie RCK, Bloomer P, Mandiwana TG, et al. (2006). Phylogenetics, biogeography and classification of, and character evolution in, gamebirds (Aves: Galliformes): effects of character exclusion, data partitioning and missing data. *Cladistics* 22: 495-532.

del Hoyo J, Elliot A and Sargatal J (1994). Handbook of the Birds of the World. Vol. 2. Lynx Editions, Barcelona, 434-557.

Deng WH and Zheng GM (2004). Landscape and habitat factors affecting Cabot's tragopan *Tragopan caboti* occurrence in habitat fragments. *Biol. Conserv.* 117: 25-32.

Dyke GJ, Gulas BE and Crowe TM (2003). Suprageneric relationships of galliform birds (Aves, Galliformes): a cladistic analysis of morphological characters. *Zoolog. J. Linnean Soc.* 137: 227-244.

Guan X, Silva P, Gyenai KB, Xu J, et al. (2009). The mitochondrial genome sequence and molecular phylogeny of the turkey, *Meleagris gallopavo*. *Anim. Genet.* 40: 134-141.

He L, Dai B, Zeng B, Zhang X, et al. (2009). The complete mitochondrial genome of the Sichuan Hill Partridge (*Arborophila rufipectus*) and a phylogenetic analysis with related species. *Gene* 435: 23-28.

IUCN (2009). IUCN Red List of Threatened Species. Gland, Switzerland. Available at [http://www.iucnredlist.org]. Accessed November 9, 2009.

Kumazawa Y and Nishida M (1993). Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. *J. Mol. Evol.* 37: 380-398.

L'Abbe D, Duhaime JF, Lang BF and Morais R (1991). The transcription of DNA in chicken mitochondria initiates from one major bidirectional promoter. *J. Biol. Chem.* 266: 10844-10850.

Larkin MA, Blackshields G, Brown NP, Chenna R, et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.

Lohse M, Drechsel O and Bock R (2007). Organellar Genome DRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr. Genet.* 52: 267-274.

Lowe TM and Eddy SR (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25: 955-964.

Mindell DP, Sorenson MD and Dimcheff DE (1998). An extra nucleotide is not translated in mitochondrial ND3 of some birds and turtles. *Mol. Biol. Evol.* 15: 1568-1571.

Monroe BL and Sibley CG (1990). A World Checklist of Birds. Yale University Press, New Haven.

Moore WS (1995). Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* 49: 718-726.

Nishibori M, Hayashi T, Tsudzuki M, Yamamoto Y, et al. (2001). Complete sequence of the Japanese quail (*Coturnix japonica*) mitochondrial genome and its genetic relationship with related species. *Anim. Genet.* 32: 380-385.

Nishibori M, Tsudzuki M, Hayashi T, Yamamoto Y, et al. (2002). Complete nucleotide sequence of the *Coturnix chinensis* (blue-breasted quail) mitochondrial genome and a phylogenetic analysis with related species. *J. Hered.* 93: 439-444.

Nishibori M, Hayashi T and Yasue H (2004). Complete nucleotide sequence of *Numida meleagris* (Helmeted guineafowl) mitochondrial genome. *J. Poult. Sci.* 41: 259-268.

Nishibori M, Shimogiri T, Hayashi T and Yasue H (2005). Molecular evidence for hybridization of species in the genus *Gallus* except for *Gallus varius*. *Anim. Genet.* 36: 367-375.

Perna NT and Kocher TD (1995). Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J. Mol. Evol.* 41: 353-358.

Quinn TW (1992). The genetic legacy of Mother Goose - phylogeographic patterns of lesser snow goose *Chen caerulescens caerulescens* maternal lineages. *Mol. Ecol.* 1: 105-117.

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.

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

Russell RD and Beckenbach AT (2008). Recoding of translation in turtle mitochondrial genomes: programmed frameshift mutations and evidence of a modified genetic code. *J. Mol. Evol.* 67: 682-695.

Saccone C, Pesole G and Sbisa E (1991). The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. *J. Mol. Evol.* 33: 83-91.

Sambrook J and Russell DW (2001). Molecular Cloning: A Laboratory Manual. 3rd edn. Cold Spring Harbor Laboratory Press, New York.

San Mauro D, Garcia-Paris M and Zardoya R (2004). Phylogenetic relationships of discoglossid frogs (Amphibia:Anura:Discoglossidae) based on complete mitochondrial genomes and nuclear genes. *Gene* 343: 357-366.

Sbisa E, Tanzariello F, Reyes A, Pesole G, et al. (1997). Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. *Gene* 205: 125-140.

Shadel GS and Clayton DA (1997). Mitochondrial DNA maintenance in vertebrates. *Annu. Rev. Biochem.* 66: 409-435.

Shen X, Tian M, Liu Z, Cheng H, et al. (2009a). Complete mitochondrial genome of the sea cucumber *Apostichopus japonicus* (Echinodermata: Holothuroidea): the first representative from the subclass Aspidochirotacea with the echinoderm ground pattern. *Gene* 439: 79-86.

Shen YY, Shi P, Sun YB and Zhang YP (2009b). Relaxation of selective constraints on avian mitochondrial DNA following the degeneration of flight ability. *Genome Res.* 19: 1760-1765.

Slack KE, Janke A, Penny D and Arnason U (2003). Two new avian mitochondrial genomes (penguin and goose) and a summary of bird and reptile mitogenomic features. *Gene* 302: 43-52.

Slack KE, Delsuc F, McLenachan PA, Arnason U, et al. (2007). Resolving the root of the avian mitogenomic tree by breaking up long branches. *Mol. Phylogenet. Evol.* 42: 1-13.

Walberg MW and Clayton DA (1981). Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. *Nucleic Acids Res.* 9: 5411-5421.

Wolstenholme DR (1992). Animal mitochondrial DNA: structure and evolution. *Int. Rev. Cytol.* 141: 173-216.

Wyman SK, Jansen RK and Boore JL (2004). Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* 20: 3252-3255.

Zhang JF, Nie LW, Wang Y and Hu LL (2009). The complete mitochondrial genome of the large-headed frog, *Limnonectes bannaensis* (Amphibia: Anura), and a novel gene organization in the vertebrate mtDNA. *Gene* 442: 119-127.

Zhang Y and Zheng G (2007). A population viability analysis (PVA) for Cabot's tragopan (*Tragopan caboti*) in Wuyanling, south-east China. *Bird Conserv. Int.* 17: 151-161.