ELSEVIER

Mitochondrion 14 (2001) 279–292

Mitochondrion

www.elsevier.com/locate/mito

Structure and patterns of sequence variation in the mitochondrial DNA control region of the great cats

Kim Jae-Heup, E. Eizirik, S.J. O'Brien, W.E. Johnson*

Laboratory of Genomic Diversity, National Cancer Institute-Frederick, Frederick, MD 21702-1201, USA

Received 20 April 2001; received in revised form 27 July 2001; accepted 30 July 2001

Abstract

Mitochondrial DNA control region structure and variation were determined in the five species of the genus *Panthera*. Comparative analyses revealed two hypervariable segments, a central conserved region, and the occurrence of size and sequence heteroplasmy. As observed in the domestic cat, but not commonly seen in other animals, two repetitive sequence arrays (RS-2 with an 80-bp motif and RS-3 with a 6–10-bp motif) were identified. The 3′ ends of RS-2 and RS-3 were highly conserved among species, suggesting that these motifs have different functional constraints. Control region sequences provided improved phylogenetic resolution grouping the sister taxa lion (*Panthera leo*) and leopard (*Panthera pardus*), with the jaguar (*Panthera onca*). Published by Elsevier Science B.V.

Keywords: Mitochondrial DNA; Control region; Heteroplasmy; *Panthera* genus; Felidae; Phylogenetic analyses

1. Introduction

A variety of molecular genetic markers have been employed in the last three decades to study population structure, intraspecific variation, and phylogenetic relationships among diverse organisms (Avise, 1994). For the last decade, mitochondrial DNA (mtDNA) has been one of the most commonly used molecular markers in vertebrates. Several of its features have led to its widespread use, including a lack of recombination, essentially maternal inheritance, high evolutionary rate, compact size, and conserved gene order (Avise, 1994; Harrison, 1989; Lopez et al., 1994; Moritz et al., 1987; Wilkinson et al., 1997). With a few exceptions, mtDNA is a closed, circular molecule of around 15–20 kb. It is composed of about 37 genes coding for 22 tRNAs, two rRNAs, and 13 mRNAs (Anderson et al., 1981; Wallace, 1986). The control region (CR), the only significant non-coding section, is located in mammals between the tRNA-Pro and tRNA-Phe genes. The CR contains promoters for polycistronic RNA transcription of genes on both the light and heavy strands, as well as the origin of DNA replication for the heavy strand. Using the human reference sequence (or Cambridge Reference Sequence; Anderson et al., 1981; Andrews et al., 1999), the CR stretches from around position 16,000 to 17,000, and then continues from position 1 to around 1000 (depending on its size).

The CR is the most rapidly evolving region of the mtDNA molecule (Aquadro and Greenberg, 1983; Brown et al., 1993; Hoelzel and Dover, 1991; Lopez et al., 1997). It is therefore one of the most commonly used markers for addressing evolutionary relationships among closely related species and/or subspecies. In addition, the most rapidly evolving portions of the

* Corresponding author. Tel.: +1-301-846-1299; fax: +1-301-846-6327.

E-mail address: johnsonw@ncifcrf.gov (W.E. Johnson).

1567-7249/01/$20.00. Published by Elsevier Science B.V.
PII: S1567-7249(01)00027-7

CR have proven to be useful for high-resolution analyses of population structure (Avise, 1994; Stoneking et al., 1991) and for individual identification (Holland and Parsons, 1999).

Functionally, sequences within the CR regulate mtDNA transcription and replication. However, although the CR has been sequenced and described in numerous species, the mechanisms of replication and transcription and their link to CR structural architecture and evolutionary dynamics remain poorly understood. The CR of most species has a central conserved region (CCR) that is surrounded by more-variable A/T rich regions. Within this CCR, several sites have been hypothesized to be important in the initiation of the transcription of mtDNA genes (Chang and Clayton, 1985; Clayton, 1992; King and Low, 1987) and to play roles in mtDNA replication. The diverse functional roles of the CR are likely to have led to contrasting patterns of selective pressures among its different segments.

In this paper, we describe structural features and patterns of molecular genetic variation in complete mtDNA CR sequences in species of the *Panthera* genus. The big cats, or 'roaring cats' of the *Panthera* genus include the lion (*Panthera leo*), jaguar (*Panthera onca*), leopard (*Panthera pardus*), tiger (*Panthera tigris*), and snow leopard (*Panthera uncia*). Evidence has been presented indicating that these species diverged from a common ancestor 3–4 million years ago (Turner, 1996). The relative evolutionary relationship among these species has been difficult to determine from past studies because of their recent and rapid radiation (Johnson et al., 1996; Johnson and O'Brien, 1997; Pecon Slattery and O'Brien, 1998). Our goal in this paper is to provide insights on mtDNA CR evolution and function through comparative analyses of sequence variation among these five *Panthera* genus species. We identify and characterize the different regions of the *Panthera* mtDNA CR and identify conserved motifs which are likely to be of functional importance. In addition, we perform phylogenetic analyses to gain insights on potential evolutionary constraints influencing CR organization and to better resolve the relationships among *Panthera* genus species.

2. Materials and methods

Mitochondrial pellets were purified from 1.0–1.5 g of frozen liver from one individual of each of the five *Panthera* genus species by differential centrifugation in a sucrose gradient (Table 1). Samples were powderized in liquid nitrogen and homogenized with 8–10 strokes at 800 revs./min of a glass–Teflon homogenizer in 5.0–7.5 ml (5 ml/g) of cold homogenization buffer (100 mM Tris–HCl (pH 7.4), 250 mM sucrose, 10 mM EDTA). Nuclei and cellular debris were removed by centrifugation at $1500 \times g$ for 10 min at 4°C. Crude mitochondrial pellets were prepared from the super-

Table 1
Animals used in this study

Sample code	Species name	Common name	Number of clones	Clone code	Clone size (bp)
Ple181	*Panthera leo*	Serengeti lion	2	Ple181-CL1	1611
				Ple181-CL2	1475
Ppa021	*Panthera pardus*	Leopard	3	Ppa021-CL1	1395
				Ppa021-CL2	1466
				Ppa021-CL3	1475
Pon011	*Panthera onca*	Jaguar	5	Pon011-CL1	1585
				Pon011-CL2	1377
				Pon011-CL3	1441
				Pon011-CL4	1574
				Pon011-CL5	1495
Pti065	*Panthera tigris*	Tiger	3	Pti065-CL1	1528
				Pti065-CL2	1309
				Pti065-CL3	1541
Pun086	*Panthera uncia*	Snow leopard	2	Pun086-CL1	1499
				Pun086-CL2	1420

natants by centrifugation at 12,000 × *g* for 15 min at 4°C. MtDNA was purified from the pellet using the Wizard Miniprep kit (Promega; Beckman et al., 1993).

We used the polymerase chain reaction (PCR) with a universal primer set (CR-up-F and R; Kocher et al., 1989) to amplify the complete CR from DNA fractions. Reactions were prepared using 10 mM Tris (pH 8.3), 50 mM KCl, 2.0 mM $Mg(OAc)_2$, 0.001% (w/v), gelatin, each dNTP at 200 μM, 2.5 U of rTh XL polymerase (Applied Biosystems, Inc.), primers at 25 pmol, and 50 ng of extracted DNA. PCR reactions consisted of a hot-start of 95°C for 2 min, followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 56°C for 60 s, and extension at 68°C for 90 s, and ending with an extension step at 72°C for 2 min.

PCR products were separated by electrophoresis on a 0.8% low melting point agarose gel and visualized with ethidium bromide (EtBr) fluorescence under UV light. PCR bands were excised and directly cloned with TOPO TA kits (Invitrogen). Cloned vectors were transformed using OneShot TOP-10 cells, which were plated on LB agar (Sambrook et al., 1989), treated with ampicilin (50 μg/ml) and X-gal for blue/white screening and grown overnight at 37°C. White colonies were picked and cultured overnight in LB broth containing ampicilin (50 μg/ml; Sambrook et al., 1989). Plasmids were isolated from 1.5 ml cultured cells by centrifugation at 10,000 × *g* for 30 s. Supernatants were removed and pellets were resuspended in 50–100 μl distilled water. Plasmids were purified using the Wizard Plus Minipreps DNA Purification System (Promega). Positive clones with inserts were selected by restriction enzyme digestion with *Eco*RI and 2–5 clones of different insert sizes were picked up for each species.

