Genetic diversity in the lion (panthera leo (Linnaeus 1758)) : unravelling the past and prospects for the future

Bertola, L.D.

Chapter 2

Genetic diversity, evolutionary history and implications for conservation of the lion (*Panthera leo*) in West and Central Africa

Journal of Biogeography (2011), 38, 1356-1367

L. D. Bertola, W. F. van Hooft, K. Vrieling, D. R. Uit de Weerd, D. S. York, H. Bauer, H. H. T. Prins, P. J. Funston, H. A. Udo de Haes, H. Leirs, W. A. van Haeringen, E. Sogbohossou, P. N. Tumenta and H. H. de Iongh

Abstract

In recent decades there has been a marked decline in the numbers of African lions (*Panthera leo*), especially in West Africa, where the species is regionally endangered. Based on the climatological history of western Africa, we hypothesize that West and Central African lions have a unique evolutionary history, which is reflected by their genetic makeup.

In this study 126 samples, throughout the lion's complete geographic range, were subjected to phylogenetic analyses. DNA sequences of a mitochondrial region, containing cytochrome *b*, tRNAPro, tRNAThr and the left part of the control region were analysed.

Bayesian, maximum likelihood and maximum parsimony analyses consistently showed a distinction between lions from West and Central Africa and lions from Southern and East Africa. West and Central African lions are more closely related to Asiatic lions than to the Southern and East African lions. This can be explained by a Pleistocene extinction and subsequent recolonization of West Africa from refugia in the Middle East. This is further supported by the fact that the West and Central African clade shows relatively little genetic diversity and is therefore thought to be an evolutionarily young clade.

The taxonomic division between an African and an Asiatic subspecies does not fully reflect the overall genetic diversity within lions. In order to conserve genetic diversity within the species, genetically distinct lineages should be prioritized. Understanding the geographic pattern of genetic diversity is key to developing conservation strategies, both for *in situ* management and for breeding of captive stocks.

Keywords: Central Africa, evolutionary history, genetic diversity, lion, *Panthera leo*, phylogenetics, phylogeography, West Africa.

Introduction

Presently, two subspecies of the lion are recognized by the International Union for the Conservation of Nature (IUCN): the African lion, *Panthera leo leo* (Linnaeus, 1758), and the Asiatic lion, *Panthera leo persica* (Meyer, 1826) (IUCN, 2008). This distinction has been confirmed in recent genetic studies (Driscoll *et al.*, 2002; Burger *et al.*, 2004; Dubach *et al.*, 2005; Barnett *et al.*, 2006a; Antunes *et al.*, 2008). However, the genetic diversity within the species is greater than this taxonomic classification implies; several studies based on genetic data have reported distinct phylogenetic groups within the African subspecies (Table 1), partially overturning earlier categorizations based on morphological traits and geographic distribution. A recent publication on lion phylogeny derived from craniometric data concluded that two major evolutionary clusters can be distinguished: sub-Saharan Africa and North Africa/Asia (Mazák, 2010), also deviating from the former Africa–Asia separation.

Table 1. Overview of the genetic studies reporting phylogenetic units within *Panthera leo*.

Authors (year)	Findings		Units distinguished	Method
Dubach *et al.* (2005)	6 maternal haplotypes		South-west Africa East of the Rift valley West of the Rift Valley Sabi Sands (RSA) (Asiatic lions not included)	Cytochrome *b* and NADH subunit 5 + 6 genetic markers
Barnett *et al.* (2006b)	11 maternal haplotypes		India North Africa West Africa Central Africa Eastern–southern Africa Southern Africa	Control region (HVR1) genetic marker
Antunes *et al*. (2008)	nuclear data:	11 microsatellite groups	India	22 microsatellites, ADA, TF (autosomal), SRY (Y-chromosome), 12S, 16S (mitochondrial) genetic markers + assessment of prevalence and genetic variation of the lion-specific feline immunodeficiency virus (FIV)
		5 ADA haplotypes	East Africa	
		3 TF haplotypes	Southern Africa	
	SRY + mitochondrial data:	1 paternal haplotype	India	
		12 maternal haplotypes	North/Central Africa Southern/East Africa Southern Africa East Africa	
	6 FIV subtypes:		Southern/East Africa Southern Africa East Africa	

RSA = Republic of South Africa, ADA = Adenosine deaminase, TF = Transferrin, SRY = Sex-determining Region Y , FIV = Feline Immunodeficiency Virus

When setting up management strategies to preserve genetic variation in a species, one has to determine what (meta)populations efforts need to be focused on. When the existing taxonomy does not sufficiently reflect the genetic diversity, a smaller scale should be used, such as evolutionarily significant units (ESUs) or management units (MUs) (Moritz, 1994). The phylogenetic approach emphasizes protection of (meta)populations with a unique evolutionary history. Insight into the geographic pattern of genetic variation is not only crucial for management of wild populations, but also for breeding of captive stocks.

The lion is classified as 'Vulnerable' on the Red List of Threatened Species (IUCN, 2008), meaning that it faces "a high risk of extinction in the wild". Ninety per cent of the estimated continental population is located in Southern and East Africa (Chardonnet, 2002; Bauer & Van Der Merwe, 2004), with many large and stable lion populations. However, in West and Central Africa lion populations are generally small and isolated (Chardonnet, 2002; Bauer & Van Der Merwe, 2004). There is an increasing number of lions in peripherally isolated populations or in wildlife parks with little to no gene flow. Lions may face genetic erosion and inbreeding in these regions (Björklund, 2003). Several studies show that inbreeding depression is much more pervasive in wild populations than previously realized (Lacy, 1997; Hedrick & Kalinowski, 2000; Keller & Waller, 2002; Tallmon *et al.*, 2004) and it has been observed that there is a strong correlation between genetic variation and reproductive parameters in lions (O'Brien, 1994). The number of mature individuals in West Africa has been estimated by two separate surveys as 850 (Bauer & Van Der Merwe, 2004) and 1163 (Chardonnet, 2002), and the lion was therefore classified as 'Regionally Endangered' according to the IUCN criteria (Bauer & Nowell, 2004).

It is known that West and Central Africa have a different climatic history compared to Southern and East Africa, as West Africa and the northern part of Central Africa were characterized by hyperarid conditions during the Holocene glacial periods (Sarnthein, 1978; Klein & Martin, 1984; Dupont *et al.*, 2000; Gasse, 2000). This may have had a significant impact on local wildlife populations, related to climatic niches and food availability, possibly resulting in the development of distinct genetic lineages in this region. A dichotomy among genetic haplotypes between West and Central Africa and Southern and East Africa has been observed in seven African bovids (Arctander *et al.*, 1999; Nersting & Arctander, 2001; Pitra *et al.*, 2002; Van Hooft *et al.*, 2002), African elephant (*Loxodonta africana*) (Eggert *et al.*, 2002), cheetah (*Acinonyx jubatus*) (Freeman *et al.*, 2001), black rhinoceros (*Diceros bicornis*) (Brown & Houlden, 2000), roan antelope (*Hippotragus equinus*) (Alpers *et al.*, 2004) and giraffe (*Giraffa camelopardalis*) (Brown *et al.*, 2007). A similar genetic pattern is expected in lions, which would illustrate the need for stronger conservation efforts for the small and isolated West and Central African lion populations.

In this study we illustrate the phylogenetic relationships between lion populations from their entire geographic range, based on a sequence analysis of a large mitochondrial region. We know of only two previous studies that have included samples from West and Central Africa in phylogenetic analyses (Barnett *et al.*, 2006a, b). With information on the genetic makeup of lions from their West and Central African range, we may be able to conclude whether these form one or more distinct groups, with possible implications for a revised phylotaxonomy. This could have consequences not only for *in situ* wildlife management, but also for the management of zoo populations and for captive breeding programmes.

Materials and Methods

For this study, scat, hair, blood or tissue samples were obtained from wild ranging lions and from captive animals in zoos. In total, 53 individuals from 15 countries were sampled (Supplemental Table S1 in the Supporting Information), and 73 sequences from GenBank (Supplemental Table S2) were added at a later stage for phylogenetic analysis. Six samples, which are indicated with question marks in the table and figures, had a doubtful origin: Angola (no. 9), Democratic Republic of the Congo (DRC) (no. 10) and Somalia (nos. 7 and 19). In an earlier study the Moroccan lions from Rabat Zoo, which were originally thought to be descendants of the extinct Barbary subspecies, were identified to contain a haplotype from Central Africa (Barnett *et al.*, 2006a). The origin of all other lions or, in the case of captive lions, the origin of their ancestors, is known.

For this study, sequences of a mitochondrial region, containing the cytochrome *b* gene, tRNAPro, tRNAThr, and the left region of the control region, were analysed. The latter part contains the HyperVariable Region 1 (HVR1), which is the most variable part of the mitochondrial genome in the genus *Panthera* (Jae-Heup *et al.*, 2001).

DNA was extracted from tissue, blood, hair and scat samples. The targeted region was amplified using the primers shown in Supplemental Table S3. Details of extraction methods, polymerase chain reaction (PCR) amplification and sequencing are given in Supplemental Information S1.

Sequences were aligned visually and deposited in the GenBank database under accession numbers GU131164–GU131185, AY781195–AY781210, and DQ018993–DQ018996. Coding regions did not contain any stop codons or nonsense mutations, nor did they contain deletions or insertions that would lead to a frame shift. No known nuclear pseudogene insertions of cytoplasmic mitochondrial DNA sequences (NUMTs) were amplified.

