动 物 学 研 究 2010，Apr. 31(2)：122−130 CN 53-1040/Q ISSN 0254-5853
Zoological Research DOI：10.3724/SP.J.1141.2010.02122

Complete Mitochondrial Genome of the Red Fox (*Vuples vuples*) and Phylogenetic Analysis with Other Canid Species

ZHONG Hua-Ming[1], ZHANG Hong-Hai[1,*], SHA Wei-Lai[1], ZHANG Cheng-De[1], CHEN Yu-Cai[2]

(1. *College of Life Science, Qufu Normal University, Qufu* 273165, *China;* 2. *Ji'nan Paomaling World of Wildlife, Ji'nan* 250000, *China*)

Abstract: The whole mitochondrial genome sequence of red fox (*Vuples vuples*) was determined. It had a total length of 16 723 bp. As in most mammal mitochondrial genome, it contained 13 protein coding genes, two ribosome RNA genes, 22 transfer RNA genes and one control region. The base composition was 31.3% A, 26.1% C, 14.8% G and 27.8% T, respectively. The codon usage of red fox, arctic fox, gray wolf, domestic dog and coyote followed the same pattern except for an unusual ATT start codon, which initiates the NADH dehydrogenase subunit 3 gene in the red fox. A long tandem repeat rich in AC was found between conserved sequence block 1 and 2 in the control region. In order to confirm the phylogenetic relationships of red fox to other canids, phylogenetic trees were reconstructed by neighbor-joining and maximum parsimony methods using 12 concatenated heavy-strand protein-coding genes. The result indicated that arctic fox was the sister group of red fox and they both belong to the red fox-like clade in family Canidae, while gray wolf, domestic dog and coyote belong to wolf-like clade. The result was in accordance with existing phylogenetic results.

Key words: Red fox; Mitochondrial genome; Canidae; Phylogenetic analysis

赤狐线粒体全基因组及系统发育分析

钟华明[1]，张洪海[1,*]，沙未来[1]，张承德[1]，陈玉才[2]

（1. 曲阜师范大学 生命科学学院，山东 曲阜 273165；2. 济南跑马岭野生动物世界，山东 济南 250000）

摘要：测定了赤狐的线粒体基因组全序列，总长度为16723 bp，碱基组成为：31.3% A、26.1% C、14.8% G、27.8% T。和大多数哺乳动物一样，赤狐的线粒体全基因组包含13个蛋白质编码基因、2个核糖体RNA基因、22个转运RNA基因和1个控制区。除ND3基因起始密码子为不常见的ATT外，赤狐与北极狐、狼、家犬、郊狼的线粒体蛋白质编码遵循相同模式。在控制区的保守序列区段1和2之间发现一段较长的富含AC的随机重复序列。为了验证赤狐与其他犬科动物的系统发育关系，利用12个重链蛋白质编码基因，分别通过邻接法和最大简约法构建了系统发育树。结果表明：赤狐与北极狐是姐妹群，它们在犬科中都属于赤狐型分支，而灰狼、家犬和郊狼则属于狼型分支，与现有的系统进化研究结果一致。

关键词：赤狐；线粒体全基因组；犬科；系统发育分析
中图分类号：Q344.13;Q786;Q959.838.09 **文献标志码**：A **文章编号**：0254-5853-(2010)02-0122-09

Mammalian mitochondrial genomes typically have a set of 13 protein-coding genes, two ribosome RNA genes (12*S* RNA and 16*S* RNA) and 22 transfer RNA genes. The gene order is highly conserved among most vertebrates (Boore, 1999). The two strands that make up the genome are commonly known as the heavy strand (H-strand) and the light strand (L-strand). Twelve of the 13 protein-coding genes locate on the H-strand and only ND6 gene are located on the L-strand. Vertebrates usually contain two non-coding regions, the major of which contains mitochondrial replication and transcription promoters. Consequently, it is known as the control region (CR) (Wolstenholme, 1992; Boore, 1999). The small non-coding region is the origin of replication

Received date: 2009-09-11; Accepted date: 2010-01-20
Foundation items: National Natural Science Foundation of China (30370218); the program for New Century Excellent Talents in University (NCET-07-0507); the Project of Science and Technology Development Plan in Shandong Province (2007GG2009011); Shandong Science Fund for Distinguished Young Scholars (2005BS02005).
收稿日期：2009-09-11；接受日期：2010-01-20
* 通讯作者 (Corresponding author)，E-mail：zhanghonghai67@126.com
第一作者简介：钟华明(1984-), 男，硕士，E-mail:monzhm@126.com

of the L-strand (OL), thought to have a functional role in replication (Shadel & Clayton, 1997).

