The distribution of genetic diversity in a taxon provides us with information about the evolutionary history that has shaped the current genetic makeup. But it may also provide us with insights about the current situation and possible (genetic) threats it may be facing. Thereby it provides a rational basis for prioritising populations for conservation effort. A flagship species like the lion, with its vast geographic range and its current vulnerability, is a great study object to illustrate this.

GENETIC DIVERSITY IN THE LION (*Panthera leo* (Linnaeus 1758))

Laura D. Bertola

Universiteit Leiden
CML Institute of Environmental Sciences
ibl Institute of Biology Leiden
ISBN: 978-94-6203-801-1

GENETIC DIVERSITY IN THE LION (*Panthera leo* (Linnaeus 1758)):

Unravelling the Past and Prospects for the Future

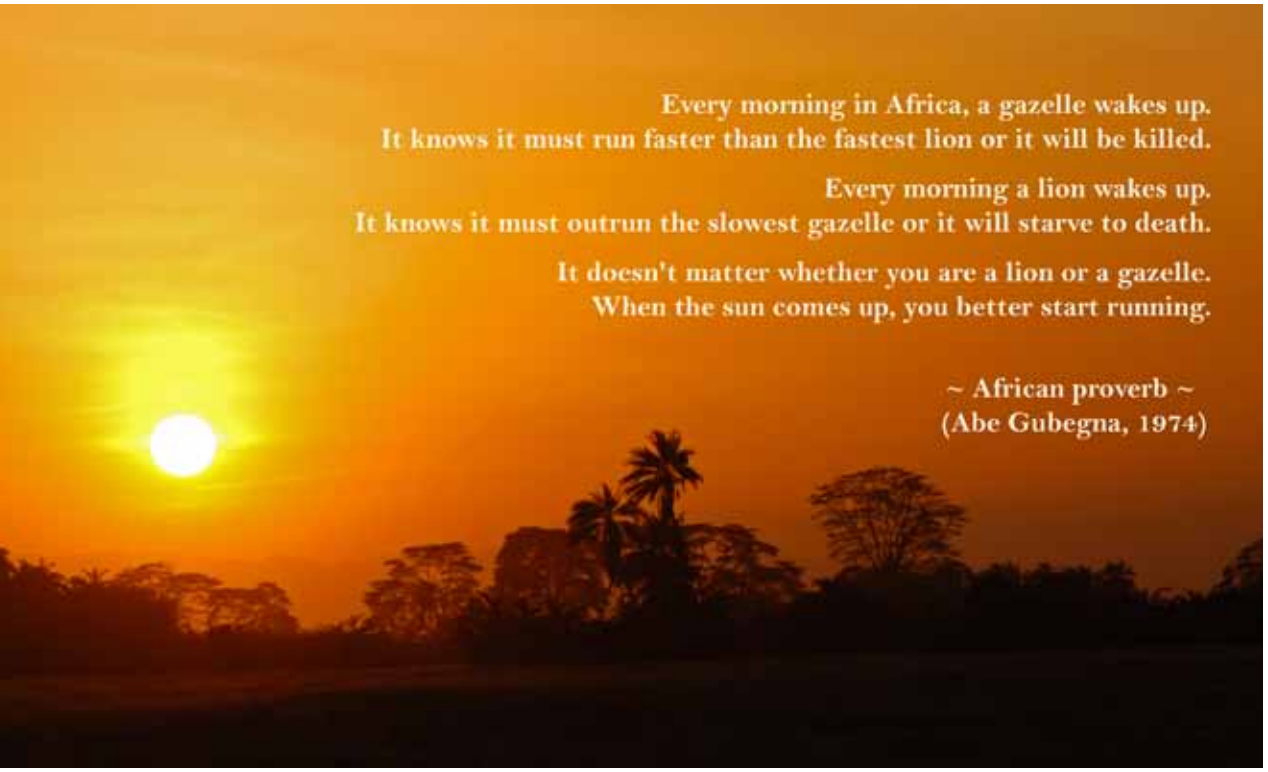
Laura D. Bertola

GENETIC DIVERSITY IN THE LION

(*Panthera leo* (Linnaeus 1758)):

Unravelling the Past and Prospects for the Future

Laura D. Bertola



Bertola, Laura D.