To increase the sample size for the phylogenetic analysis, 28 cytochrome *b* sequences (sample group 4) from five countries (Dubach *et al.*, 2005) and 45 control region sequences (sample group 5) from 19 countries (Barnett *et al.*, 2006a) were obtained from GenBank (Supplemental Table S2). To gain more insight in the recent evolutionary history of the lion, control region sequences from extinct lion populations were also included. Figure 1 shows the localities of origin of the samples processed in our laboratory and the sequences obtained from GenBank that were combined for phylogenetic analyses. Sequences were divided into three sets for the analysis to obtain sequences of the same length (Table 2): cytochrome *b* + control region (A), cytochrome *b* (B), and control region (C). Samples of which only partial sequences could be obtained were either included in a subset of analyses or were completely excluded and were only used for direct sequence comparison (Supplemental Table S4). Sequences from the Moroccan samples (no. 18) contain an insertion of 80 bp (also visible as a longer PCR product on the gel). This insertion proved to be a duplication of 1382–1462 bp and was treated as one mutational event in every analysis. A second region, present in all samples (including the GenBank samples), was excluded from the analysis based on unknown homology. This region contains a repeat of cytosines of variable length at 1382–1393 bp.

For Bayesian and maximum likelihood analyses three outgroup species were added: two sequences of tiger (*Panthera tigris*: EF551003 and DQ151550), leopard (*P. pardus*: NC_010641) and snow leopard (*P. uncia*: EF551004). In addition, one sequence of extinct European cave lion (*P. leo spelaea*: DQ899900) and one sequence of extinct American cave lion (*P. leo atrox*: DQ899912) were added for the analysis of the control region.

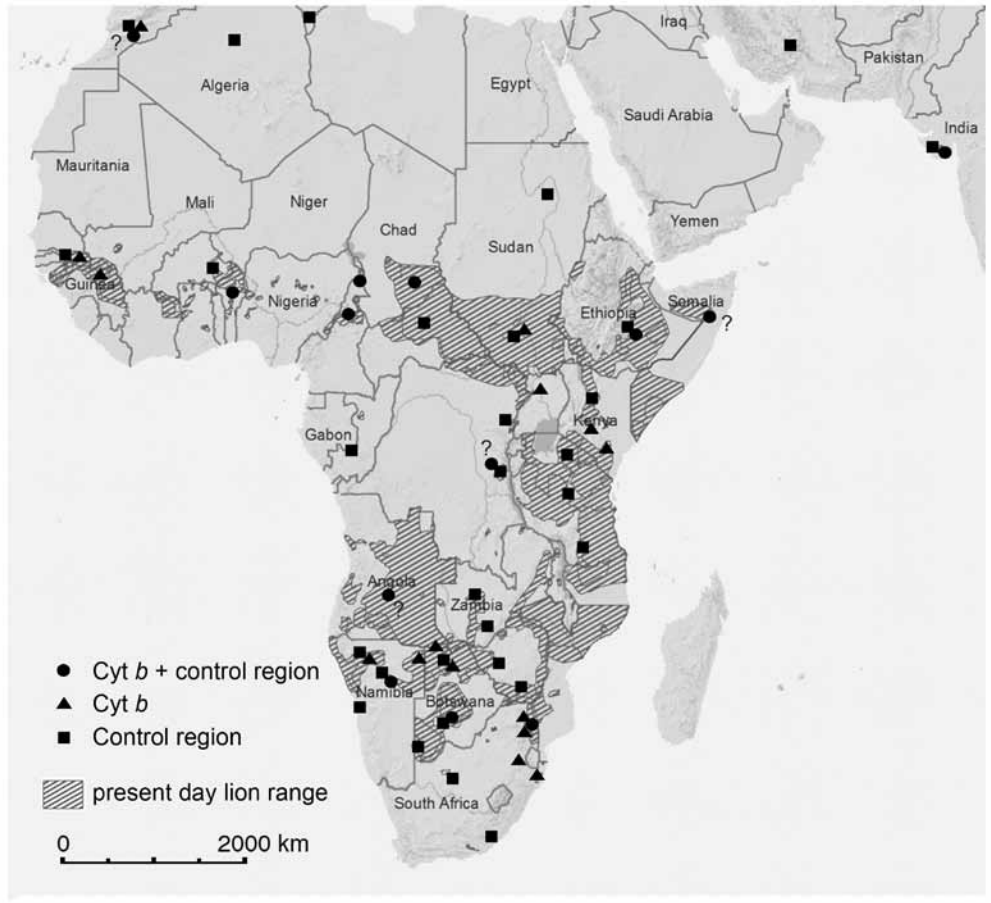


Figure 1. Map showing the origin of the lion (*Panthera leo*) samples that were used for the phylogenetic analyses. Dots indicate the samples from which cytochrome *b* (cyt *b*), tRNAThr, tRNAPro and the left domain of the control region sequences are known, triangles are cytochrome *b* sequences, and the squares show sample locations from which only a part of the control region was sequenced. For several samples only the country of origin was known, and there was no information available on the exact locality. In these cases the geographical centre of the known lion range within the country is indicated. Lion range data from IUCN (2008).

Table 2. Overview of the sets into which the lion (*Panthera leo*) sequences were subdivided for the phylogenetic analyses.

Sets for analyses	Genetic region	Position	Samples
A	Cytochrome *b*, tRNAPro, tRNAThr, control region	1-1764 bp	1-2a, 2c-15, 17-19
B	Cytochrome *b*	1-1140 bp	1-14, 16-23, 25-32
C	Control region	1355-1570 bp	1-2a, 2c-15, 17-19, 34-73

MrBayes v. 3.1.2 (Huelsenbeck & Ronquist, 2001) was used for the Bayesian analyses of each of the sets of sequences. The appropriate models for molecular evolution were determined using MrModeltest2 (v. 2.3) (Nylander, 2004). Stationary nucleotide frequencies of the HKY85 rate matrix were set to a flat Dirichlet distribution for the substitution rate priors and the state frequency priors.

The Markov chain Monte Carlo search was continued for 1,000,000 generations, sampling every 100 generations, and the first 2500 trees were discarded as burn-in.

Clusters of samples with an identical haplotype for the marker(s) studied were pooled and analysed as a single sample to reduce the time needed for analysis. Maximum likelihood (ML) analyses were performed using PAUP* 4.0 (Swofford, 2000). Heuristic ML searches [single random addition sequence, tree bisection–reconnection (TBR) without steepest descent] were performed for 100 bootstrap replicates. In each bootstrap replicate, all parameter settings were estimated by PAUP*, except for the base frequencies for which empirical data were used.

For each set of sequences, a haplotype network was generated, using TCS v. 1.21: phylogenetic network estimation using statistical parsimony (http://darwin.uvigo.es/software/tcs.html).

The samples from Angola (no. 9), DRC (no. 10), Somalia (nos. 7 and 19) and Morocco (nos. 18 and 20) were excluded from the isolation-by-distance analysis because of their doubtful origin (see above). Two matrices were generated for each of the sets of sequences: one with the genetic distances between the samples, expressed in the number of variable sites in the sequences, the other with the geographical (Euclidean) distance. For some samples only the country of origin was known. In these cases the coordinates of the geographic centre of the lion range within the country was chosen. The Isolation by Distance Web Service (IBDWS) v. 3.15 was used for performing a Mantel test for matrix correlation between genetic and geographic distance (http://ibdws.sdsu.edu/~ibdws/).

Results

The HKY85 model was chosen as the model for DNA evolution by MRMODELTEST, supported by hierarchical likelihood ratio tests and the Aikake information criterion for each of the sets of sequences. Rate variation across sites was modelled allowing invariable sites in all sets. Phylogenetic trees with posterior probability (PP) values derived from Bayesian analysis are shown in Figure 2 (cytochrome *b* + control region, and cytochrome *b* alone).

Phylogenetic trees derived from maximum likelihood (ML) analyses are shown in Figure 3 (cytochrome *b* + control region, and cytochrome b alone). Samples that share the same haplotype are joined on one branch, in clusters that are identical to the clusters found in the maximum parsimony analyses (see below).

Both Bayesian and ML analyses of cytochrome *b* + control region sequences (Figures 2A and 3A) support four basal clades: (1) the two Botswanan samples (PP >0.95; bootstrap value >70%), (2) a southern clade with lions from Namibia (PP >0.95; bootstrap value>70%) and the Republic of South Africa (RSA) (PP >0.95; bootstrap value >70%), (3) Ethiopian and Somalian samples (bootstrap value >70%), and (4) a geographically widespread clade, grouping lions from West and Central Africa, also including Angola and India (PP >0.95; bootstrap value >70%). In the ML analysis, the first three branches form a polytomy within the sister group of the widespread West and Central Africa clade, while in the Bayesian tree all branches have an equally basal position. Within the West and Central Africa group, the India clade is well supported in both analyses (PP>0.95; bootstrap value>70%). The branch leading to the rest of the group has significant branch support in the Bayesian analysis (PP >0.95), and the position of the two Benin samples in this clade remains unresolved in both analyses.