Belonging to Canivora, Canidae, Vuples, red fox (*Vuples vuples*) is one of the worldwide distributed mammals. Comparing complete animal mitochondrial genome sequences is now common for phylogenetic reconstruction and as a model for genome evolution (Wei et al, 2008). Mitochondrial DNA sequences were also used to clarify phylogenetic relationships within Canidae (Geffen et al, 1992; Wayne et al, 1997). Sequence data from three mitochondrial genes (Wayne et al, 1997) suggest that Canidae falls to four monophyletic groups: (1) the wolf- and jackal-like canids; (2) the red fox-like canids; (3) the South American foxes; and (4) the maned wolf (*Chysocyon brachyurus*) and bush dog (*Speothos venaticus*). Analysis of various morphological and mitochondrial DNA data (Zrzavy & Ricankova, 2004) as well as phylogenetic analysis using six nuclear loci combined with mitochondrial data (Bardeleben, 2005) both agree with the three clades, including the red fox-like canids, the South American foxes, and the wolf-like canids.

This paper reports a complete mitochondrial genome of red fox, which was also compared with that of other canids to discuss the red fox mitochondrial genome structure and evolution. A phylogenetic tree was reconstructed to confirm the evolution position of red fox in Canidae.

1 Materials and Methods

1.1 Samples and DNA extracting

A blood sample of 200 μL from a female red fox was obtained from Beijing Zoo. Total genomic DNA was extracted following the method of Sambrook & Russell (2001) and the obtained DNA solution was preserved under −20 ℃ for next use.

1.2 PCR and sequencing

Five pairs of primers were designed based on the sequences of wolf and coyote, which were available online with the accession number of NC008092 and NC008093, respectively. Primer sequences are shown in Tab. 1.

Tab. 1 Five primer pairs for the first PCRs

Primer pair	Primer	Sequence(5'→3')
1	1-F	TCCCTCTAGAGGAGCCTGTTC
	1-R	GGGTATGGGCCCGATAGCTT
2	2-F	GGCGGATAAAAGAGTTACTTTGATAGAG
	2-R	GCGAATTTAACTTTGACAAAGTCATGT
3	3-F	GAAGAAAGGAAGGAATCGAACC
	3-R	GCGTAGGGATGATAATTTTTAGCATT
4	4-F	GTATTTGCTGCCTGCGAAGC
	4-R	TAGTGGTGGGATTGGTTGTGC
5	5-F	GGGTATTGCTCAGTAGCCATAGC
	5-R	GGTTTGCTGAAGATGGCGGTATAT

PCR was implemented in a 50 μL system including 1 μL DNA, 5 μL 10 × Long-PCR Buffer, 4 μL dNTP mixtures (2.5 mmol/L each NTP), 15 mmol $MgCl_2$; pfu polymerase 1 μL (5 U/μL); 2 μL primers (10 μmol/L). After degeneration of 95℃ for 5 min, 35 cycles of 95℃ degeneration for 30 s, 57℃ annealing for 30 s, and 72℃ extension for 4.5 min were ran. Final extension was ran at 72℃ for 10 min. PCR products were purified from a 1% agarose gel and were then used as DNA templates in secondary PCR. Secondary PCR was conducted with 10 pairs of primers (Tab. 2). PCR conditions were: initial degeneration of 95℃ for 5 min, then 35 cycles of 95℃ degeneration for 30 s, 57℃ annealing for 30 s, 72℃ extension for 2.5 min and a final extension step of 72℃ for 4.5 min. Reacting system and purifying methods are the same as the first PCR. The purified PCR products were then sent out for sequencing.

1.3 Sequence assembly, annotation and analysis

Three times of sequencing were performed to make sure correct target sequences were amplified. The DNAMAN 6.0 program (Lynnon Biosoft, Quebec, Canada) was used for editing and assembling the sequence. The positions of protein coding genes, rRNA genes and non-coding regions were located through alignment with that of wolf and coyote using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Start and stop codons were identified using the vertebrate mitochondrial code except some potential incomplete stop codons.

Tab. 2 Ten primer pairs used for secondary PCRs

Primer pair	Primer	Sequence(5'→3')
1	1-F	TCCCTCTAGAGGAGCCTGTTC
	1-R	TCCGAGGTCACCCCAACC
2	2-F	GACGAGAAGACCCTATGGAGC
	2-R	GGGTATGGGCCCGATAGCTT
3	3-F	GGCGGATAAAAGAGTTACTTTGATAGAG
	3-R	GCCTACTATACCGGCTCATGC
4	4-F	GCTCAGCCATTTTACCTATGTTC
	4-R	GCGAATTTAACTTTGACAAAGTCATGT
5	5-F	GAAGAAAGGAAGGAATCGAACC
	5-R	GCGAAGAGTTGTAGTGAAATCATAT
6	6-F	GCTACCTAATGACCCACCAAAC
	6-R	GCGTAGGGATGATAATTTTTAGCATT
7	7-F	GTATTTGCTGCCTGCGAAGC
	7-R	CGCTTATCTGGAGTTGCACC
8	8-F	CCGCAAGAACTGCTAATTCATG
	8-R	TAGTGGTGGGATTGGTTGTGC
9	9-F	GGGTATTGCTCAGTAGCCATAGC
	9-R	GCAGAATTTCAGCTTTGGGTG
10	10-F	CGCGATGAAGAGTCTTTGTAGTAT
	10-R	GGTTTGCTGAAGATGGCGGTATAT