GENETIC DIVERSITY IN THE LION (*Panthera leo* (Linnaeus 1758)):
Unravelling the Past and Prospects for the Future

PhD thesis Leiden University, The Netherlands

Cover design	: Laura Bertola
Lay-out	: René Glas (www.reneglas.com)
Photos	: Laura Bertola
Printed	: Whörmann Print Service, Zutphen
ISBN	: 978-94-6203-801-1

GENETIC DIVERSITY IN THE LION
(*Panthera leo* (Linnaeus 1758)):
Unravelling the Past and Prospects for the Future

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 18 maart 2015
klokke 13:45 uur

door

Laura Diana Bertola

geboren te Delft
in 1984

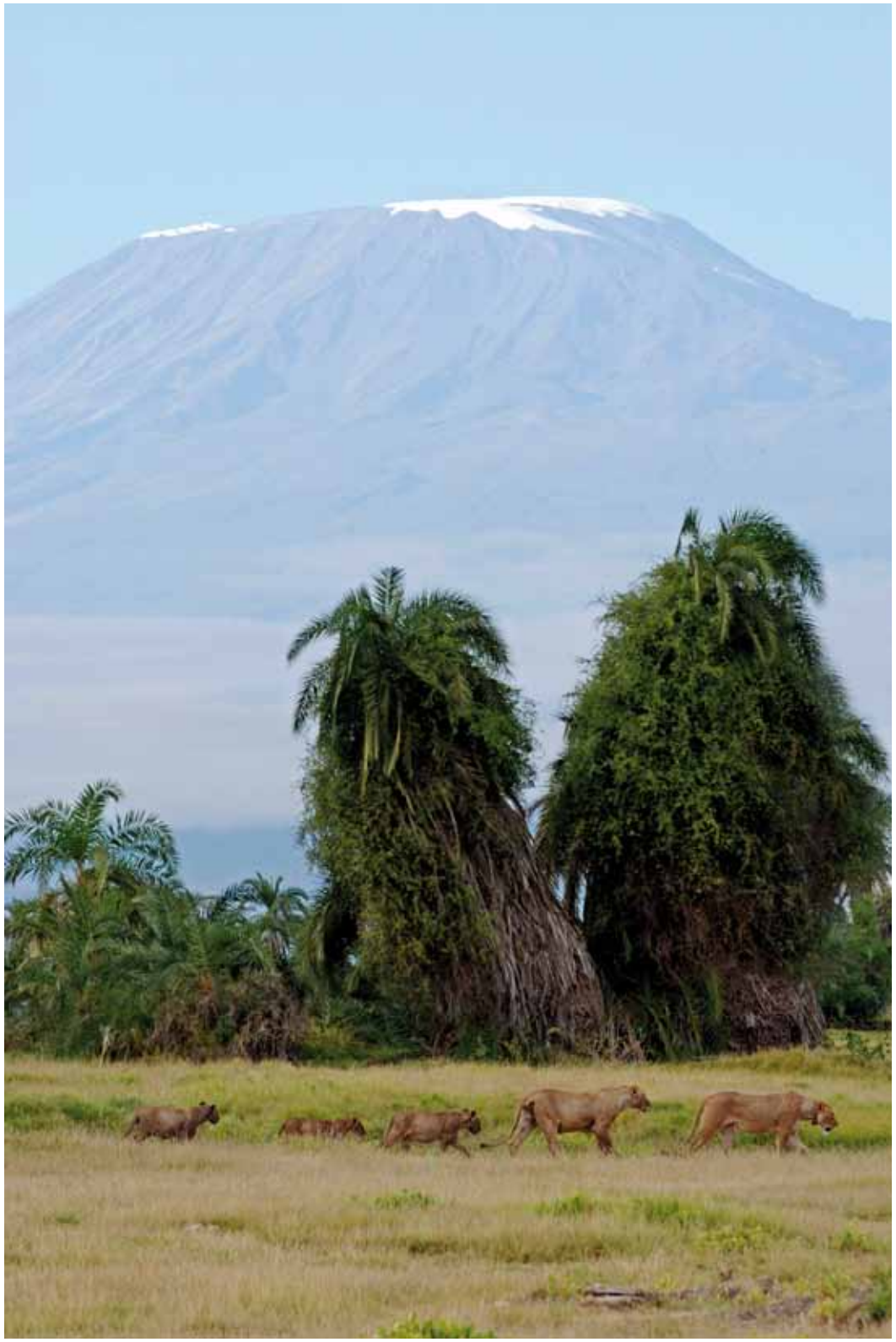
Promotie commissie

Promotor:	Prof. dr. Geert R. de Snoo (Leiden University)
Co-promotores:	Prof. dr. ir. Hans H. de Iongh (Leiden University / Antwerp University)
	Dr. Klaas Vrieling (Leiden University)
Overige leden:	Prof. dr. Michael Hofreiter (University of Potsdam)
	Prof. dr. Herbert H.T. Prins (Wageningen University)
	Dr. Ross Barnett (University of Copenhagen)

The investigations were supported by the Division for Earth and Life Sciences (ALW) with financial aid from the Netherlands Organization for Scientific Research (NWO) (project no. 820.01.002).

This research was conducted under the auspices of the Graduate School for Socio-Economic and Natural Sciences of the Environment (SENSE).

Contents



Chapter 1

General Introduction

Global biodiversity is rapidly declining, largely due to human-induced processes, such as land conversion, habitat degradation and overexploitation (Pimm *et al.* 1995, 2014; Pimm & Raven 2000; Butchart *et al.* 2010; Dirzo *et al.* 2014). The conservation of biodiversity is crucial since biodiversity contributes to stability and resilience in ecosystems, following the diversity-stability hypothesis (McNaughton 1977; Pimm 1984; Tilman & Downing 1994; Peterson *et al.* 1998; McCann 2000; MacArthur 2008). We depend on environmental services, provided by well-functioning ecosystems (Millenium Ecosystem Assessment 2005). The scale at which biodiversity can be studied, ranges from entire ecosystems to intraspecific diversity at the genetic level (UNEP-CBD 2006). Insights into the spatial distribution of this diversity provide us with information regarding evolutionary processes which have shaped these patterns, and may provide us with guidelines on how to conserve it.