We sequenced miniprepped plasmid DNA on an automated DNA sequencer (ABI 377) using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Inc.), M13 primers and amplification primers (Kocher et al., 1989), including additional species-specific internal primers (RS2-F, TACCCGTACTGTGCTTGCCC; RS2-R, AGCACTTTCGGACAGTTGAG; CCR-F; CTCAACTGTCCGAAAGTGCTT; CCR-R, CCTGTGGAAGCAATAGG AATT; RS3-F, ATTCCTATTGCTCCACAGG; RS3-R, AATCAAAAAGTTCGCA TGTG; Fig. 1). Internal primers for sequencing were designed from highly conserved segments of the clones, using the virtual PCR program Amplify-2.53 (Engels, 1997). Sequences were aligned using Clustal-X (Thompson et al., 1997) and verified by eye. The nucleotide composition, sequence identity and distribution of

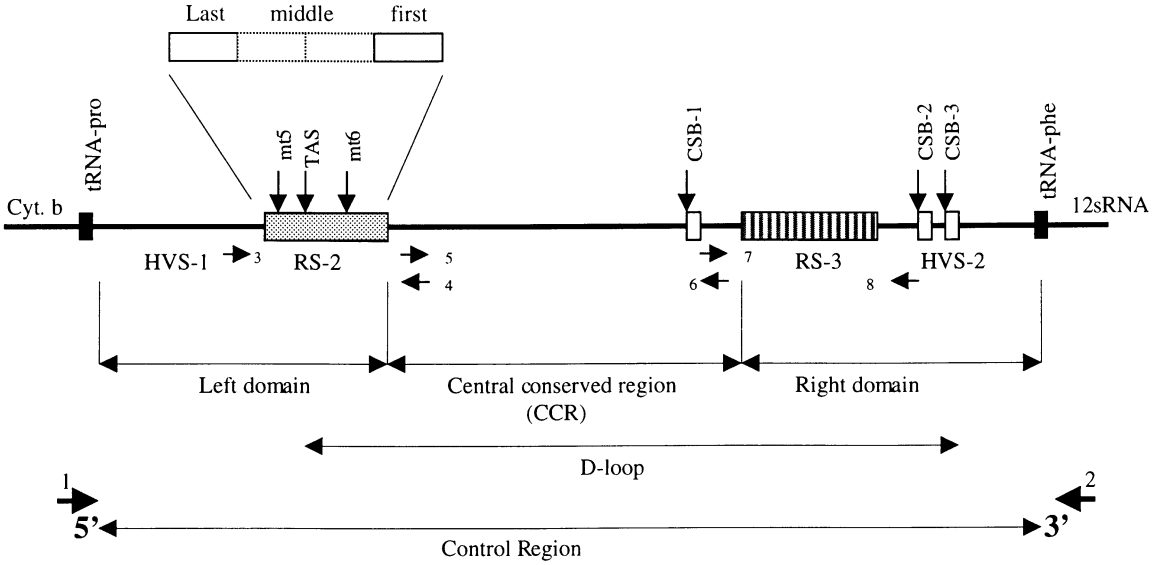


Fig. 1. Schematic of the *Panthera* mitochondrial CR showing conserved blocks, the location of repetitive sequences (RS), and other defined domains. Numbers represent the amplification and sequencing primers (described in text).

variable sites were characterized with MEGA (Kumar et al., 1993). Phylogenetic relationships among *Panthera* species were estimated using PAUP* 4.0b (Swofford, 1998), with maximum parsimony (MP), minimum evolution (ME), and maximum likelihood (ML) approaches using complete mtDNA CR

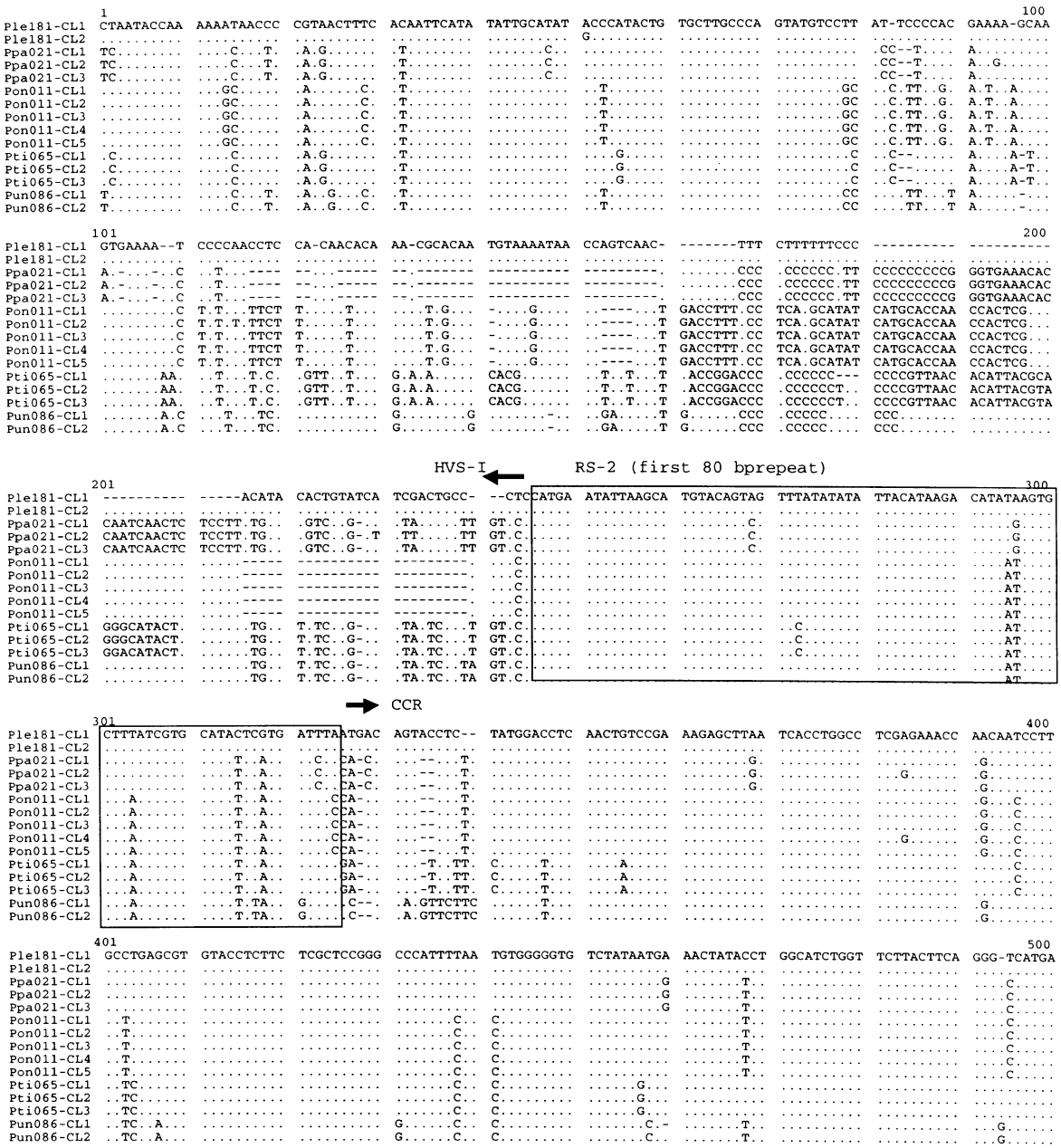


Fig. 2. Aligned sequences of the mitochondrial CR of 15 clones of five *Panthera* genus species (labeled as in Table 1). The alignment contains the sequences of HVS-1, 2, CCR, and first repeat unit of RS-2. RS-3 was not included. Hyphens (-) denote gaps inserted to maximize similarity among the sequences.

sequences (excluding 99 sites at the 5′ flank of the RS-2, which could not be aligned). Separate analyses were also conducted with each of the individual RS-2 repeats. Aligned sequences were deposited in GenBank.