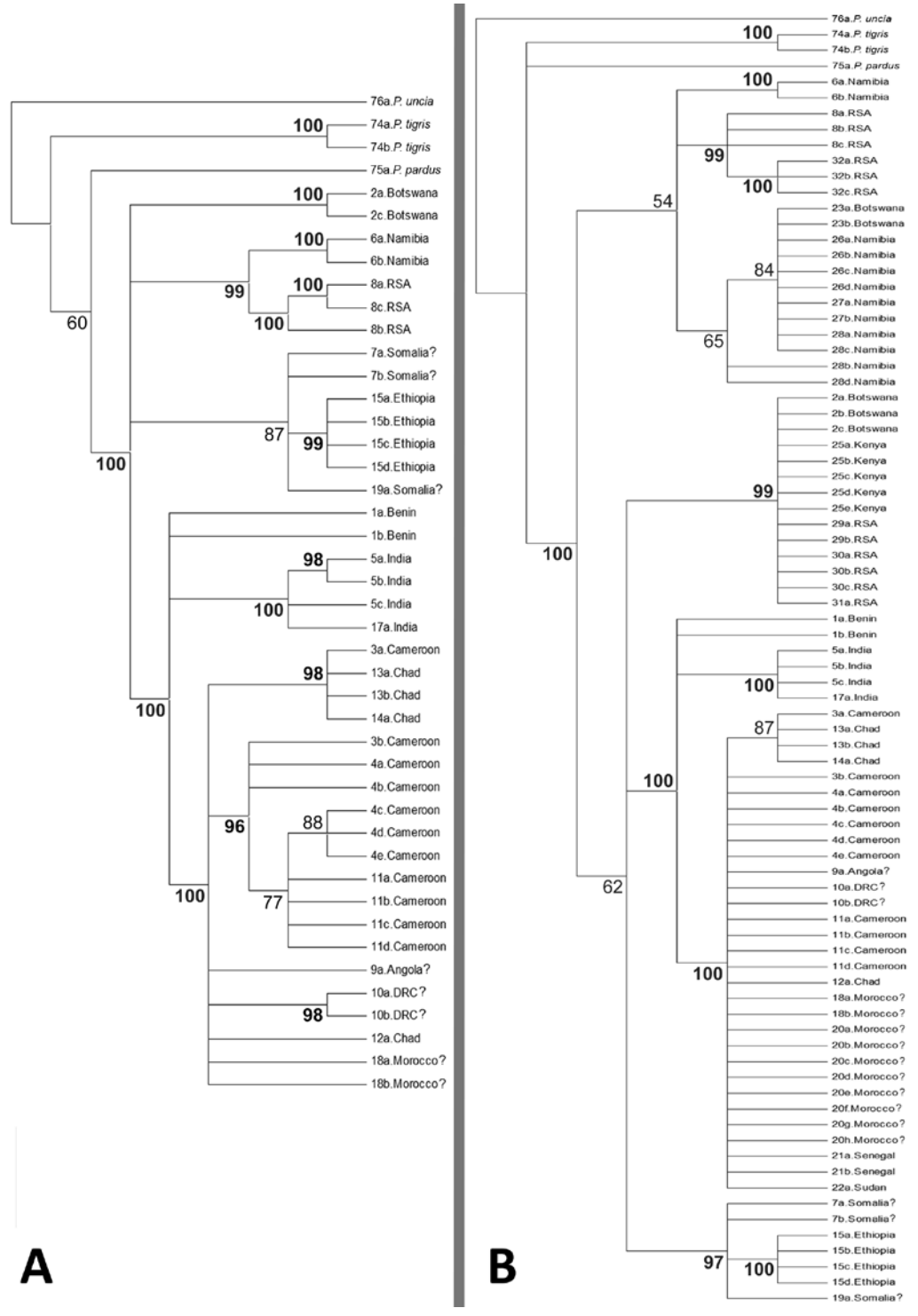


Figure 2. Phylogenetic trees resulting from Bayesian analysis of two sets of lion (*Panthera leo*) sequences: (A) cytochrome *b* + control region, (B) cytochrome *b* alone. The numbers represent the percentages for Bayesian posterior probability (PP). DRC = Democratic Republic of the Congo, RSA = Republic of South Africa.

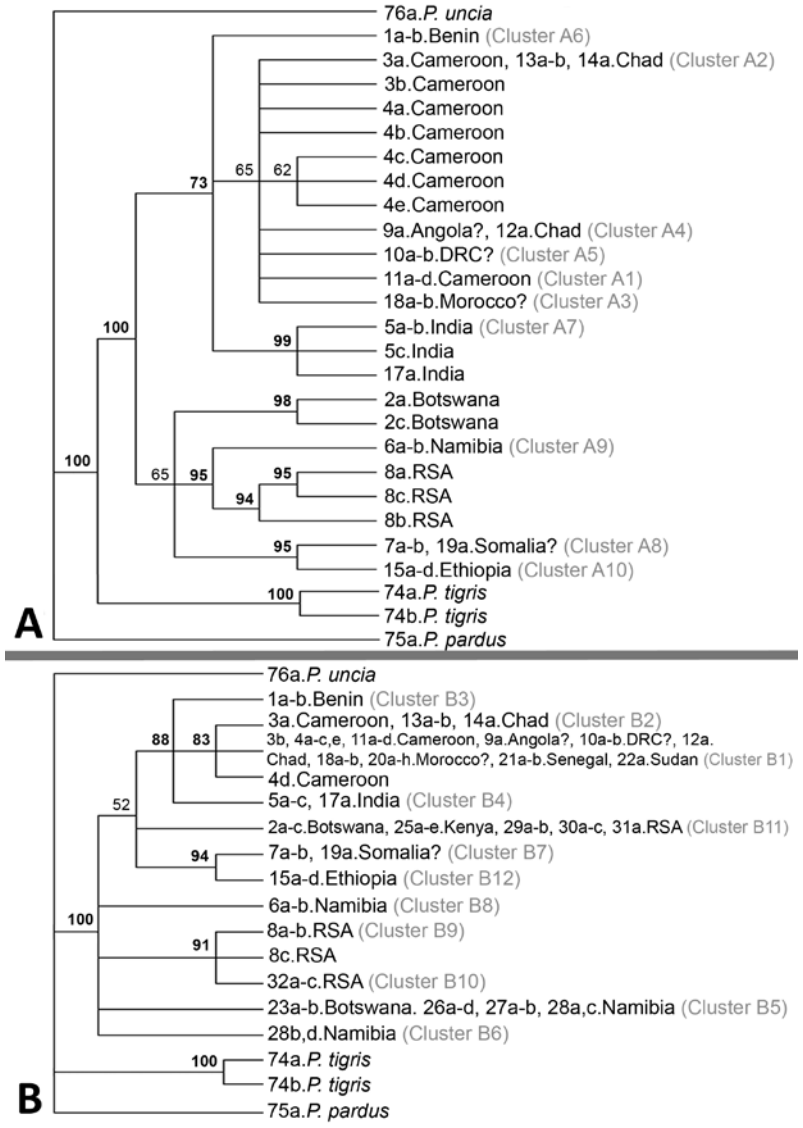


Figure 3. Phylogenetic trees resulting from maximum likelihood (ML) analysis of two sets of lion (*Panthera leo*) sequences: (A) cytochrome *b* + control region, and (B) cytochrome *b* alone. The numbers indicate the percentage for bootstrap support. Identical sequences were pooled. These clusters correspond with the clusters distinguished in the maximum parsimony analysis (Figure 4). DRC = Democratic Republic of the Congo, RSA = Republic of South Africa.

Among the West and Central Africa lions, the Bayesian analysis gives significant branch support (PP >0.95) for (1) a subclade with the two DRC samples, (2) a subclade containing ten Cameroon samples, and (3) a subclade with all Chad samples together with one from Cameroon.

The tree based on Bayesian analysis of cytochrome *b* sequences (Figure 2B) shows a basal split into two clades: (1) a clade from southern African countries, and (2) a clade with samples from West, Central and East Africa, plus Angola and India, and Botswana (no. 2) and RSA (nos. 29–31) samples.

There is significant branch support (PP >0.95) within the southern African clade for two out of three subclades: (1) a subclade containing two Namibia lions (no. 6), and (2) a subclade consisting of samples from RSA (no. 8) and RSA (no. 32) (PP >0.95). The ML analysis tree (Figure 3B) shows a basal polytomy with (1) samples from West, Central and East Africa, plus Angola and India, and Botswana (no. 2) and RSA (nos. 29–31), (2) Namibia (no. 6), (3) a well-supported branch containing RSA samples (no. 8 and no. 32) (bootstrap value >70%), (4) one containing sequences from Botswana and Namibia, and (5) Namibia (28b,d). The clade that contains the sequences from West, Central and East Africa shows significant branch support in both analyses (PP >0.95; bootstrap value >70%) for two subclades: (1) the Ethiopian (PP >0.95) and Somalian samples, and (2) a subclade containing all samples from West and Central Africa, including Angola and India. The third branch, leading to the samples from Botswana (no. 2), Kenya and RSA (nos. 29–31) is significantly supported by the Bayesian analysis (PP >0.95). Within the West and Central African subclade, the branch leading to the clade with the Indian lions is significantly supported in the Bayesian tree (PP >0.95) and the Benin samples have an unresolved position in both analyses.

The trees of the control region sequences (not shown) are not well resolved. The well-supported clades contain the two extinct cave lion subspecies *P. leo spelaea* and *P. leo atrox* (PP >0.95; bootstrap value> 70%), and in the case of the Bayesian analysis there is significant support (PP >0.95) for a branch with RSA samples (no. 8a,c).