21 tRNAs genes were identified by tRNAscan-SE Search Server v.1.21 (http://lowelab.ucsc.edu/tRNAscan-SE/). The tRNASer (AGC) gene not found by tRNAscan-SE was determined by comparison with the two canids mentioned above. RNA structure 4.6 (Lowe & Eddy, 1997) was used to speculate the second structure of the 22 tRNAs. MEGA v.4.0 (Tamura et al, 2007**)** was used to compute base composition, substitution and codon usage.

The mitochondrial genome sequence of the red fox was submitted into GenBank with the accession number of GQ374180.

1.4 Phylogenetic analysis

Four complete protein sequences available for *Canidae* were downloaded from GenBank, and tiger (*Panthera tigris*) was used as an outgroup for phylogenetic analyses (Tab. 3). All protein coding genes except the ND6 gene were chosen to reconstruct phylogenetic tree. Multiple alignments were performed by ClustalX 1.83 (Thompson et al, 1997) with the default parameters. The results of alignments were also carefully checked and edited by eye. Phylogenetic analyses were conducted using neighbor-joining (NJ) and maximum parsimony (MP). The NJ analysis was performed in MEGA 4.0 using Kimura 2-parameter as the nucleotide substitution model. Statistical confidence was assessed by bootstrap analysis with 1000 replications. The MP analysis was performed in PAUP*4.0b10 (Swofford, 2003) using a heuristic search. The robustness of the

Tab. 3 The taxa and sequence's accession number of the six species used for phylogenetic analyses

Species	Common name	Accession No. and references
Vuples vuples	Red fox	GQ374180,This paper
Alopex lagopus	Arctic fox	AH014073, Delisle & Strobeck, 2005
Canis lupus	Gray wolf	NC_008092, Bjornerfeldt et al, 2006
Canis lupus familiaris	Dog	NC_002008, Kim et al, 1998
Canis latrans	Coyote	NC_008093, Bjornerfeldt et al, 2006
Panthera tigris	Tiger	EF551003, unpublished

tree was tested with 100 bootstrap replications.

2 Results

2.1 Mitochondrial genome structure

The total length of red fox mitochondrial genome is 16 723 bp, a little shorter than that of gray wolf(1 6729 bp) and coyote (16 724 bp). The red fox mitochondrial genome shares high similarity with those of three other canids: gray wolf (84.93%), domestic dog (84.86%) and coyote (84.87%). Similar to other animals (Boore, 1999), the red fox mitochondrial genome has 13 protein-coding genes, 22 tRNA genes, two rRNA genes (12*S* rRNA, 16*S* rRNA) and one control region (CR). Except for the protein coding gene ND6 and 8 tRNA genes [tRNAGln, tRNAAla, tRNAAsn, tRNACys, tRNATyr, tRNASer (UCN), tRNAGlu, tRNAPro] encoded at the L-strand, the rest genes are encoded on the H-strand. Overall organization of the genome is comparable to the situation of other vertebrates(Tab. 4). As shown in Fig. 1, the genes are so compactly arranged that few gaps were found between them and some genes even overlap each other. The longest gap was found between COII and tRNALys by 16 nucleotides, while ATPase8 and ATPase6 overlap each other by 33 nucleotides, which is the longest one.

The overall base composition of the red fox mitochondrial genome is 31.3% A, 26.1% C, 14.8% G and 27.8% T, respectively (Tab. 5).The consequent order would be A>T>C>G, which was in coincidence with the order of other mammals. The A+T content is 59.1%, which indicated an A−T rich in the red fox mitochondrial genome.

2.2 Protein coding genes

Except ATA for ND2, ND5 and ATT for ND3, other protein coding genes use the start codon ATG. When it comes to stop codons, ND1, COI, COII, ATPase8, ATPase6, ND4L, ND5 and ND6 terminate with TAA, while ND2 and Cytb terminated with TAG and AGA respectively. Besides, three incomplete stop codons were detected: T (COIII), TA (ND3) and T (ND4). Codon usage analysis (Tab. 6) indicates obvious codon bias and CUA (Leu), AUA (Met), AUU (Ile) are the most used three codons. Codon degenerate was also found by codon usage analysis. Both four-fold degenerate codons (e.g. codons specify Ala and Pro) and two-fold degenerate codons (e.g. codons specify Asn and Tyr) were found. Leu and Ser are found to be specified by both four-fold degenerate codons and two-fold degenerate codons.