3. Results

3.1. Overall control region patterns

Two to five cloned PCR products of different

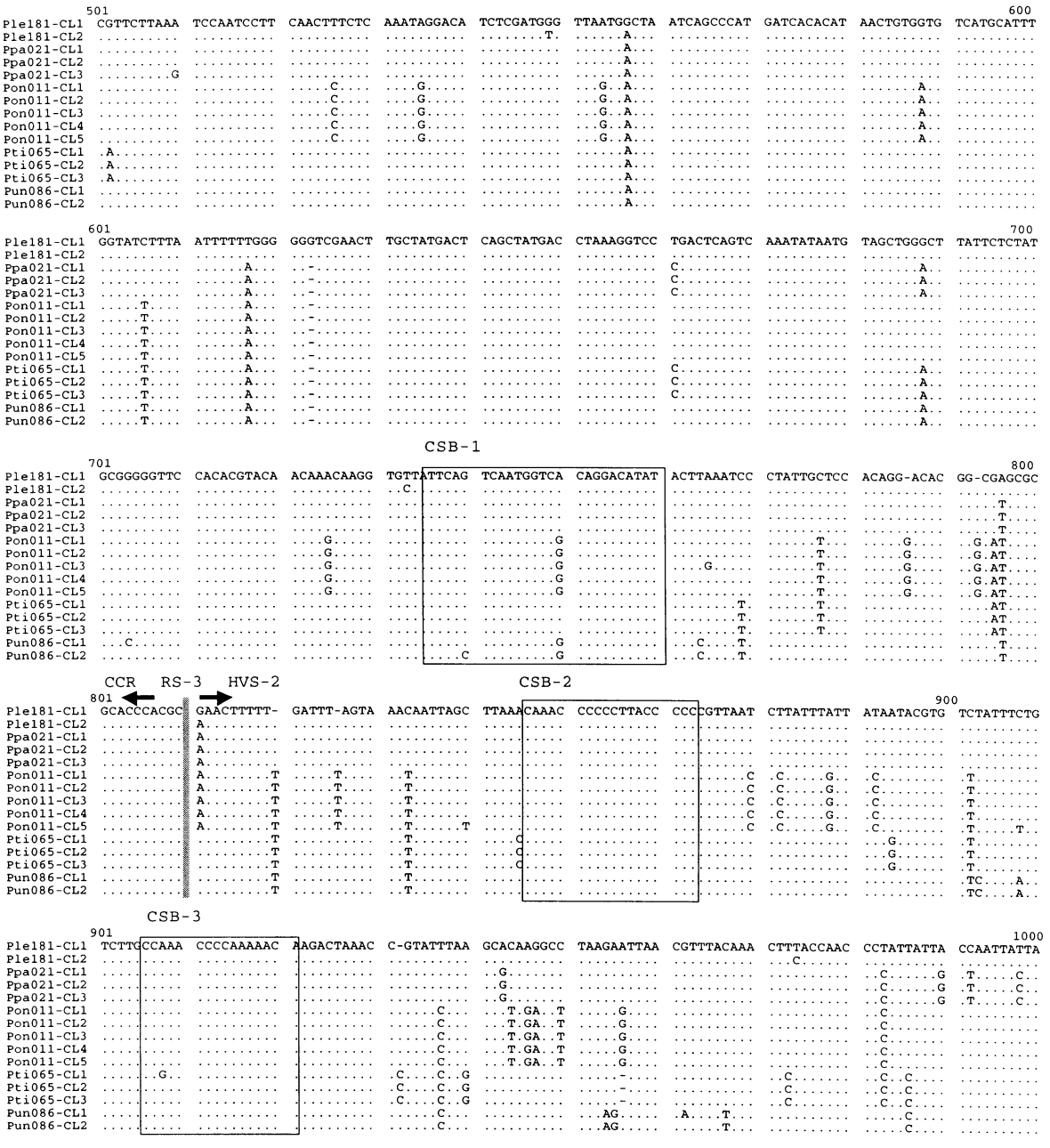


Fig. 2. (*continued*)

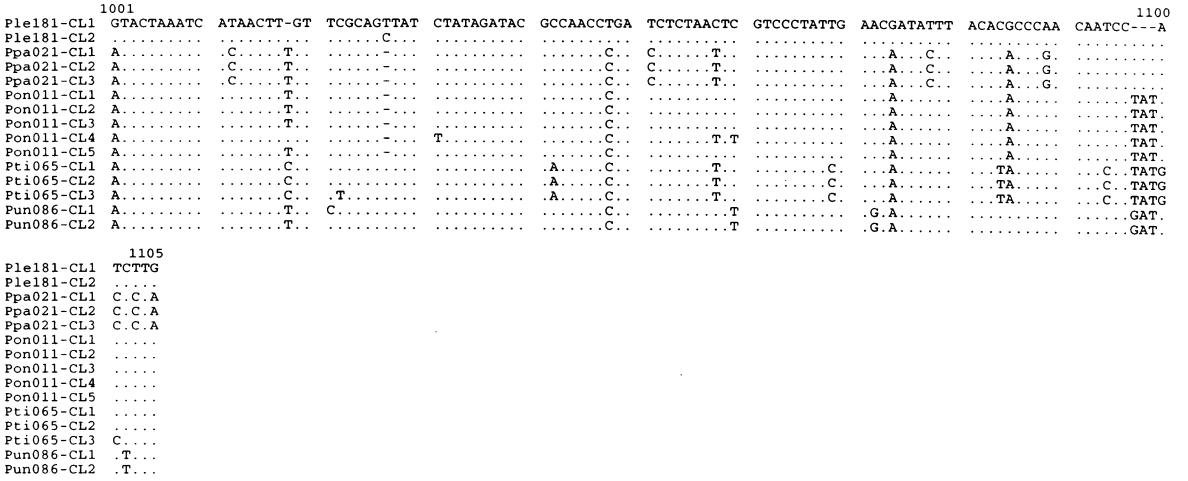


Fig. 2. (*continued*)

lengths of the whole mtDNA CR were sequenced per individual (Fig. 2). The CR between tRNA-pro and tRNA-phe of the 15 clones from the five *Panthera* species ranged from 1309 in tiger (Pti065-CL2) to 1585 bp in jaguar (Pon011-CL1), including repetitive elements (Table 1). These 15 complete sequences were used in all CR comparisons.

We compared the *Panthera* sequences with those of the domestic cat (*Felis catus*; Lopez et al., 1996), other carnivores (Hoelzel et al., 1994) and the pig (*Sus scrofa*; Ghivizzani et al., 1993) to identify the location of previously described sequence blocks. We identified a CCR and two more-variable A/T rich flanking sequences, a left domain (L-domain) at the 5′ end of the CR, and a right domain (R-domain) at the 3′ end of the CR in all five *Panthera* species (Figs. 1 and 2). Hypervariable segments (HVS) and repetitive segments (RS) were characterized in both the L-domain (HVS-1 and RS-2) and in the R-domain (HVS-2 and RS-3; Fig. 1). As in the domestic cat, the two RS observed in the *Panthera* CR were very distinct from each other. Transcriptional promoters of the H-strand and L-strand, termination associated sequence (TAS) elements, and some conserved sequence blocks (CSBs) that have been implicated in H-strand replication (Chang and Clayton, 1986; Clayton, 1982; Gemmell et al., 1996; Sbisa et al., 1997; Wilkinson and Chapman, 1991) were also identified (Fig. 1). These include mt5 (Ohno et al., 1991), mt6 (Kumar et al., 1995), and CSB-1, 2, and 3 (Ghivizzani et al., 1993).

Hot spots, or stretches of sequences with numerous substitutions and insertions/deletions among species, were apparent within the CCR, the L-domain, and the R-domains (Fig. 3). In places, such as the 3′ end of the HVS-1, the amount of variation made it impossible to generate reliable alignments among species.