A haplotype network was generated for each of the sets of sequences (Figure 4). The patterns resulting from analysis on cytochrome *b* + control region (Figure 4A) and from analysis on cytochrome *b* alone (Figure 4B) are strongly consistent. In both cases there is a clear distinction between West and Central African lions and Southern and East African lions, indicated by numerous mutations between the two groups. In general, variation amongst the West African lions is relatively small, with many individuals sharing the same haplotype, and little distance between the different haplotypes. Indian samples branch off close to the West and Central African group. As was the case in previous phylogenetic analyses, Angolan lions share their haplotype with (or cluster close to) lions from West and Central African countries. Lions from Southern and East African regions show more variation, illustrated by numerous mutations between the different haplotypes.

The haplotype network derived from the control region (Figure 4C) shows a more complex structure. A short loop is formed by the extinct lion populations from North Africa and the Middle East.

Because of partial sequences, Guinea (no. 16), Kenya (no. 24) and Uganda (no. 33) were excluded from these analyses. Comparing these partial sequences to the rest of the samples, it is very likely that the samples from Kenya and Uganda would cluster with the samples from Somalia and Ethiopia. The Guinean sample shows two-point mutations that are not present in any of the other sequences and an one-point mutation they only share with the Benin samples. Based on the rest of the sequence, Guinea is likely to be positioned close to samples from Benin and Cameroon. In all three cases the partial sequences seem to be related to sequences of close or neighbouring countries.

A Mantel test and a linear regression analysis were performed for matrix correlation between genetic and geographic distances for each set of sequences (Supplemental Figure S1). The R^2 value is the

highest for the analysis of cytochrome *b* + control region, 0.349. For cytochrome *b* alone and the control region (not shown), the R^2 values were 0.311 and 0.150, respectively. All these values are highly significant (*F*-test, $P < 0.0001$).

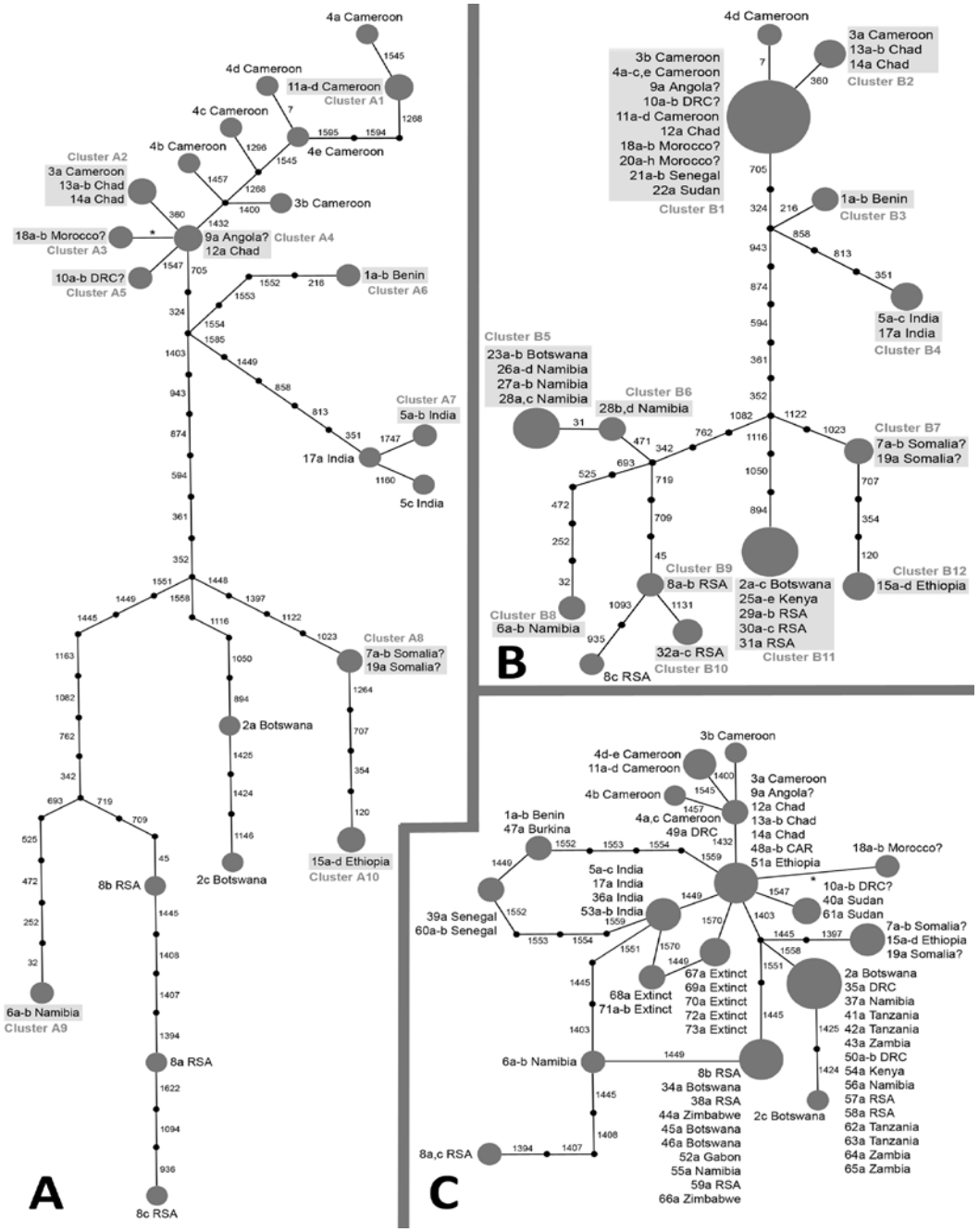


Figure 4. Haplotype networks for each of the analysed sets of lion (*Panthera leo*) sequences: (A) cytochrome *b* + control region, (B) cytochrome *b* alone, and (C) control region alone. The numbers indicate the location of each mutation. In (a) and (c), one of these mutations is indicated by *, representing the 80 bp insert found in two of the Moroccan samples. Clusters correspond to the pooled sequences used for maximum likelihood analysis (Figure 3). DRC = Democratic Republic of the Congo, CAR = Central African Republic, RSA = Republic of South Africa.

Discussion

In this study, the divergence of mitochondrial sequences of the cytochrome *b* gene, tRNAThr, tRNAPro and the left domain of the control region was assessed in a large number of lion individuals from different populations. The analyses are consistently showing similar patterns when using diverse algorithms. In general, samples from neighbouring countries cluster together and there is a distinction between West and Central Africa, and Southern and East Africa.

This can partially be explained by the unique climatological history of western Africa, leading to a dichotomy as has been witnessed in other African mammals (Arctander *et al.*, 1999; Brown & Houlden, 2000; Freeman *et al.*, 2001; Nersting & Arctander, 2001; Eggert *et al.*, 2002; Pitra *et al.*, 2002; Van Hooft *et al.*, 2002; Alpers *et al.*, 2004; Brown *et al.*, 2007). The low genetic diversity in and between the West and Central African lion populations indicate that they have a shorter evolutionary history than the more diverse Southern and East African lions. We hypothesize that this is caused by regional extinction, followed by recolonization. During the Late Pleistocene, 40-18 thousand years ago (ka), large parts of West and Central Africa were characterized by hyperarid conditions (Sarnthein, 1978; Dupont *et al.*, 2000; Gasse, 2000). The resulting lack of prey might have led to regional extinction of lions. This hypothesis is supported by several studies on large mammals, based on genetic research (Arctander *et al.*, 1999; Van Hooft *et al.*, 2002; Alpers *et al.*, 2004) and fossil data (Klein & Martin, 1984). This bottleneck in lion populations coincides with the well-known cheetah bottleneck (Menotti-Raymond & O'Brien, 1993; Driscoll *et al.*, 2002) and Late Pleistocene megafaunal extinctions that occurred over much of the globe (Cardillo & Lister, 2002; Barnosky *et al.*, 2004; Lyons *et al.*, 2004). More humid conditions 15–11 ka (Gasse, 2000) probably made recolonization of West and Central Africa possible.

Because of the strong relationship between West and Central African lions and Asiatic lions, it is likely that recolonization took place from refugia in close geographic proximity to India, which may have been located in the Middle East. Historical records suggest that there was a continuous Eurasian–North African lion population, which was distributed from Morocco through the Middle East to India (Blanford, 1876; Vogt & Specht, 1889; Flower & Lydekker, 1891). The extinction of the lion in Europe, Middle East and North Africa has effectively severed Asiatic lion gene flow to Africa (Mazák, 1970).

A complementary argument for the observed pattern in lion genetic diversity is the location of current natural barriers such as the African rainforest and the Rift Valley (Pitra *et al.*, 2002; Burger *et al.*, 2004; Dubach *et al.*, 2005; Barnett *et al.*, 2006b), as already proposed by Barnett *et al.* (2006b), and the connective Sahel savanna belt, which sustains numerous lion populations (Bauer & Van Der Merwe, 2004).

The Ethiopian samples show dispersion in the analyses: no. 15 clusters with samples from East Africa in every cytochrome *b* and cytochrome *b* + control region analysis, and, to a lesser extent, also in the control region analyses; however, no. 51 shows a closer genetic relationship to samples from West and Central Africa. It is possible that no. 15 comes from a population east from the Rift Valley, while no. 51 was sampled west of the Rift Valley, and is therefore connected to the Sahelian belt. Samples from DRC and Botswana also show some dispersion, reflecting genetic diversity within these countries. Botswana no. 2 groups with Kenya and RSA, while Botswanan no. 23 (Moremi GR) shows close genetic relationship with the Namibian samples. The same dichotomy has been described by Antunes *et al.* (2008).