The total length of 13 protein coding genes is 11 411 bp, accounting for 68.24% of the complete sequence of the red fox mitochondrial genome. As shown in Fig.1, three pairs of protein coding genes, ATP8/ATP6, ND4L/ND4 and ND5/ND6 overlap each other by 33 bp, 7 bp and 17 bp respectively. For the 13 protein-coding genes of the red fox mitochondrial DNA, the nucleotide distribution at three codon positions differ significantly (P<0.01).A strong bias against G (only 9.2%) at the third codon position was observed. The frequency of A+T in all protein-coding genes ranges from 53.6% to 61.6%.

2.3 Ribosomal RNA and transfer RNA genes

The 12S rRNA gene of the red fox mitochondrial genome was located between $tRNA^{Phe}$ gene and $tRNA^{Val}$ gene with a length of 957 bp, longer than that of gray wolf (954 bp) and coyote (955 bp) due to different numbers of insertions. The base composition is 35.6% A, 22.2% C, 18.6% G and 23.6% T respectively. The 16S rRNA gene was located between $tRNA^{Val}$ gene and $tRNA^{Leu(UUR)}$ gene. The length is 1579 bp, which was identical to coyote but shorter than gray wolf (1580 bp). The base composition is 36.8% A, 20.6% C, 17.6% G and 25.0% T respectively. The computed A+T content of the 12S rRNA and 16S rRNA gene are 59.2% and 61.8%, which also indicates an A−T rich.

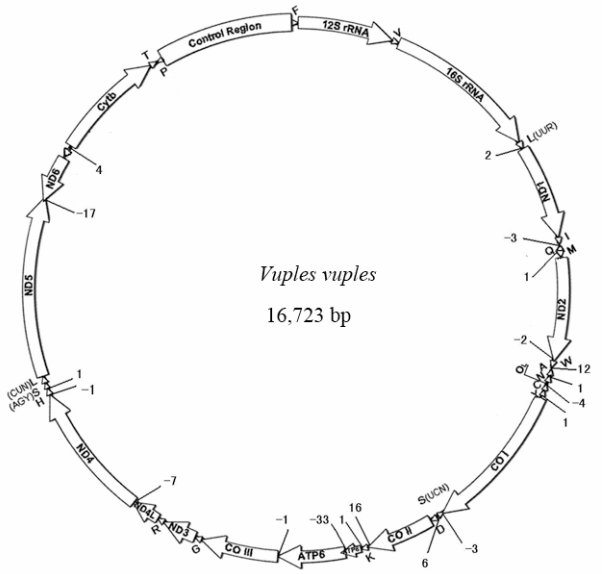

Fig. 1 Organizations of red fox mitochondrial genome. Transfer RNA genes are depicted by their one-letter amino acid codes

Numbers indicate non-coding nucleotides between genes or gene overlap (negative values). Arrows indicate orientation on (+) strand (clockwise) or (−) strand (counterclockwise).

Tab. 4 Organizations and characteristics of red fox mitochondrial genome

Name of gene	Position number		Size(bp)	Intervals	Codon		Anti-codon	Strand
	Start	Stop			Start	Stop		
$tRNA^{Phe}$	1	69	69	0			GAA	H
12SrRNA	70	1026	957	0				H
$tRNA^{Val}$	1027	1093	67	0			TAC	H
16S rRNA	1094	2672	1579	0				H
$tRNA^{Leu(UUR)}$	2673	2747	75	2			TAA	H
ND1	2750	3706	957	0	ATG	TAA		H
$tRNA^{Ile}$	3706	3774	69	-3			GAT	H
$tRNA^{Gln}$	3772	3845	74	1			TTG	L
$tRNA^{Met}$	3847	3916	70	0			CAT	H
ND2	3917	4960	1044	-2	ATA	TAG		H
$tRNA^{Trp}$	4959	5026	68	12			TCA	H
$tRNA^{Ala}$	5039	5107	69	1			TGC	L
$tRNA^{Asn}$	5109	5179	71	0			GTT	L
OL	5180	5217	38	-4				L
$tRNA^{Cys}$	5214	5281	68	0			GCA	L
$tRNA^{Tyr}$	5282	5349	68	1			GTA	L
COI	5351	6895	1545	-3	ATG	TAA		H
$tRNA^{Ser(UCN)}$	6893	6961	69	6			TGA	L
$tRNA^{Asp}$	6968	7035	68	0			GTC	H
COII	7036	7719	684	16	ATG	TAA		H
$tRNA^{Lys}$	7737	7803	67	1			TTT	H
ATPase8	7805	8008	204	-33	ATG	TAA		H
ATPase6	7966	8646	681	-1	ATG	TAA		H
COIII	8646	9429		0	ATG	T		H
$tRNA^{Gly}$	9430	9497	68	0			TCC	H
ND3	9498	9844	347	0	ATT	TA		H
$tRNA^{Arg}$	9845	9913	69	0			TCG	H
ND4L	9914	10210	297	-7	ATG	TAA		H
ND4	10204	11582	1379	-1	ATG	T		H
$tRNA^{His}$	11582	11650	69	1			GTG	H
$tRNA^{Ser(AGY)}$	11652	11710	59	0				H
$tRNA^{Leu(CUN)}$	11711	11780	70	0			TAG	H
ND5	11781	13601	1821	-17	ATA	TAA		H
ND6	13585	14112	528	0	ATG	TAA		L
$tRNA^{Glu}$	14113	14181	69	4			TTC	L
Cytb	14186	15325	1140	0	ATG	AGA		H
$tRNA^{Thr}$	15326	15395	70	0			TGT	H
$tRNA^{Pro}$	15395	15460	66	0			TGG	L
D-loop	15461	16723	1263	0				