There were several examples of interspecific and intra-individual size heteroplasmy among the clones (Table 1; Fig. 2). Jaguar clones were the most variable in size, ranging from 1377 to 1585 bp. Although most of this size variation could be attributed to differences in the number of repeats in RS-2 and RS-3, there were other insertions/deletions evident from the alignments.

3.2. Central conserved region patterns

The CCR between the end of RS-2 and the beginning of RS-3 ranged from 476 to 479 bp among species. Most of the sequences were very conserved among the species and within individuals, but there was intra-individual sequence variation (heteroplasmy) at seven positions. Of the numerous sequence blocks that have been identified in the CCR across various species (Hoelzel et al., 1994; Stacy et al., 1997; Wilkinson et al., 1997), we only recognized sequence block D (King and Low, 1987) and conserved sequence block 1 (CSB-1; Ghivizzani et

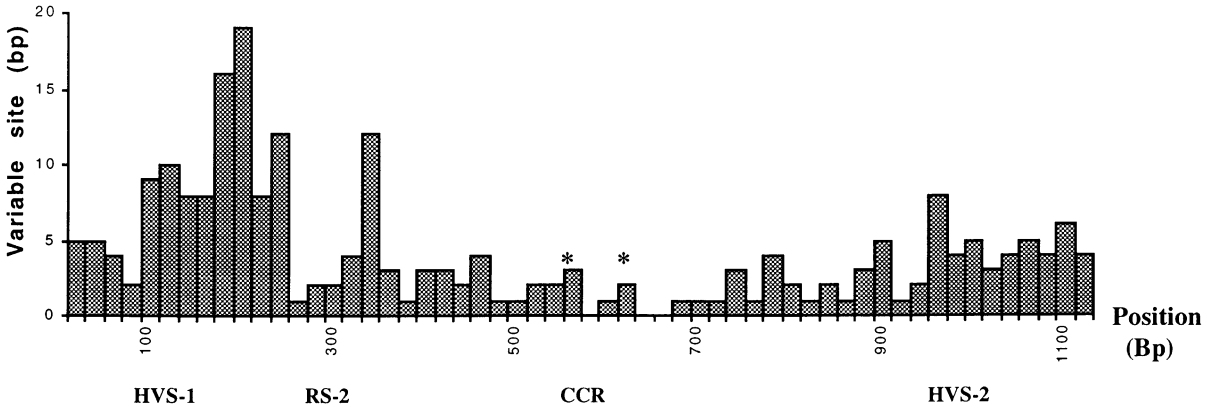


Fig. 3. Distribution pattern of variable sites among CR sequences of five *Panthera* genus species. Each column represents 20 base windows and asterisks (*) denote the most conserved regions. RS-3 was not included in the comparison.

al., 1993) in the *Panthera* sequences. We also identified conserved sequence blocks 2 and 3 (CSB-2 and CSB-3), but these were located in the 3′ end of the RS-3 repetitive array (Fig. 1).

3.3. HVS patterns

We compared sequences of the two HVS segments among the five *Panthera* species and their heteroplasmic clones. The tiger had the longest HVS-1 segment (232–236 bp) compared with the other four species (188–200 bp). Differences in sequence length were due to insertions/deletions starting 80 bp downstream from tRNA-pro. The smaller HVS-1 segments of the lion, leopard, and jaguar may have been generated by separate deletion events. There was both interspecific and intra-individual sequence variation in HVS-1 among the clones. The HVS-2 segment ranged in size from 303–305 in the jaguar to 289–298 bp in the other species. The HVS-1 displayed more interspecific sequence variation than the HVS-2 segment (Figs. 2 and 3), but similar numbers of heteroplasmic sites (seven of 245 sites compared with 13 of 295).

3.4. Repetitive sequence patterns

Panthera mtDNA CR contained two stretches of repetitive sequences, which we labeled as RS-2 and RS-3 based on their locations and sequence characteristics following Hoelzel et al. (1994). RS-2, which flanked the HVS-1, had an 80-bp repeat motif (Fig. 4) and RS-3 had a short 6–10-bp motif (Fig. 5). The number of repeats varied among species and clones from the same individual. There were 1–4 RS-2 repeats/clone, with most of them (80%) having two or three. A lion clone (Ple181-CL1) had four repeats (320 bp), while a leopard (Ppa021-CL1) and tiger clone (Pti065-CL2) had a single repeat. If a clone had more than one copy of the repeat motif, we classified the 5′ repeat as the last (L), the 3′ repeat as the first (F), and intervening repeats as middle (M), based on the assumed direction of H-strand replication. If the clones had a single RS-2 repeat, we called this the first repeat.

The last and middle RS-2 repeats were very similar within the same clone, among clones from the same individual, and among species (97.4% average similarity). The first repeat was less similar (93.3% average similarity). Within this first repeat, the first 48 bp (5′ to 3′) were highly conserved among all clones (98% similarity) relative to the final 32 bp (87% similarity; Fig. 4). The consensus sequence of the presumptive TAS (Doda et al., 1981) and mt5 (Ohno et al., 1991) sequence blocks was recognizable within the first 40 bp (5′ to 3′) of all RS-2 repeats. In contrast, mt6 (Kumar et al. 1995) was recognized only in the last and middle repeats (Fig. 4).

Phylogenetic analyses of the 80-bp motif among the *Panthera* cats and six other felid species clustered the first (3′) repeat from all species into a single group with considerable bootstrap support, 72% for MP and 82% for ME (Fig. 6). The only exception was the cheetah (Aju078), in which the first 80-bp unit was

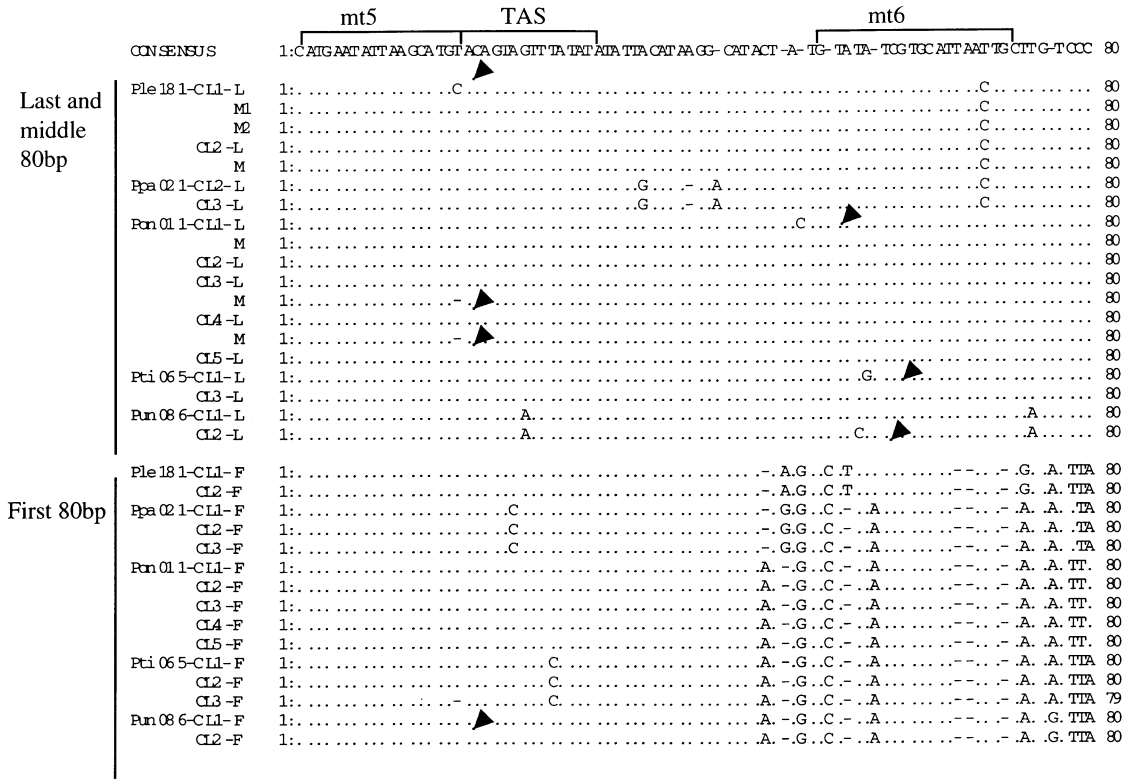


Fig. 4. Alignment of all 80-bp repeat sequences (RS-2) from the HVS-1 of 2–5 clones from a lion (Ple181), jaguar (Pon011), leopard (Ppa021), tiger (Pti065), and snow leopard (Pun086). The top group of sequences consists of the middle (M) and last (5′) (L) repeats, and the bottom group of sequences are the 3′ (F: first) repeats. Arrows represent heteroplasmic sites. Proposed conserved domains (mt5, mt6, and a TAS element) are labeled.

identical to the second. Within this group of 3′ repeats, there was also consistent support for a monophyletic lineage composed of all five *Panthera* species (85% MP and 80% ME bootstrap support) and for the additional clustering of the lion and leopard (81% MP and 61% ME). In contrast, there were no consistent patterns among the last or middle repeats, although repeats from clones of the same individual sometimes clustered together.