Our results confirm that the lions that are thought to be of Moroccan origin share their haplotype mainly with Central African countries, which was already discussed by Barnett *et al.*, 2006b. Angola (no. 9) was positioned in the West and Central Africa group in every analysis. The Angolan sample shows little genetic relationship to samples from neighbouring countries such as Namibia, Botswana, Zambia, Tanzania and Uganda. Earlier published articles that include pedigrees (Steinmetz *et al.*, 2006) show that there is no certainty about the purity of the maternal line of the Angolan lions that are presently held in European zoos. This also explains why a similar pattern was found with the Angolan sample analysed by Antunes *et al.* (2008).

The isolation-by-distance analysis resulted in a highly significant correlation between genetic and geographic distance. A better model would be developed if possible migration routes as opposed to linear distances are used. Unfortunately, these routes are difficult to assess and probably changed extensively during the last millennia. We think that the inclusion of the Indian samples does not lead to an abnormally high correlation, as these samples show relatively little genetic differences when compared to West and Central Africa, despite the distance. The data points derived from the Indian samples do not form a separate group in the isolation-by-distance analysis, even in the analyses that do not include intermediate extinct lion populations from North Africa and the Middle East. It is also debatable if a linear model gives the best fit for the observed correlation, since it is expected that the variable sites in a genetic region can become saturated.

In this study lions from West and Central African countries are well represented, while samples from these regions were rare in other studies (Dubach *et al.*, 2005; Barnett *et al.*, 2006a,b; Antunes *et al.*, 2008; Mazák, 2010). West African countries were included in two previous studies, but only part of the control region was analysed, and samples connecting West to Central Africa were absent (Barnett *et al.*, 2006a,b). In general, a pattern was found of two major clades, one being located west of the Rift valley, and one confined to East and Southern Africa (Barnett *et al.*, 2006b). In the same study it was concluded that sub-Saharan lions are basal amongst modern lions, being in line with the high genetic diversity we observe in Southern and East Africa.

The data from Barnett *et al.* (2006b) seem to indicate that West African lions are more closely related to lions from Southern and East Africa, than they are to Central African lions. India falls between West and Central Africa, while one would expect West and Central Africa to be directly related. This pattern is less explicit after incorporation of these sequences to our data set.

Antunes *et al.* (2008) do not include any West or Central African countries. In the mDNA analysis samples from Angola, Morocco and Zimbabwe fall in one clade, close to the India clade. But all samples in this group were derived from captive individuals, and the Moroccan samples that were included are likely to contain a Central African haplotype as has previously been described by Barnett *et al.* (2006a). The purity of the Angolan lineage in the samples used by Antunes *et al.* (2008) is questionable, considering the pedigree of captive Angolan lions in European zoos (Steinmetz *et al.*, 2006). A similar explanation is hypothesized for the analysed Zimbabwean sample, which was also derived from a zoo. Sequences derived from wild-ranging Zimbabwean lions that were included in our study (control region) cluster with sequences from lions from neighbouring countries, and not with those from West and Central African lions.

In line with the pattern described by Dubach *et al.* (2005) we confirm the distinct position of

populations west of the Rift Valley, which were represented in the study of Dubach *et al.* by two sequences from Uganda. The distinct position of some RSA populations is also supported. The cytochrome *b* haplotype networks (MP analysis) shows that at least nine-point mutations in the cytochrome *b* gene make up the difference between lions from Timbavati and those from other regions in RSA.

We also support the conclusions of Mazák (2010), where one sub-Saharan Africa cluster and one North-Africa/Asia cluster are distinguished. Due to low sample size for West and Central African lions in that study, their taxonomic and phylogenetic position remained largely unresolved. Our data show that lions from this region should be considered to be part of a cluster also including North Africa and India.

The risk of extinction is often underestimated because all populations are considered to belong to a single (sub)species, and are managed as such. Management policies that are based on taxonomic divisions that insufficiently reflect genetic lineages within the taxon may lead to the disappearance of distinct lineages within the species. Therefore, it is important to focus on conservation strategies at a different scale, such as evolutionarily significant units (ESUs) or management units (MUs) (Moritz, 1994). These provide a rational basis for prioritizing populations for conservation.

In view of our results, we argue that the existing taxonomy, with the African and Asiatic lion as the only subspecies, does not sufficiently reflect the genetic diversity of this species. Several clades in Southern and East Africa show more variety in the studied genetic areas and show less relatedness to the West and Central African lions than to the Asiatic lion. Numerous subspecies are recognized in other African mammals which show this dichotomy (Arctander *et al.*, 1999; Nersting & Arctander, 2001; Pitra *et al.*, 2002; Van Hooft *et al.*, 2002; Eggert *et al.*, 2002; Freeman *et al.*, 2001; Brown & Houlden, 2000; Alpers *et al.*, 2004; Brown *et al.*, 2007).

In this study, 126 lion sequences were analysed using a number of phylogenetic approaches. The consistent pattern that emerged shows a clear distinction between West and Central African lions (including India) on the one hand, and Southern and East African lions on the other. This pattern is most likely to be explained by the climatological history of western Africa and current environmental connections and barriers for lion dispersal. The hyperarid conditions during Holocene glacial periods may have led to the regional extinction of the lion in West and Central Africa, followed by subsequent recolonization from refugia in the Middle East. This would explain why West and Central African lions seem to be closely related to Indian lions, and why they show relatively little genetic diversity. This may indicate that this is an evolutionarily young branch, in comparison to the Southern and East African lions, which show much more diversity.

Understanding the geographic pattern of genetic variation within species is critical for conservation management, not only for wild populations, but also for breeding of captive stocks. Most zoos only distinguish between accepted subspecies, which do not necessarily reflect the overall genetic diversity of the species. Based on our results, existing management strategies should be reconsidered and West and Central Africa's lions should not only be prioritized based on their current endangered situation, but also based on their genetic distinctness, their different level of genetic variation and their unique evolutionary history.

Acknowledgements

Samples were kindly provided by B. Chardonnet (France), N. Vanherle (CURESS, Chad), Saleh Adam and selected members of the West and Central African lion network (ROCAL), Ralph Buij and Barbara Croes (Cameroon), Dierenpark Amersfoort (The Netherlands), Diergaarde Blijdorp (The Netherlands), Zoo Basel (Switzerland), Safaripark Beekse Bergen (The Netherlands), Ouwehands Dierenpark (The Netherlands), Burgers' Zoo, (The Netherlands), Sanaa Zoo (Yemen), Antwerp Zoo – Planckendael (Belgium), Sables d'Olonne Zoo (France), Arabia's Wildlife Centre (United Arab Emirates) and Rabat Zoo (Morocco). We are grateful for the advise provided by Sarel van der Merwe (Chair) and selected members of the African Lion Working Group (ALWG). We also thank R. Glas for his assistance in processing of the samples. The investigations were (in part) supported by the Division for Earth and Life Sciences (ALW) with financial aid from the Netherlands Organization for Scientific Research (NWO). Further financial support that contributed to the provision of samples was received from WWF-INNO Fund, Mohamed bin Zayed Species Conservation Fund, Prins Bernhard Natuurfonds, Dutch Zoos Conservation Fund and the IUCN NL Ecosystem Grant Programme.

References

Alpers, D.L., Van Vuuren, B.J., Arctander, P. & Robinson, T.J. (2004) Population genetics of the roan antelope (*Hippotragus equinus*) with suggestions for conservation. *Molecular Ecology*, **13**, 1771-1784.

Antunes, A., Troyer, J.L., Roelke, M.E., Pecon-Slattery, J., Packer, C., Winterbach, C., Winterbach, H., Hemson, G., Frank, L.G., Stander, P., Siefert, L., Driciru, M., Funston, P.J., Alexander, K.A., Prager, K.C., Mills, G., Wildt, D., Bush, M., O'Brien, S.J. & Johnson, W.E. (2008) The evolutionary dynamics of the lion *Panthera leo* revealed by host and viral population genomics. *PLoS Genetics*, **4**, e1000251.

Arctander, P., Johansen, C. & Coutellec-Vreto, M.A. (1999) Phylogeography of three closely related African bovids (tribe Alcelaphini). *Molecular Biology and Evolution*, **16**, 1724-1739.

Barnett, R., Yamaguchi, N., Barnes, I. & Cooper, A. (2006a) Lost populations and preserving genetic diversity in the lion *Panthera leo*: implications for its ex situ conservation. *Conservation Genetics*, **7**, 507-514.

Barnett, R., Yamaguchi, N., Barnes, I. & Cooper, A. (2006b) The origin, current diversity and future conservation of the modern lion (*Panthera leo*). *Proceeding of the Royal Society B: Biological Sciences*, **273**, 2119-2125.

Barnosky, A.D., Koch, P.L., Feranec, R.S., Wing, S.L. & Shabel, A.B. (2004) Assessing the causes of late Pleistocene extinctions on the continents. *Science*, **306**, 70-75.

Bauer, H. & Nowell, K. (2004) Endangered classification of the lion in West Africa. *Cat News*, **41**, 3-36.

Bauer, H. & Van Der Merwe, S. (2004) Inventory of free-ranging lions *Panthera leo* in Africa. *Oryx*, **38**

Bauer, H., Nowell, K. & Packer, C. (2008) *Panthera leo*. *IUCN Red List of Threatened Species. Version 2010.4.* Available at: http://www.iucnredlist.org (accessed 12 April 2010).

Björklund, M. (2003) The risk of inbreeding due to habitat loss in the lion (*Panthera leo*). *Conservation Genetics*, **4**, 515-523.