The length of the red fox mitochondrial tRNA genes range from 59 bp to 75 bp. Among the adjacent tRNA genes, only tRNAIle and tRNAGln overlap each other by 3 bp. Except for tRNASer(AGY) lack the dihydrouridine stem and loop (DHU Stem and loop), other tRNAs can fold into typical cloverleaf secondary structure (Fig.2). The average A+T content in all tRNA genes is 63.6%, higher than that of protein-coding and rRNA genes (Tab.

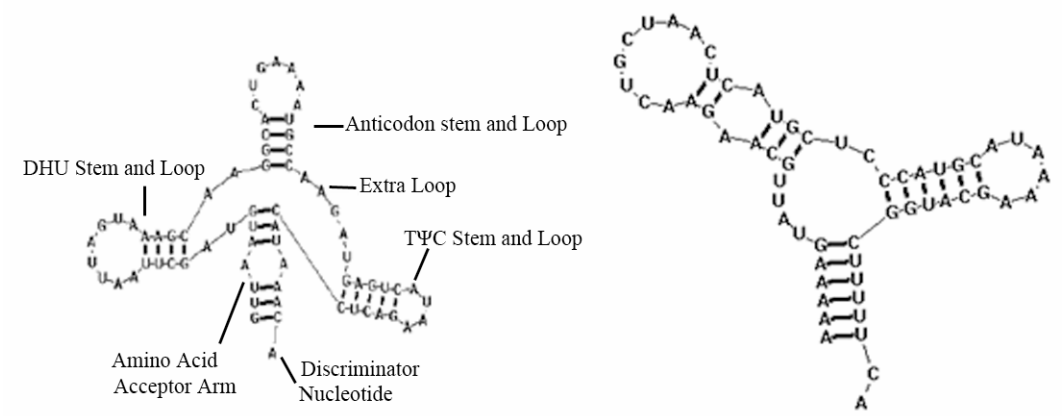

Fig. 2 Typical cloverleaf secondary structure (Left, e.g. $tRNA^{Phe}$) and the secondary structure of $tRNA^{Ser(AGY)}$ (Right)
Six major elements are included in the typical structure of tRNA: Anticodon stem and loop, TΨC Stem and loop, Amino Acid Acceptor Arm, Extra Loop, Discriminator Nucleotide and finally a DHU Stem and loop. Whereas the secondary structure of $tRNA^{Ser(AGY)}$ lack of the DHU Stem and loop.

Tab. 5 Comparison of base compositions in different genes (or regions) of the red fox mitochondrial genome

Gene/region	Base composition (%)				
	A	C	G	T	A+T
Protein coding					
1st	31.1	24.7	21.8	22.5	53.6
2nd	19.4	26.2	12.2	42.2	61.6
3rd	36.6	30.1	9.2	24.1	60.7
Total	29	27	14.4	29.6	58.6
tRNA	34.7	20.6	15.9	28.9	63.6
12*S* rRNA	34.7	22.2	18.6	23.6	59.2
16*S* rRNA	34.7	20.6	17.6	25	61.8
D-loop	34.7	29.2	15.8	25.1	55.5
Overall	34.7	26.1	14.7	27.7	59.1

5).