The phylogenetic patterns observed among the 80-bp repeats were compared with those estimated using the sequence variation in HVS-1, the first (3′) repeat of RS-2, CCR, and HVS-2. Nucleotides from positions 117–215 were excluded from this analysis because the large amount of variation in this area made alignment difficult. Clones from each individual clustered together with 100% bootstrap support. As observed with the 80-bp repeats, among the five *Panthera* genus species, the lion and leopard were sister taxa (97% MP and 96% ME bootstrap support) that shared a common ancestor with the jaguar (74% MP and 67% ME bootstrap support). Together, these three species shared ancestry with the tiger and snow leopard.

RS-3 differed greatly from RS-2, consisting of an array of short tandem, 6–8-bp repeats that varied considerably among clones and species. Among all the clones, there were from 28 to 49 repeats, and a total of 238–376 bp in length (Fig. 5). The repeats of the RS-3 arrays had similar nucleotide patterns that were interrelated to each other in a hierarchical manner based on alternating purines and pyrimidines. The fundamental repeat core of the repeats was TACACG, from which the 12 other repeat motifs could be derived by a process of duplication, deletion and occasional substitutions. The 5′ RS-3 repeat motif

		Number of repeat (bp)
Lion (*Panthera leo*)		
Ple181-CL1	fjbba babab ababa babab ababa caaaa aaaaa aaaaa aaaaa aa cg	49 (330)
Ple181-CL2	fjbba babab ababa babab aba caaaa aaaaa aaaaa cg	40 (274)
Leopard (*P. pardus*)		
Ppa021-CL1	hfabi bbbck abbck abbck abbbb bbbbb bbbbb b c ccccc cc cg	46 (362)
Ppa021-CL2	hfabb bbbck abbck abbck abbbb bbbbb bbbbb b ccccc cc cg	45 (354)
Ppa021-CL3	hfabb bbbck abbck abbck abbbb bbbbb bbbbb b c ccccc cc cg	46 (362)
Jaguar (*P. onca*)		
Pon011-CL1	fabbb bbbbb bcccc bbccf abccc fabbb bbbbb bbb ccccc cc cg	47 (372)
Pon011-CL2	fabbb bbbb ccc fabbb bbbbb ccccc cc cg	32 (254)
Pon011-CL3	fabbb b cc fabbb bbbbb bbb ccccc cc cg	30 (238)
Pon011-CL4	fabbb bbbbb bcccc bcccf abccc fabbb bbbbb bbb ccccc cc cg	47 (372)
Pon011-CL5	fabbb bbbbb bcccc bcccf abccc labbb bbbbb bbb ccccc cc cg	47 (372)
Tiger (*P. tigris*)		
Pti065-CL1	facbb bbbcc bbbcc fcccc bbccf abbbb bbbbb bbbbb bb cc cg	46 (366)
Pti065-CL2	facbb bbbcc bbbcc cc bb bbb bb cc cg	28 (224)
Pti065-CL3	facbb bbbcc ebbcc ccccc bbccf bbbbb bbbbb bbbbb bbb cc cg	47 (376)
Snow-leopard (*P. uncia*)		
Pun086-CL1	hdabb cbbbc bccbb bbfab bbbfa abbbb bbbbc caaaa aaaaa cc cg	49 (368)
Pun086-CL2	hdabb cbbbc bccbb bbfab bbbfa a aaaa aaaaa cc cg	39 (288)

Fig. 5. Encoded repetitive array sequences (RS-3) of 2–5 clones from five *Panthera* cat species. Sequences are labeled as in Table 1. (a), TACACG; (b), TACACACG; (c), TATACACG; (d), TACACACG; (e), TGCACACG; (f), TATACGCG; (g), TATACACATG; (h), TATACGTG; (i), CACACACG; (j), CACATGTG; (k), TACGCACG; (l), TATGCGCG.

was either TATACGCG or TATACGTG among species and was fixed in all clones of a single individual (Fig. 5). The 3′ motif of RS-3 was invariable in all species and individuals (TATACACGTATACACATG). There was no pattern (or phylogenetic signal) apparent in the internal RS-3 repeats, which varied in both number and motif among species and clones.