Blanford, W.T. (1876) *Zoology and Geology,* Vol. II Macmillan and Co, London.

Brown, D.M., Brenneman, R.A., Koepfli, K.P., Pollinger, J.P., Milá, B., Georgiadis, N.J., Louis, E.E.J., Grether, G.F., Jacobs, D.K. & Wayne, R.K. (2007) Extensive population genetic structure in the giraffe. *BMC Biology*, **5**, 57.

Brown, S.M. & Houlden, B.A. (2000) Conservation genetics of the black rhinoceros (*Diceros bicornis*). *Conservation Genetics*, **1**, 365-370.

Burger, J., Rosendahl, W., Loreille, O., Hemmer, H., Eriksson, T., Gotherstrom, A., Hiller, J., Collins, M.J., Wess, T. & Alt, K.W. (2004) Molecular phylogeny of the extinct cave lion *Panthera leo spelaea*. *Molecular Phylogenetics and Evolution*, **30**, 841-849.

Cardillo, M. & Lister, A. (2002) Death in slow lane. *Nature*, **419**, 440-441.

Chardonnet, P. (2002) *Conservation of the African lion: contribution to a status survey.* International Foundation for the Conservation of Wildlife, Paris, and Conservation Force, Metairie, Louisiana, LA.

Driscoll, C.A., Menotti-Raymond, M., Nelson, G., Goldstein, D. & O'Brien, S.J. (2002) Genomic microsatellites as evolutionary chronometers: a test in wild cats. *Genome Research*, **12**, 414-423.

Dubach, J., Patterson, B.D., Briggs, M.B., Venzke, K., Flamand, J., Stander, P., Scheepers, L. & Kays, R.W. (2005) Molecular genetic variation across the southern and eastern geographic ranges of the African lion. *Conservation Genetics*, **6**, 15-24.

Dupont, L.M., Jahns, S., Marret, F. & Ning, S. (2000) Vegetation change in equatorial West Africa: time-slices for the last 150 ka. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **155**, 95-122.

Eggert, L.S., Rasner, C.A. & Woodruff, D.S. (2002) The evolution and phylogeography of the African elephant inferred from mitochondrial DNA sequence and nuclear microsatellite markers. *Proceeding of the Royal Society B: Biological Sciences*, **269**, 1993-2006.

Flower, W.H. & Lydekker, R. (1891) *An introduction to the study of mammals living and extinct*. Adam and Charles Black, London.

Freeman, A., Machugh, D., Mckeown, S., Walzer, C., Mcconnell, D. & Bradley, D. (2001) Sequence variation in the mitochondrial DNA control region of wild African cheetahs (*Acinonyx jubatus*). *Heredity*, **86**, 355-362.

Gasse, F. (2000) Hydrological changes in the African tropics since the Last Glacial Maximum. *Quaternary Science Reviews*, **19**, 189-211.

Hedrick, P.W. & Kalinowski, S.T. (2000) Inbreeding depression in conservation biology. *Annual Review of Ecology and Systematics*, **31**, 139-162.

Huelsenbeck, J.P. & Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754-755.

IUCN (2008) *Panthera leo*. *IUCN Red List of Threatened Species* (ed. by K. Nowell, C. Breitenmoser-Wursten, U.C.R.L.A. Breitenmoser and M.G.M.A.T. Hoffmann). IUCN, Gland, Switzerland

Jae-Heup, K., Eizirik, E., O'Brien, S.J. & Johnson, W.E. (2001) Structure and patterns of sequence variation in the mitochondrial DNA control region of the great cats. *Mitochondrion*, **1**, 279-292.

Keller, L.F. & Waller, D.M. (2002) Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, **17**, 230-241.

Klein, R.G. & Martin, P.S. (1984) *Mammalian extinctions and stone age people in Africa*. The University of Arizona Press, Tucson, AZ.

Lacy, R.C. (1997) Importance of genetic variation to the viability of mammalian populations. *Journal of Mammalogy*, **78**, 320-335.

Lyons, S.K., Smith, F.A. & Brown, J.H. (2004) Of mice, mastodons and men: human-mediated extinctions on four continents. *Evolutionary Ecology Research*, **6**, 339-258.

Mazák, J.H. (2010) Geographical variation and phylogenetics of modern lions based on craniometric data. *Journal of Zoology,* **281**, 194-209.

Mazák, V. (1970) The Barbary lion, *Panthera leo leo* (Linnaeus, 1758): some systematic notes, and an interim list of the specimens preserved in European museums. *Zeitschrift für Säugetierkunde*, **35**, 34-45.

Menotti-Raymond, M. & O'Brien, S.J. (1993) Dating the genetic bottleneck of the African cheetah. *Proceedings of the National Academy of Sciences USA* **90**, 3172-3176.

Moritz, C. (1994) Defining 'evolutionarily significant units' for conservation. *Trends in Ecology and Evolution*, **9**, 373-375.

Nersting, L.G. & Arctander, P. (2001) Phylogeography and conservation of impala and greater kudu. *Molecular Ecology*, **10**, 711-719.

Nylander, J.A.A. (2004) *MrModeltest v2.3*. Program distributed by the author. Evolutionary Biology Centre, Uppsala University, Uppsala.

O'Brien, S. (1994) A role for molecular genetics in biological conservation. *Proceedings of the National Academy of Sciences USA*, **91**, 5748-5755.

Pitra, C., Hansen, A.J., Lieckfeldt, D. & Arctander, P. (2002) An exceptional case of historical outbreeding in African sable antelope populations. *Molecular Ecology*, **11**, 1197-1208.

Sarnthein, M. (1978) Sand deserts during glacial maximum and climatic optimum. *Nature*, **272**, 43-46.

Steinmetz, A., Eulenberger, K., Thielebein, J., Buschatz, S., Bernhard, A., Wilsdorf, A., Grevel, V. & Ofri, R. (2006) Lens-anomalies and other ophthalmic findings in a group of closely-related Angola lions (*Panthera leo bleyenberghi*). *Zoo Biology*, **25**, 433 - 439.

Swofford, D.L. (2000) *PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4*. Sinauer Associates, Sunderland, MA.

Tallmon, D.A., Luikart, G. & Waples, R.S. (2004) The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology and Evolution*, **19**, 489-496.

Van Hooft, W.F., Groen, A.F. & Prins, H.H.T. (2002) Phylogeography of the African buffalo based on mitochondrial and Y-chromosomal loci: Pleistocene origin and population expansion of the Cape buffalo subspecies. *Molecular Ecology*, **11**, 267-279.

Vogt, C. & Specht, F. (1889) *The natural history of animals*. Blackie and Son, London.

Data accessibility

All sequence data generated in this study have been submitted to GenBank. Accession numbers are listed in Supplemental Table S1 and S2.

Supporting Information

Supporting information which is not included here may be found in the online version of this article and is available upon request.

Supplemental Table S1. Overview of the lion samples analysed in this study.

Samplegroup	No.	Female Ancestry	No. Individuals	Accession	Sequence	Type	Origin
1	1	Benin	2	GU131164 - GU131165	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	blood	Wild capture
	2	Botswana	3	GU131166 - GU131168	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	scat	Dierenpark Amersfoort, The Netherlands
	3	Cameroon - Bénoué NP	2	GU131169 - GU131170	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	blood	Wild capture
	4	Cameroon - Waza NP	5	GU131171 - GU131175	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	blood/ear	Wild capture
	5	India - Gir forest	3	GU131176 - GU131178	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	scat	Diergaarde Blijdorp, The Netherlands
	6	Namibia	2	GU131179 - GU131180	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	blood	Zoo Basel, Switserland
	7	Somalia?	2	GU131181 - GU131182	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	scat	Safaripark Beekse Bergen, The Netherlands
	8	RSA - Kruger NP - Timbavati GR	3	GU131183 - GU131185	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	blood	Ouwehands Dierenpark, The Netherlands
2	9	Angola?	1	AY781201	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	scat	Burgers' Zoo, The Netherlands
	10	DRC?	2	DQ018993 - DQ018994	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	liver	Diergaarde Blijdorp, The Netherlands
	11	Cameroon - Waza NP	4	AY781202 - AY781205	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	skin	Wild capture
	12	Chad - Zakouma NP	1	AY781200	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	blood	Wild capture
	13	Chad - Zakouma NP	2	AY781198 - AY781199	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	blood	Wild capture
	14	Chad - Zakouma NP	1	AY781197	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	blood	Wild capture
	15	Ethiopia	4	AY781207 - AY781210	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	blood	Sanaa Zoo, Yemen
	16	Guinea	1	DQ018996	partial Cytochrome *b* (bad quality sample)	scat	Wild capture
	17	India - Gir forest	1	AY781206	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	scat	Plankendael, Belgium
	18	Morocco?	2	AY781195 - AY781196	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	hair/scat	Sables d'Olonne Zoo, France
	19	Somalia?	1	DQ018995	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	blood	Breeding Centre, United Arab Emirates
3	20	Morocco?	8	DQ022294 - DQ022301	Cytochrome *b*	blood	Rabat Zoo, Morocco
	21	Senegal	2	DQ022291 + DQ022293	Cytochrome *b*	blood	Rabat Zoo, Morocco
	22	Sudan	1	DQ022292	Cytochrome *b*	blood	Rabat Zoo, Morocco

DRC = Democratic Republic of the Congo, RSA = Republic of South Africa. West Africa N=5, Central Africa N=18, East Africa N=7, Southern Africa N=9, North Africa=10, India N=4.