2.4 Non-coding regions

Locating between $tRNA^{Pro}$ and $tRNA^{Phe}$ genes, the CR of the red fox mitochondrial genome has a length of 1263 bp, shorter than that of gray wolf (1269 bp) and coyote (1266 bp). The base composition is 29.9% A, 29.2% C, 15.8% G and 25.1% T, respectively. The extended termination associated sequence (ETAS) located at the 5' end of the CR and contain a termination associated sequence (TAS). This domain contains a core motif of and its complement and fold into a 'hairpin' structure; the central conserved domain (CD) domain includes five conserved blocks (CSB-F, E, D, C, B); the conserved sequence block (CSB) domain contains three conserved blocks (CSB-1, 2, 3). A region of tandem repeat region called RS3 (Douzery & Randi, 1997) was determined by Tandem Repeats Finder program (Benson, 1999) at region 645−981bp. Similar to some mammals investigated so far (Fumagalli et al, 1996; Hoelzel et al, 1994; Ketmaier & Bernardini, 2005), it was located between CSB-1 and CSB-2. Particularly, the base composition of A+C is 70.0%, indicating an A-C rich in this region. The single repeat is a motif of TACACACG (8 bp). There are totally 37 copies, 32 of which are complete repeats. Besides, two repeats show the same transition of A→G, C→T and G→A at position 4, 5 and 8 respectively and the sequence is TAC**GT**AC**A**. Two repeats show the same transition of T→C, A→G and C→T at position 1, 4, 5 respectively and the sequence is **C**AC**GT**ACG. One repeat shows a transition of A→G and C→T at position 4 and 5. Among this tandem repeat region, we found an alternative array of repeats. In this case there were 14 repeats with the motif of ACGTACACACGTACACACGT (20 bp). Except for one repeat with a transition of G→A, T→C, A→G and C→T at position 3, 4, 7 and 8 respectively, the other 13 repeats follow the complete repeat.

Tab. 6 Codon usage and codon bias (RSCU) in red fox mitochondrial genome

AA	Codon	Count	RSCU	AA	Codon	Count	RSCU	AA	Codon	Count	RSCU	AA	Codon	Count	RSCU
Ala	GCA	97	1.45	Glu	GAA	64	1.36	Lys	AAA	84	1.65	Ter	AGA	1	0.4
	GCC	93	1.39		GAG	30	0.64		AAG	18	0.35		AGG	0	0
	GCG	15	0.22	Gly	GGA	82	1.53	Met	ATA	189	1.51		TAA	8	3.2
	GCT	63	0.94		GGC	50	0.93		ATG	61	0.49		TAG	1	0.4
Arg	CGA	40	3.64		GGG	41	0.76	Phe	TTC	137	1.2	Thr	ACA	108	1.44
	CGC	7	0.64		GGT	42	0.78		TTT	92	0.8		ACC	101	1.35
	CGG	8	0.73	His	CAC	75	1.46	Pro	CCA	68	1.4		ACG	11	0.15
	CGT	10	0.91		CAT	28	0.54		CCC	68	1.4		ACT	79	1.06
Asn	AAC	84	1.16	Ile	ATC	154	0.91		CCG	7	0.14	Trp	TGA	88	1.69
	AAT	61	0.84		ATT	183	1.09		CCT	51	1.05		TGG	16	0.31
Asp	GAC	48	1.28	Leu	CUA	230	2.34	Ser	AGC	33	0.69	Tyr	TAC	73	1.08
	GAT	27	0.72		CUC	94	0.96		AGT	21	0.44		TAT	62	0.92
Cys	TGC	19	1.52		CUG	63	0.64		TCA	86	1.8	Val	GTA	74	1.58
	TGT	6	0.48		CUU	81	0.83		TCC	67	1.4		GTC	48	1.03
Gln	CAA	73	1.64		TTA	94	0.96		TCG	14	0.29		GTG	25	0.53
	CAG	16	0.36		TTG	27	0.28		TCT	66	1.38		GTT	40	0.86

The putative light strand replication (OL) was detected in the region WANCY, which contains five tRNA genes: $tRNA^{Trp}$, $tRNA^{Ala}$, $tRNA^{Asn}$, $tRNA^{Cys}$ and $tRNA^{Tyr}$. This region is 38 bp long and overlaps the $tRNA^{Cys}$ gene by 4 bp. It can fold into a stem-loop secondary structure with a 12 bp stem and a loop of 14 nucleotides.

2.5 Phylogenetic analysis

The same topology was obtained from both NJ tree and MP tree (Fig. 3). The bootstrap values are very high and all up to 100. The red fox and arctic fox fall into one clade, while gray wolf, domestic dog and coyote fall into another one. The red fox is a sister group with arctic fox. The coyote diverged after the emergence of red fox and arctic fox and prior to gray wolf and dog.