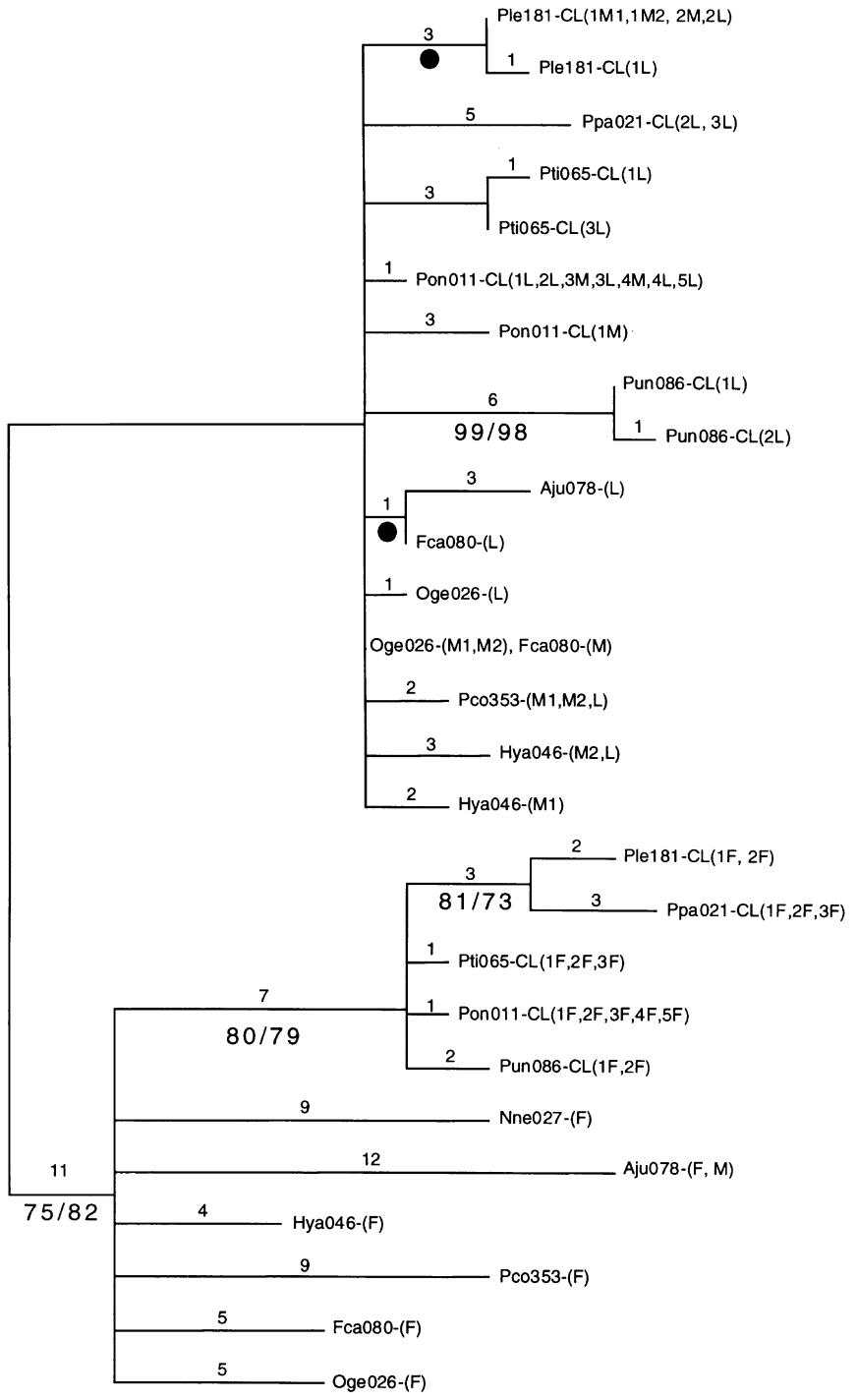
4. Discussion

4.1. Overall control region patterns

The arrangement of the mtDNA CR were similar in the lion, leopard, jaguar, tiger, and snow leopard, and recapitulated the patterns described in the domestic cat (Lopez et al., 1996). We recognized several previously described segments, including two hypervariable segments (HVS-1 and HVS-2), a CCR, and two repetitive sequence arrays (RS-2 and RS-3). Compared with other mammalian species, there were several notable differences in this arrangement, however. For example, although 40–160-bp RS-2 motifs have been described in numerous species (e.g. Buroker et al., 1990; Hayasaka et al., 1991; Yang et al., 1994; Wilkinson et al., 1997), while comparatively short <20-bp RS-3 motifs have been described in other species (e.g. Desjardins and Morais, 1990; Mignotte et al., 1990; Ghivizzani et al., 1993; Hoelzel et al., 1994; Savolainen et al., 2000; Schmitz et al., 2000), the presence of both sets of repeats in a single species is rare. Our results suggest that these two sets of repeats are present in all major cat lineages, making the Felidae family an interesting group in which to investigate the origin and evolution of these sequences.

Sequence variation also followed patterns previously described in other species, varying greatly among the different CR sections (Fig. 3). The variation was highest in HVS-1 (71.3%), followed by HVS-2 (89.4%). Among the clones of *Panthera* individuals, heteroplasmic sites were distributed randomly in the HVS-1, CCR, and HVS-2, but were not observed in the TAS or CSBs. Of 27 heteroplasmic sites, only one (at position 384) varied among clones of more than one species. In each species, there was one clone that differed from the consensus sequence at heteroplasmic sites in at least two and as many as five sites.

Improvements in sequence technology have facilitated the recognition of heteroplasmy (Steighner and Holland, 1998; Underhill et al., 1997). The proportion of heteroplasmic point mutations has been shown to vary among individuals (Byrne et al., 1991; Hamazaki et al., 1993; Howell et al., 1994; Hurko et al., 1990), among tissues (Matthews et al., 1994), and to increase


Ple181-CL(1M1,1M2, 2M,2L)
3
1
Ple181-CL(1L)
5
Ppa021-CL(2L, 3L)
1
Pti065-CL(1L)
3
Pti065-CL(3L)
1
Pon011-CL(1L,2L,3M,3L,4M,4L,5L)
3
Pon011-CL(1M)
Pun086-CL(1L)
6
99/98
1
Pun086-CL(2L)
3
Aju078-(L)
1
Fca080-(L)
1
Oge026-(L)
Oge026-(M1,M2), Fca080-(M)
2
Pco353-(M1,M2,L)
3
Hya046-(M2,L)
2
Hya046-(M1)
2
Ple181-CL(1F, 2F)
3
81/73
3
Ppa021-CL(1F,2F,3F)
1
Pti065-CL(1F,2F,3F)
7
80/79
1
Pon011-CL(1F,2F,3F,4F,5F)
2
Pun086-CL(1F,2F)
9
Nne027-(F)
11
12
Aju078-(F, M)
75/82
4
Hya046-(F)
9
Pco353-(F)
5
Fca080-(F)
5
Oge026-(F)

with age (Macmillan et al., 1993; Matthews et al., 1994; Michikawa et al., 1999). Liver, from which the mtDNA in this study was extracted, has been documented to have higher rates of heteroplasmy compared with other tissues (Hamazaki et al., 1993).

4.2. Repetitive sequence patterns

The most striking sources of felid CR size variation were the two contrasting repeat arrays (RS-2 and RS-3). Among clones, RS-2 ranged in size from 80 to 320 bp (1–4 repeats) and RS-3 from 224 to 376 bp (28–49 repeats). The structure of felid RS-2 is most similar to that of the bat (Wilkinson et al., 1997). The bat repeat is composed of a 78–85-bp sequence located between the 5′ end of the CR and the CCR. As in the bat, each repeat of RS-2 contains a TAS.

MtDNA CR size variation and repeat turnover may be the result of the replication process. In vertebrate mtDNA, a third strand of DNA, the D-loop, is synthesized off the light strand (L-strand) and displaces the heavy strand (H-strand) of the molecule. The D-loop, which is initiated near the CSBs (1, 2, 3) and ends near the terminal associated sequence (TAS; Clayton, 1982), may result from an aborted attempt at the replication of the H-strand (Chang and Clayton, 1985). TAS at different locations could lead to D-loop strands of different lengths in different molecules (Doda et al., 1981). However, the existence of multiple TASs alone is not enough to produce length variations as no DNA is lost or gained in the nascent H-strand (Arnason and Rand, 1992).

The phylogenetic tree with the separated RS-2 repeat element of each clone and other related species in the Felidae (clouded leopard, cheetah, Geoffroy's cat, domestic cat, and jaguarundi; Fig. 6) suggests that the first (3′) repeat in the tandem array has been subjected to a different evolutionary process compared with the other units in the series, as has been seen in other species (e.g. Wilkinson and Chapman, 1991). They form a clear monophyletic cluster in the tree, indicating either a common origin not shared by the other repeats, which is unlikely, or sequence constraints leading to the maintenance of different motifs. The topological hierarchy among these repeats is consistent with current hypotheses of the evolutionary relationships among the sampled species, in contrast with the other repeats, which show no similar pattern. On the contrary, there is clear indication of close similarity (or identity) among repeats of the same individual, suggesting that recent episodes of duplication involving all repeats but the first one is quite common. The only indication of a recent episode of duplication involving the first repeat is observed in the cheetah (*Acinonyx jubatus*), where the two repeats on the 3′ side of the array were identical (Aju078-M and Aju078-F).

The other repetitive sequence in the CR of big cats is RS-3, which has relatively short tandem repeat motifs. These repeat units have common structural motifs that are related to each other. They appear to be based on a 'TACACG' motif, similar to the 'ACGT' model proposed by Hoelzel et al. (1994). In our data, the CA, CG, and TA are the most common pairs of nucleotides. Combinations of these make up most of the RS-3 motifs. A process of duplication, deletion, and substitution of the CA motif can explain the generation of the basic repeat elements seen in this study (Fig. 5). One or two stepwise changes account for most of the elements. For example, if we presume that replication is occurring only from the 3′ end, the first core motif might be TATACACATG (motif G). All of the additional motif elements can be derived from a combination of deletions, duplications of CA or TA, or conversions from CA to TA.