Supplemental Table S2. Overview of the lion sequences obtained from GenBank that were used to increase the sample size for the phylogenetic analyses.

Sample group	No.	Female Ancestry	No. Individuals	Accession	Sequence	Type	Origin
4	23	Botswana - Moremi GR	2	AF384815	Cytochrome *b*	Genbank entry	Dubach *et al.*, 2005
	24	Kenya - Aberdare NP	1	AF384809	Cytochrome *b*	Genbank entry	Dubach *et al.*, 2005
	25	Kenya - Tsavo East NP	5	AF384817	Cytochrome *b*	Genbank entry	Dubach *et al.*, 2005
	26	Namibia - Bushmanland	4	AF384813	Cytochrome *b*	Genbank entry	Dubach *et al.*, 2005
	27	Namibia - Caprivi Strip	2	AF384814	Cytochrome *b*	Genbank entry	Dubach *et al.*, 2005
	28	Namibia - Etosha NP	4	AF384811-AF384812	Cytochrome *b*	Genbank entry	Dubach *et al.*, 2005
	29	RSA - Fannie Roberts GR	2	AF384816	Cytochrome *b*	Genbank entry	Dubach *et al.*, 2005
	30	RSA - Hluhluwe-Umfolozi, NP	3	AF384818	Cytochrome *b*	Genbank entry	Dubach *et al.*, 2005
	31	RSA - Kapama GR	1	AF384816	Cytochrome *b*	Genbank entry	Dubach *et al.*, 2005
	32	RSA - Kruger NP - Sabi Sands	3	AF384810	Cytochrome *b*	Genbank entry	Dubach *et al.*, 2005
	33	Uganda	1	AF384809	Cytochrome *b*	Genbank entry	Dubach *et al.*, 2005
5	34	Botswana	1	DQ899922	control region (haplotype W)	Genbank entry	Barnett *et al.*, 2006
	35	DRC	1	DQ899921	control region (haplotype V)	Genbank entry	Barnett *et al.*, 2006
	36	India - Gir forest	1	DQ899919	control region (haplotype T)	Genbank entry	Barnett *et al.*, 2006
	37	Namibia	1	DQ899921	control region (haplotype V)	Genbank entry	Barnett *et al.*, 2006
	38	RSA	1	DQ899922	control region (haplotype W)	Genbank entry	Barnett *et al.*, 2006
	39	Senegal	1	DQ899918	control region (haplotype S)	Genbank entry	Barnett *et al.*, 2006
	40	Sudan	1	DQ899920	control region (haplotype U)	Genbank entry	Barnett *et al.*, 2006
	41	Tanzania	1	DQ899923	control region (haplotype X)	Genbank entry	Barnett *et al.*, 2006
	42	Tanzania - Serengeti	1	DQ899921	control region (haplotype V)	Genbank entry	Barnett *et al.*, 2006
	43	Zambia	1	DQ899921	control region (haplotype V)	Genbank entry	Barnett *et al.*, 2006
	44	Zimbabwe	1	DQ899922	control region (haplotype W)	Genbank entry	Barnett *et al.*, 2006
	45	Botswana	1	DQ248049	control region (haplotype 5)	Genbank entry	Barnett *et al.*, 2006
	46	Botswana - Moremi GR	1	DQ248049	control region (haplotype 5)	Genbank entry	Barnett *et al.*, 2006
	47	Burkina	1	DQ248047	control region (haplotype 3)	Genbank entry	Barnett *et al.*, 2006
	48	CAR	2	DQ248050	control region (haplotype 6)	Genbank entry	Barnett *et al.*, 2006
	49	DRC	1	DQ248051	control region (haplotype 7)	Genbank entry	Barnett *et al.*, 2006
	50	DRC - L. Edward	2	DQ248046	control region (haplotype 2)	Genbank entry	Barnett *et al.*, 2006
	51	Ethiopia	1	DQ248050	control region (haplotype 6)	Genbank entry	Barnett *et al.*, 2006
	52	Gabon	1	DQ248049	control region (haplotype 5)	Genbank entry	Barnett *et al.*, 2006
	53	India - Gir forest	2	DQ248053	control region (haplotype 9)	Genbank entry	Barnett *et al.*, 2006
	54	Kenya	1	DQ248046	control region (haplotype 2)	Genbank entry	Barnett *et al.*, 2006
	55	Namibia - Etosha Pan	1	DQ248049	control region (haplotype 5)	Genbank entry	Barnett *et al.*, 2006
	56	Namibia - Walvis Bay	1	DQ248046	control region (haplotype 2)	Genbank entry	Barnett *et al.*, 2006
	57	RSA - Kalahari	1	DQ248046	control region (haplotype 2)	Genbank entry	Barnett *et al.*, 2006
	58	RSA - Kalahari Gemsbok NP	1	DQ248046	control region (haplotype 2)	Genbank entry	Barnett *et al.*, 2006
	59	RSA - King William's Town	1	DQ248049	control region (haplotype 5)	Genbank entry	Barnett *et al.*, 2006
	60	Senegal	2	DQ248048	control region (haplotype 4)	Genbank entry	Barnett *et al.*, 2006
	61	Sudan - Nubia	1	DQ248052	control region (haplotype 8)	Genbank entry	Barnett *et al.*, 2006
	62	Tanzania - Serengeti	1	DQ248046	control region (haplotype 2)	Genbank entry	Barnett *et al.*, 2006
	63	Tanzania - Tanganyika	1	DQ248045	control region (haplotype 1)	Genbank entry	Barnett *et al.*, 2006
	64	Zambia	1	DQ248046	control region (haplotype 2)	Genbank entry	Barnett *et al.*, 2006
	65	Zambia - Kafue NP	1	DQ248046	control region (haplotype 2)	Genbank entry	Barnett *et al.*, 2006
	66	Zimbabwe - Tsholotsho	1	DQ248049	control region (haplotype 5)	Genbank entry	Barnett *et al.*, 2006
	67	Extinct - Barbary	1	DQ899916	control region (haplotype Q)	Genbank entry	Barnett *et al.*, 2006
	68	Extinct - Iran	1	DQ899917	control region (haplotype R)	Genbank entry	Barnett *et al.*, 2006
	69	Extinct - Algeria	1	DQ248055	control region (haplotype 11)	Genbank entry	Barnett *et al.*, 2006
	70	Extinct - Barbary	1	DQ248055	control region (haplotype 11)	Genbank entry	Barnett *et al.*, 2006
	71	Extinct - Iran	2	DQ248054	control region (haplotype 10)	Genbank entry	Barnett *et al.*, 2006
	72	Extinct - North Africa	1	DQ248055	control region (haplotype 11)	Genbank entry	Barnett *et al.*, 2006
	73	Extinct - Tunisia	1	DQ248055	control region (haplotype 11)	Genbank entry	Barnett *et al.*, 2006

DRC = Democratic Republic of the Congo, CAR = Central African Republic, RSA = Republic of South Africa. West Africa N=4, Central Africa N=9, East Africa N=13, Southern Africa N=36, North Africa/Middle East=8, India N=3.

Supplemental Table S3. Primers used for PCR amplification and sequencing.

Sample group	Region	Primername	Sequence (5'-3')	Origin
1	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	F: 1F	CGTTGTACTTCAACTATAAGAACTT	own design
		R: 1R	ATGGGATTGCTGATAGGAGATTAG	own design
		F: 2F	GTGGGGCCAAATATCCTTTT	own design
		R: 2R	GAAGGCCTAGGATATCTTTGATTG	own design
		F: 2bF	CATGAAACATTGGAATCGTATTGTTGTTC	own design
		R: 2bR	AGCTCTTTCGGACAGTTGAG	own design
		F: 3F	GACTCAGATAAAATTCCATTCCA	own design
		R: 3R	CATTATTCCTCGCTGTTTGG	own design
		F: 4F	CAATTATCCCTGCCCTCCA	own design
		R: 4R	TTTTTGGTTTACAAGACCAAGGTA	own design
		F: 5F	AAATCGCCTCCTCAAATGAA	own design
		R: 6R	AGCTCTTTCGGACAGTTGAG	own design
2	Cytochrome *b*, tRNAPro, tRNAThr, partial control region	F: L14724	CGAAGCTTGATATGAAAAACCATCGTTG	Cracraft *et al.*,1998
		R: H15915	AACTGCAGTCATCTCCGGTTTACAAGAC	Cracraft *et al.*,1998
		F: 1F	CGTTGTACTTCAACTATAAGAACTT	own design
		R: 1R	ATGGGATTGCTGATAGGAGATTAG	own design
		F: 2bF	CATGAAACATTGGAATCGTATTGTTGTTC	own design
		R: 2bR	AGCTCTTTCGGACAGTTGAG	own design
		F: 3bF	CCTATTCTCACCAGACCTATTAGGAGAT	own design
		R: 4bF	CCTGACCCTGACATGAATTG	own design
3	Cytochrome *b*	F: L14724	CGAAGCTTGATATGAAAAACCATCGTT	Irwin *et al.*, 1991
		R: CB141H	TGGCCCCACGGTAAGACATAT	Burger *et al.*, 2004
		F: CB17L	ATGGGATTGCTGATAGGAGGTTG	Burger *et al.*, 2004
		R: CB1912H	AAGGCCTAGGATATCTTTGATTGTA	Burger *et al.*, 2004
		F: CB19L	GATTCTTTGCCTTCCACTTCAT	Burger *et al.*, 2004
		R: CB211H	GAGGGCAGGGATAATTGCTAAG	Burger *et al.*, 2004
		F: CB10L	CCGCTACTAGGAATCAGAATA	Burger *et al.*, 2004
		R: H15915	AACTGCAGTCATCTCCGGTTTACAAGA	Irwin *et al.*, 1991

Supplemental Table S4. (online only) Overview of the variable sites of cytochrome *b*, tRNAThr, tRNAPro and the left domain of the control region.