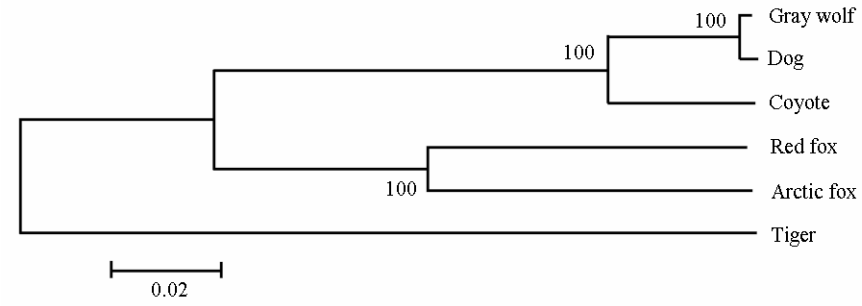

Fig. 3 Phylogenetic relationships by neighbor-joining (NJ) and maximum parsimony (MP) methods inferred from concatenated 12 protein-coding genes

Same topology was obtained and all bootstrap values are 100.

3 Discussion

3.1 Mitochondrial genome evolution

The size and structure of mitochondrial genome of vertebrates are highly conserved. The length ranges from 15.7 to 19.5 kb and the red fox, gray wolf and coyote all fall within this. There are only several nucleotides differences in length among these three canids. The arrangement of multiple genes in red fox is in line with most vertebrates and the overall organization of the

canids genome is the same, which indicates that the mitochondrial genome are evolutionary highly conserved. Base composition bias was determined in the red fox mitochondrial genome. A−T rich was found not only in whole genome level, but also in rRNA and tRNA genes. Therein, A+T frequency in tRNA genes is the highest one. However, the tandem repeat region determined in the CR is A−C rich.

Comparison of the codon usage of red fox with arctic fox, gray wolf, dog and coyote suggests that they follow the same pattern except the unusual ATT start codon in the red fox. The unusual ATT start codon was also observed in fur seal (Lin et al, 2002) and giant panda (Peng, 2007). Three types of incomplete stop codons were determined in the red fox mitochondrial genome. The phenomenon of terminating with incomplete stop codons was also found in other mammals (Gissi et al, 1998; Kim et al, 1998; Ursing & Arnason et al, 1998; Xu et al, 1996; Peng et al, 2007; Gu et al, 2007). Codon bias and degeneration which occur in both nuclear and mitochondrial genome were found.

The length of tRNA genes of red fox follows the vertebrates' typical 59 − 75 bp. Still similar to other vertebrates, all tRNAs can fold into typical cloverleaf secondary structure except tRNASer$^{(AGY)}$,which lacks the DHU stem and loop. Moreover, the anti-codons of the 22 tRNA genes are the same as in other vertebrates.

Extensive tandem repeats in the mtDNA control region have been found in many vertebrate species. Xiao (2006) suggest that tandem repeats may have resulted from slipped-strand mispairing during mtDNA replication and they may play an important role in regulating transcription and replication of the mitochondrial genome. Different unit size and copy number of tandem repeat region are responsible for size variation among CR sequences from different species, or even from individuals of the same species (Zhang & Hewitt 1997). However, the CR sizes of these three canids vary little, which indicate that they have similar tandem repeat region size and our alignments proved this. In despite, they have different patterns of repetition.

3.2 Phylogenetic analysis

Phylogenetic results determined by 12 concatenated protein-coding genes were in accordance with the two of the three well-defined clades within Canidae: the red fox-like canids and the wolf-like canids. However, the 12 protein-coding genes used in our phylogenetic analysis all come from one mitochondrial genome. Namely, they are inherited as a single linkage group (haplotype) and provide only one independent estimate of the species tree (Moore, 1995). Thus, further researches by nuclear sequence or combined nuclear-mitochondrial data are required to validate or reinforce our results (Yu & Zhang, 2006).

References:

Bardeleben C, Moore RL, Wayne RK. 2005. A molecular phylogeny of the Canidae based on six nuclear loci [J]. *Mol Phylogenet Evol*, **37**: 815-831.

Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences [J]. *Nucleic Acids Res*, **27**: 573-580.

Bjornerfeldt S, Webster MT, Vila C. 2006. Relaxation of selective constraint on dog mitochondrial DNA following domestication [J]. *Genom Res*, **16**(8): 990-994.

Boore JL. 1999. Animal mitochondrial genomes [J]. *Nucleic Acids Res*, **27**: 1767-1780.

Delisle I, Strobeck C. 2005. A phylogeny of the Caniformia (order Carnivora) based on 12 complete protein-coding mitochondrial genes [J]. *Mol Phylogenet Evol*, **37**(1): 192-201.

Douzery E, Randi E. 1997. The mitochondrial control region of Cervidae: evolutionary patterns and phylogenetic content [J]. *Mol Biol Evol*, **14**: 1154-1166.