Fig. 6. Phylogenetic relationships among RS-2 repeats from various felid species, estimated using MP with PAUP*. Each branch is labeled as in Table 1 and Fig. 4 by individual, clone, and repeat (F, first 3′ repeat; M, middle repeat; L, last 5′ repeat). Identification labels for other felid species are as follows: Nne, *Neofelis nebulosa*; Aju, *Acinonyx jubatus*; Hya, *Herpailurus yagouaroundi*; Pco, *Puma concolor*; Fca, *Felis catus* (Lopez et al., 1996); Oge, *Oncifelis geoffroyi*. The phylogenetic tree represents a strict consensus from a heuristic search with simple taxon addition, constrained to generate a maximum of 2000 trees (length, 89; CI, 0.629). The ME analysis, which produced a concordant topology, was performed with PAUP* using Kimura two-parameter distances and a neighbor-joining algorithm followed by a branch-swapping procedure. Values above branches are the number of steps in the consensus MP tree. The numbers below branches are the bootstrap support (from 100 replications) values from MP/ME analyses (MP bootstrap: 100 replications with heuristic search, simple taxon addition, constrained to a maximum of 500 trees; ME bootstrap: using neighbor-joining). The black circles indicate branches with bootstrap support of >50% using at least one of the methods.

At a higher level of organization, repeat elements combine to form 'compound units' that are reiterated (e.g. 'ab' in the lion and 'bb' or 'cc' in the tiger). Most of the variation within arrays can be characterized as repetition at one of these three levels: sub-element (repetition of CA or CG), element, or compound unit, as seen in other taxa (Hoelzel et al., 1994; Schmitz et al., 2000). Additional variation appears to be generated by point mutations and the spreading of these variant elements through the array. The cat RS-3 repeat is upstream of the presumed start of the D-loop so that slippage at the beginning of its synthesis may generate this RS-3 array. Local expansion by slippage is obvious in several arrays, such as the 'aaaa' pattern in the lion array.

Since RS-2 and RS-3 may have important functional roles, they are possibly under selective pressure, both positive and negative. For example, multiple repeats may aid in binding proteins involved in mtDNA replication. The number of repeats is probably also under selective pressure. For example, organelles that are larger because they have more RS-2 and RS-3 repeats might replicate more slowly, and thus be at a selective disadvantage compared with smaller genomes within the same individual. Selection pressures are probably complex, however, since the sequences may undergo selection at multiple levels, including competition among molecules, cells, and individuals (Wilkinson et al., 1997).

Selection among individuals may favor an increase in repeat numbers of RS elements for at least two of the reasons proposed by Wilkinson et al. (1997). One is that multiple repeats could compensate for deleterious mutations during the life of an individual. MtDNA is well known for its high mutation rate (Brown, 1985; Li, 1991), although potential DNA repair mechanisms have been reported (Jacobs and Holt, 1998; Nilsen et al., 1997). Alternatively, concerted evolution caused by repeat duplication and deletion could eliminate damaged repeat sequences.

The number of tandem repeats may also vary with age and with generation time. For example, Solignac et al. (1987) found an increase in the frequency of longer repeats in a *Drosophila* germ line during the adult life of the fly. It has also been proposed that the mean percentage of longer mtDNA molecules will increase in a long generation system compared with a short generation system (Casane et al., 1997).

RS-2 and RS-3 are similar in one significant way. The $3'$ end of both are highly conserved among species, suggesting that the first RS-2 motif and the c (TATACACG) and g (TATACACATG) motifs of RS-3 have conserved functional roles. The other repetitive elements seem to be under different selective pressures and may be more influenced by processes such as concerted evolution.

The results of this study demonstrate that the mtDNA CR patterns of variation among *Panthera* genus species are complex. Portions of the CR are under strong selective constraint, showing no change across several million years of evolutionary time. Other portions are evolving so rapidly that they can be impossible to align among individuals of the same species (as in Eizirik et al., 1998) and are variable among clones from the same individual.

Acknowledgements

The authors would like to thank G.S. Wilkinson and N. Yuhki for helpful comments on the project and the manuscript. The authors are also grateful to J. Martenson, S. Cevario, G. Pei, and A. Snyder for laboratory assistance. The content of this publication does not necessarily reflect the views or policies of the DHHS, nor does the mention of trade name, commercial products, or organization imply endorsement by the U.S. Government.

References

Anderson, S., Bankier, A.T., Barrell, B.G., et al., 1981. Sequence and organization of the human mitochondrial genome. Nature 290, 457–465.

Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M., Howell, N., 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat. Genet. 23, 147.

Aquadro, C.F., Greenberg, B.D., 1983. Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. Genetics 103, 287–312.

Arnason, E., Rand, D.M., 1992. Heteroplasmy of short tandem repeats in mitochondrial DNA of Atlantic cod, *Gadus morhua*. Genetics 132, 211–220.

Avise, J.C., 1994. Molecular Markers, Natural History and Evolution, Chapman and Hall, New York.

Beckman, K.B., Smith, M.F., Orrego, C., 1993. Purification of Mitochondrial DNA with Wizard Minipreps DNA Purification System, Promega Note, Vol. 43. pp. 10–13.

Brown, W.M., 1985. The mitochondrial genome of animals. In: Macintyre, R.J. (Ed.). Molecular Evolutionary Genetics, Plenum Press, New York, pp. 95–130.

Brown, J.R., Beckenbach, A.T., Smith, M.J., 1993. Intraspecific DNA sequence variation of the mitochondrial control region of white sturgeon (*Acipenser transmontanus*). Mol. Biol. Evol. 10, 326–341.

Buroker, N.E., Brown, J.R., Gilbert, T.A., Ohara, P.J., Beckenbach, A.T., Thomas, W.K., Smith, M.J., 1990. Length heteroplasmy of sturgeon mitochrondrial-DNA, An illegitimate elongation model. Genetics 124, 157–163.

Byrne, E., Trounce, I., Marzuki, S., et al., 1991. Functional respiratory chain studies in mitochondrial cytopathies: support for mitochondrial DNA heteroplasmy in myoclonus epilepsy and ragged red fibers syndrome. Acta Neuropathol. 81, 318–323.

Casane, D., Dennebouy, N., de Rochambeau, H., Mounolou, J.C., Monnerot, M., 1997. Nonneutral evolution of tandem repeats in the mitochondrial DNA control region of Lagomorphs. Mol. Biol. Evol. 14, 779–789.

Chang, D.D., Clayton, D.A., 1985. Priming of mammalian mitochondrial DNAs begins at the light strand promoter. Proc. Natl. Acad. Sci. USA 82, 351–355.

Chang, D.D., Clayton, D.A., 1986. Identification of primary transcription start sites of mouse mitochondrial DNA: accurate in vitro initiation of both heavy- and light-strand transcriptions. Mol. Cell. Biol. 6, 1446–1453.

Clayton, D.A., 1982. Replication of animal mitochondrial DNA. Cell 28, 693–705.

Clayton, D.A., 1992. Transcription and replication of animal mitochondrial DNAs. Int. Rev. Cytol. 141, 217–232.

Desjardins, P., Morais, R., 1990. Sequence and gene organization of the chicken mitochondrial genome, a novel gene order in higher vertebrates. J. Mol. Biol. 212, 599–634.

Doda, J.N., Wright, C.T., Clayton, D.A., 1981. Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. Proc. Natl. Acad. Sci. USA 78, 6116–6170.

Eizirik, E., Bonatto, S.L., Johnson, W.E., et al., 1998. Phylogeographic patterns and evolution of the mitochondrial DNA control region in two neotropical cats (Mammalia, Felidae). J. Mol. Evol. 47, 613–624.

Engels, B., 1997. Amplify: Software for PCR, version 2.53B. University of Wisconsin, WI.

Gemmell, N.J., Western, P.S., Watson, J.M., Graves, J.A., 1996. Evolution of the mammalian mitochondrial control region – comparisons of control region sequences between monotreme and therian mammals. Mol. Biol. Evol. 13, 798–808.

Ghivizzani, S.C., Mackay, S.L.D., Madsen, C.S., Laipis, P.J., Hauswirth, W.W., 1993. Transcribed heteroplasmic repeated sequences in the porcine mitochondrial DNA D-loop region. J. Mol. Evol. 37, 36–47.

Hamazaki, S., Koshiba, M., Sugiyama, T., 1993. Organ distribution of mutant mitochondrial tRNA Aleu (UUR) gene in a MELAS patient. Acta Pathol. Jpn. 43, 187–191.

Harrison, R.G., 1989. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends Ecol. Evol. X, 6–11.