Supplemental Information S1. Details of DNA isolation and sequence analysis.

Genetic analyses of sample group 1 (nos. 1-8) were performed at Leiden University (The Netherlands), sample group 2 (nos. 9-19) at the University of Antwerp (Belgium), and sample group 3 (nos. 20-23) at the Hillsdale College (USA).

Tissue, hair and scat samples were either kept in 100% ethanol or stored in a fridge or freezer for transportation. Blood samples from sample groups 2 and 3 were heparinized and those from sample group 1 were stored in a buffer (0.15 M NaCl, 0.05 M Tris-HCl, 0.001 M EDTA, pH = 7.5) at -80 °C.

For the DNA extraction from tissue and blood samples the DNeasy Blood & Tissue kit (Qiagen) (sample group 1 and 2), the Puregene kit (Gentra) (sample group 2) and the GenElut Blood Genomic DNA kit (Sigma) (sample group 3) were used. DNA extraction from the Moroccan hair and scat samples from sample group 2 was performed at the Dr. Van Haeringen Laboratorium (Wageningen), using a procedure based on guanidine thiocyanate and diatomaceous earth. From the scat samples from sample group 1 DNA was extracted following a protocol also used for ancient DNA extraction from bone and teeth (Rohland & Hofreiter, 2007).

In sample group 1, cytochrome *b*, tRNAThr, tRNAPro and the left domain of the control region were amplified from the blood and tissue samples using primers 1F-1R (first 439 bp) and 2bF-2bR (last 1326 bp) (Supplemental Table S3). DNA from scat and two blood samples, 6b (Namibia) and 8a (Republic of South Africa, RSA), proved to be degraded, so that internal primers had to be designed. For these samples DNA was amplified using 1F-1R, 2F-2R, 3F-3R, 4F-4R, 5F-6R (Supplemental Table S3). All internal primers were designed using the web-based software Primer3v. 0.4.0 (Rozen & Skaletsky, 2000). The PCR amplification profile that was used for all of these extractions included an initialization step of 94 °C for 4 minutes, 60 cycles of 20 seconds at 93 °C, 30 seconds at 55 °C and 30 seconds at 72 °C, ending with a final elongation step of 72 °C for 10 minutes and a final hold of 15 °C. Three samples from Waza NP and both Bénoué NP samples (both Cameroon) were sequenced in Leiden using the MegaBACE 1000 DNA automated analyzer (Amersham). The other sequence data were obtained from Macrogen Inc., Amsterdam, The Netherlands.

PCR amplification of DNA in sample group 2 was performed in either a single or two-step PCR with the Multiplex PCR kit (Qiagen, hot start, single PCR or first step) and the PCR Core System I kit (Promega, second step). In all cases negative controls were included. Cytochrome *b*, tRNAThr, tRNAPro and the left domain of the control region were amplified using primer pairs L14724 (Irwin *et al.*, 1991) - H15915 (cytochrome *b*) (Cracraft *et al.*, 1998), 1F-1R and 2bF-2bR (Supplemental Table S3). In the second PCR, 1 µl of PCR product was used as a template in a total volume of 50 µl. No multiple bands resulted from the PCR. Before sequencing, DNA products were cleaned with the GFX PCR DNA and Gel Band Purification kit (Amersham) and the EXO-SAP-IT kit (Amersham). Application of the latter kit and DNA sequencing was performed at the Genetic Service Facility of the Flanders Interuniversity Institute for Biotechnology (VIB, University of Antwerp). An Applied Biosystems 3730 DNA Analyzer in combination with ABI PRISM® BigDye™ Terminator cycle sequencing kit was used. For sequencing, two additional forward primers 3bF (starting at 757 bp) and 4bF (starting at 995 bp) were used (Supplemental Table S3).

PCR amplification in sample group 3 only targeted cytochrome *b* and was performed with PCR reagents from Invitrogen, using primer pairs L14724 (Irwin *et al.*, 1991) - CB141H, CB17L - CB1912H, CB19L - CB211H and CB10L (Burger *et al.*, 2004) - H15915 (Irwin *et al.*, 1991) (Supplemental Table S3). Before sequencing, DNA products were cleaned with Centricon YM-100 Centrifugal Filter Devices (Millipore). DNA sequencing was performed with the ABI PRISM® BigDye™ Terminator cycle sequencing kit and the resulting sequences were run on an Applied Biosystems 310 Genetic Analyzer (Hillsdale College).

The partial cytochome *b* sequences that are suspected to be of nuclear origin, described in Janczewski *et al.* (1995) and Hsieh *et al.* (2001) differ from their mitochondrial homologue by a large number (>11%) of point mutations, which is rare for mitochondrial cytochrome *b*. We observed none of this. Our sequences showed high sequence dissimilarity with another NUMT observed in cat and tiger described in Cracraft *et al.* (1998) and Kim *et al.* (2006). In only one case a lion nuclear pseudogene was amplified in an earlier stage of this study (not shown). The sequence showed 95% sequence similarity to a known tiger pseudogene and only 89% sequence similarity to validated lion sequences from mitochondrial origin. Pseudogene contamination among our mitochondrial DNA sequences is even more unlikely if considering the fact that some of our cytochrome *b* haplotypes were also observed in Dubach *et al.* (2005), where they specifically validated the mitochondrial origin of their sequences.

References

Burger, J., Rosendahl, W., Loreille, O., Hemmer, H., Eriksson, T., Gotherstrom, A., Hiller, J., Collins, M.J., Wess, T. and Alt, K.W. (2004) Molecular phylogeny of the extinct cave lion *Panthera leo spelaea*. *Molecular Phylogenetics and Evolution*, **30,** 841-849.

Cracraft, J., Feinstein, J., Vaughn, J. & Helm-Bychowski, K. (1998) Sorting out tigers (*Panthera tigris*): mitochondrial sequences, nuclear inserts, and conservation genetics. *Animal Conservation*, **1**, 139-150.

Dubach, J., Patterson, B.D., Briggs, M.B., Venzke, K., Flamand, J., Stander, P., Scheepers, L. & Kays, R.W. (2005) Molecular genetic variation across the southern and eastern geographic ranges of the African lion. *Conservation Genetics*, **6**, 15-24.

Hsieh, H.M., Chiang, H.L., Tsai, L.C., Lai S.Y., Huang N.E., Linacre A. and Lee J.C. (2001) Cytochrome *b* gene for species identification of the conservation animals. *Forensic Science International*. **122**, 7-18.

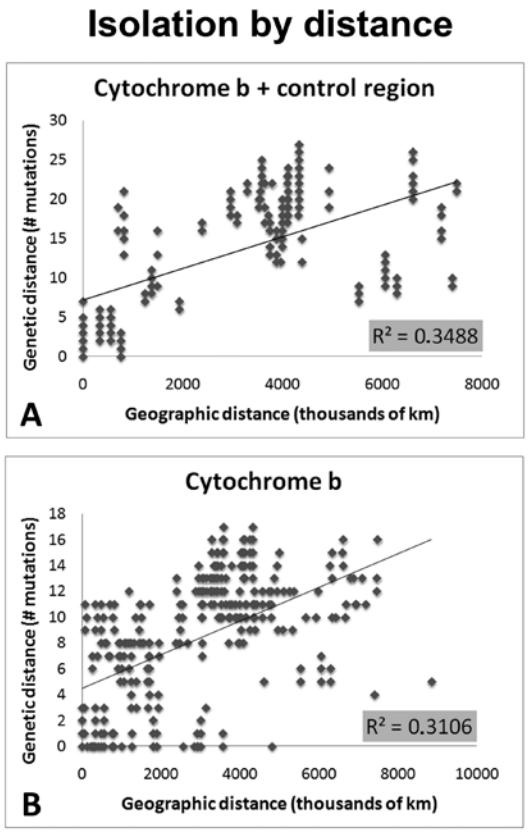
Irwin, D.M., Kocher, T.D. & Wilson, A.C. (1991) Evolution of the cytochrome *b* gene of mammals. *Journal of Molecular Evolution*, **32**, 128-144.

Janczewski, D., Modi, W.S., Stephen, J.C., O'Brien, S.J. Molecular evolution of mitochondrial 12S RNA and cytochrome *b* sequences in the pantherine lineage of Felidae. Molecular Biology and Evolution. 1995; 12 (4): pp. 690-707.

Kim, J.H., Antunes, A., Luo, S.J., Menninger, J., Nash, W.G., O'Brien, S.J. & Johnson, W.E. (2006) Evolutionary analysis of a large mtDNA translocation (*numt*) into the nuclear genome of the *Panthera* genus species. *Gene,* **366**, 292-302.

Rohland, N. & Hofreiter, M. (2007) Ancient DNA extraction from bones and teeth. *Nature Protocols*, **2**, 1756-1762.

Rozen, S. & Skaletsky, H. (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology*, **132**, 365-86.



Supplemental Figure S1. Graphs derived from the isolation-by-distance analysis, showing the relationship between geographic distance and genetic distance (R^2 values are added for linear regression) for two sets of lion (*Panthera leo*) sequences: (A) cytochrome *b* + control region and (B) cytochrome *b* alone.