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

Geffen E, Mercure A, Girman DJ, Macdonald DW, Wayne RK. 1992. Phylogenetic relationships of the fox-like canids: mitochondrial DNA restriction fragment, site and cytochrome β sequence analyses [J]. *J Zool*, **228**: 27-39.

Gissi C, Gullberg A, Arnason U. 1998. The complete mitochondrial DNA sequence of the Rabbit (*Oryctolagus cuniculus*) [J]. *Genomics*, **50**: 161-169.

Gu ZL, Zhao XB, Li N, Wu CX. 2007. Complete sequence of the yak (*Bos grunniens*) mitochondrial genome and its evolutionary relationship with other ruminants[J]. *Mol Phylogenet Evol,* **42**: 248-255.

Hoelzel AR, Lopez JV, Dover GA, O'Brien SJ. 1994. Rapid evolution of heteroplasmic repetitive sequence in the mitochondrial DNA control region of carnivores [J]. *J Mol Evol*, **39**: 191-199.

Jeon KW. (Eds.). Mitochondrial Genomes[M]. New York: Academic Press, 173-216.

Ketmaier V, Bernardini C. 2005. Structure of the mitochondrial control region of the Eurasian Otter (*Lutra lutra*; Carnivora, Mustelidae): Patterns of genetic heterogeneity and implications for conservation of the species in Italy [J]. *J Hered*, **96**(4): 318-328.

Kim KS, Seong EL, Ho WJ, Ji HH. 1998. The complete nucleotide sequence of the domestic dog (*Canis familiaris*) mitochondrial genome [J]. *Mol Phylogenet Evol*, **10**: 210-220.

Lin YH, McLenachan PA, Gore AR, Phillips MJ, Ota R, Hendy MD, Penny D. 2002. Four new mitochondrial genomes and the increased stability of evolutionary trees of mammals from improved taxon sampling [J].*Mol Biol Evol*, **19** (12): 2060–2070.

Moore WS. 1995. Inferring phylogenies from mtDNA variation: Mitochondrial gene trees versus nuclear gene trees [J]. *Evolution*,

49(4): 718-726.

Peng R, Zeng B, Meng XX, Yue BS, Zhang ZH, Zou FD. 2007. The complete mitochondrial genome and phylogenetic analysis of the giant panda (*Ailuropoda melanoleuca*) [J]. *Gene*, **397**: 76-83.

Sambrook J, Russell DW. 2001. Molecular Cloning: A Laboratory Manual[M]. NY: Cold Spring Harbor Press.

Shadel GS, Clayton DA. 1997. Mitochondrial DNA maintenance in vertebrates [J]. *Ann Rev Biochem*, **66**: 409-435.

Swofford DL. 2003. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)[CP/DK]. Version 4.0b10. Sunderland, Massachusetts: Sinauer Associates,

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0 [J]. *Mol Biol Evol*, **24**: 1596-1599.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The Clustal X windows interface: Xexible strategies for multiple sequence alignment aided by quality analysis tools[J]. *Nucleic Acids Res*, **25**(24): 4876-4882.

Ursing BM, Arnason U. 1998. The complete mitochondrial DNA sequence of the pig (*Sus scrofa*) [J]. *Mol Phylogenet Evol*, **47**: 302-306.

Wayne RK, Geffen E, Girman DJ, Koeppfli KP, Lau LM, Marshall CR. 1997. Molecular systematics of the Canidae [J]. *Syst Biol*, **46**: 622-653.

Wei L, Wu XB, Jiang ZG. 2008. The complete mitochondrial genome structure of snow leopard *Panthera uncia* [J]. *Mol Biol Rep*, **36**(5): 871-878.

Wolstenholme DR. 1992. Animal mitochondrial DNA: Structure and evolution[J]. *Int Rev Cytol*, **141**: 173-216.

Xiao B, Ma F, Sun Y, Li QW. 2006. Comparative analysis of complete mitochondrial DNA control region of four species of Strigiformes [J]. *Act Genet Sin*, **33** (11): 965-974.

Xu X, Gullberg A, Arnason U. 1996. The complete mitochondrial DNA (mtDNA) of the donkey and mtDNA comparisons among four closely related mammalian species-pairs [J]. *J Mol Evol*, **43**(5): 438-446.

Yu L, Zhang YP. 2006. Summary of phylogeny in mammalian order Carnivora [J]. *Zool Res*, **27** (6): 657-665.

Zhang DX, Hewitt GM. 1997. Insect mitochondrial control region: A review of its structure, evolution and usefulness in evolutionary studies [J]. *Biochem Syst Ecol*, **25**: 99-120.

Zrzavy J, Ricankova V. 2004. Phylogeny of recent Canidae (Mammalia, Carnivora): Relative reliability and utility of morphological and molecular datasets [J]. *Zool Scr*, **33**: 311-333.