Hayasaka, K., Ishida, T., Horai, S., 1991. Heteroplasmy and polymorphism in the major noncoding region of mitochondrial-DNA in Japanese monkeys, association with tandemly repeated sequences. Mol. Biol. Evol. 8, 399–415.

Hoelzel, A.R., Dover, G.A., 1991. Evolution of the cetacean mitochondrial D-loop region. Mol. Biol. Evol. 8, 475–493.

Hoelzel, A.R., Lopez, J.V., Dover, G.A., O'Brien, S.J., 1994. Rapid evolution of a heteroplasmic repetitive sequence in the mitochondrial DNA control region of carnivores. J. Mol. Evol. 39, 191–199.

Holland, M.M., Parsons, T.J., 1999. Mitochondrial DNA sequence analysis — validation and use for forensic casework. Forensic Sci. Rev. 11, 22–50.

Howell, N., Xu, M., Halvorson, S., Bodis-Wollner, I., Sherman, J., 1994. A heteroplasmic LHON family: tissue distribution and transmission of the 11778 mutation. Am. J. Hum. Genet. 55, 203.

Hurko, O., Johns, D.R., Rutledge, S.L., et al., 1990. Heteroplasmy in chronic external ophthalmoplegia: clinical and molecular observations. Pediatr. Res. 28, 542.

Jacobs, H.T., Holt, I.J., 1998. The mitochondrial genetic system. In: Singh, K.K. (Ed.). Mitochondrial DNA Mutations in Aging, Disease and Cancer, Springer, New York, pp. 43–84.

Johnson, W.E., O'Brien, S.J., 1997. Phylogenetic reconstruction of the Felidae using 16S rRNA and NADH-5 mitochondrial genes. J. Mol. Evol. 44, S98–S116.

Johnson, W.E., Dratch, P.A., Martenson, J.S., O'Brien, S.J., 1996. Resolution of recent radiations within three evolutionary lineages of Felidae using mitochondrial restriction fragment length polymorphism variation. J. Mamm. Evol. 3, 97–120.

King, T.C., Low, R.L., 1987. Mitochondrial DNA displacement loop structure depends on growth state in bovine cells. J. Biol. Chem. 262, 6214–6220.

Kocher, T.D., Thomas, W.K., Meyer, A., et al., 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc. Natl. Acad. Sci. USA 86, 6196–6200.

Kumar, S., Tamura, K., Nei, M., 1993. MEGA: Molecular Evolutionary Genetic Analysis, The Pennsylvania State University, University Park, PA.

Kumar, S., Suzuki, H., Onoue, S., Suzuki, S., Hattori, N., 1995. Rat mitochondrial mtDNA-binding proteins to inter-specifically conserved sequences in the displacement loop region of vertebrate mtDNAs. Biochem. Mol. Biol. Int. 36, 973–981.

Li, W.H., 1991. Molecular Evolution, Sinauer Associates, Sunderland, MA.

Lopez, J.V., Yuhki, N., Masuda, R., Modi, W., O'Brien, S.J., 1994. *Numt*, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. J. Mol. Evol. 39, 174–190.

Lopez, J.V., Cevario, S., O'Brien, S.J., 1996. Complete nucleotide sequences of the domestic cat (*Felis catus*) mitochondrial genome and a transposed mtDNA tandem repeat (Numt) in the nuclear genome. Genomics 33, 229–246.

Lopez, J.V., Culver, M., Stephens, J.C., Johnson, W.E., O'Brien, S.J., 1997. Rates of nuclear and cytoplasmic mitochondrial

DNA sequence divergence in mammals. Mol. Biol. Evol. 14, 277–286.

Macmillan, C., Lach, B., Shoubridge, E.A., 1993. Variable distribution of mutant mitochondrial DNAs (tRNA Leu[3243]) in tissues of symptomatic relative with MELAS: the role of mitotic segregation. Neurology 43, 1586–1590.

Matthews, P.M., Hopkins, J., Brown, R.M., Stephenson, J.B.P., Hilton-Jones, D., Brown, G.K., 1994. Comparison of the relative levels of the 3243 (A → G) mtDNA mutation in heteroplasmic adult and fetal tissues. J. Med. Genet. 31, 41–44.

Michikawa, Y., Mazzucchelli, F., Bresolin, N., Scrlato, G., Attardi, G., 1999. Aging-dependent large accumulation of point mutations in the human mtDNA control region for replication. Science 286, 774–779.

Mignotte, F., Gueride, M., Champagne, A.-M., Mounolou, J.-C., 1990. Direct repeats in the non-coding region of rabbit mitochondrial DNA. Eur. J. Biochem. 194, 561–571.

Moritz, C., Dowling, T.E., Brown, W.M., 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annu. Rev. Ecol. Syst. 18, 269–292.

Nilsen, H., Otterlei, M., Haug, T., et al., 1997. Nuclear and mitochondrial uracil-DNA glycocylases are generated by alternative splicing and transcription from different positions in the UNG gene. Nucleic Acids Res. 25, 750–755.

Ohno, K., Tanaka, M., Suzuki, H., Ohbayashi, T., Ikebe, S., 1991. Identification of a possible control element, Mt5, in the major noncoding region of mitochondrial DNA by intraspecific nucleotide conservation. Biochem. Int. 24, 263–271.

Pecon Slattery, J., O'Brien, S.J., 1998. Patterns of Y and X chromosome DNA sequence divergence during the Felidae radiation. Genetics 148, 1245–1255.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Savolainen, P., Arvestad, L., Lundeberg, J., 2000. mtDNA tandem repeats in domestic dogs and wolves: mutation mechanism studied by analysis of the sequence of imperfect repeats. Mol. Biol. Evol. 17, 474–488.

Sbisa, E., Tanzariello, F., Reyes, A., Pesole, G., Saccone, C., 1997. Mammalian mitochondrial D-loop structural analysis: identification of new conserved sequences and their functional and evolutionary implications. Gene 205, 125–140.

Schmitz, J., Ohme, M., Zischler, H., 2000. The complete mitochondrial genome of *Tupaia belangeri* and the phylogenetic affiliation of Scandentia to other Eutherian orders. Mol. Biol. Evol. 17, 1334–1343.

Solignac, M., Gebnermont, J., Monnerot, M., Mounolou, J.C., 1987. *Drosophila* mitochondrial genetics: evolution of heteroplasmy through germ line cell divisions. Genetics 117, 687–696.

Stacy, J.E., Jarde, P.E., Steen, H., Ims, R.A., Puvis, A., Jakobsen, K.S., 1997. Lack of concordance between mtDNA gene flow and population density fluctuations in the bank vole. Mol. Ecol. 6, 751–759.

Steighner, R.J., Holland, M., 1998. Amplification and sequencing of mitochondrial DNA in forensic casework. Methods Mol. Biol. 98, 213–223.

Stoneking, M., Hedgecock, D., Higuchi, R.G., Vigilant, L., Erlich, H.A., 1991. Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. Am. J. Hum. Genet. 48, 370–382.

Swofford, D.L., 1998. PAUP*, Version 4.0, Sinauer, Sunderland, MA.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The Clustal-X windows interface: flexible strategies for multiple alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Turner, A., 1996. The big cats and their fossil relatives, Columbia University Press, New York, NY.

Underhill, P.A., Jin, L., Lin, A.A., et al., 1997. Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. Genome Res. 10, 996–1005.

Wallace, D.C., 1986. Mitochondrial genes and disease. Hosp. Pract. 21, 77–92.

Wilkinson, G.S., Chapman, A.M., 1991. Length and sequence variation in Evening Bat D-loop mtDNA. Genetics 128, 607–617.

Wilkinson, G.S., Mayer, F., Kerth, G., Petri, B., 1997. Evolution of repeated sequence arrays in the d-loop region of bat mitochondrial DNA. Genetics 147, 1035–1048.

Yang, Y.G., Lin, Y.S., Wu, J.L., Hui, C.F., 1994. Variation in mitochondrial DNA and population structure of the Taipei treefrog *Rhacophorus taipeianus* in Taiwan. Mol. Ecol. 3, 219–228.