Genetic diversity at the intraspecific level further contains the evolutionary potential, and therefore, to some degree, the resilience of a species. The goal to preserve species as dynamic entities with the potential to adapt to changing environmental factors, illustrates why recognizing the genetic diversity contributes to conservation planning. This follows from the notion that conservation should not focus on preserving particular objects (species or ecosystems), but rather on preserving the processes of life (Bowen 1999). In addition, genetic diversity may also reflect other types of diversity of importance for conservation, e.g. ecological, behavioural, morphological and demographical diversity, and provides natural resources for breeding.

The aim of this thesis is to gain insight into the intraspecific genetic diversity of an ecological and cultural umbrella species: the lion (*Panthera leo* (Linnaeus, 1758)). The distribution of this diversity reflects the evolutionary history on one hand, but may also provide guidance on how to retain this for the future on the other. The lion serves as a model for illustrating general phylogeographic patterns on the African continent, and provides us with a framework in which we can make recommendations for conservation practices.

The importance of conserving carnivores

Many carnivores have suffered from population declines and range contractions in the past two centuries (Woodroffe 2000; Ceballos & Ehrlich 2002; Ray *et al.* 2005b; Morrison *et al.* 2007; Ripple *et al.* 2014). At the same time, large carnivores are known to play an important role in maintaining rich and resilient ecosystems (Fretwell 1987; Miller *et al.* 2001; Terborgh *et al.* 2001; Sala 2006; Johnson *et al.* 2007; Bruno & Cardinale 2008; Letnic *et al.* 2009, 2012; Beschta & Ripple 2009; Schmitz *et al.* 2010; Strong & Frank 2010; Estes *et al.* 2011). Elimination of top predators from a community, may lead to the reorganization of trophic webs and biodiversity loss, following from trophic cascade theory (Hairston *et al.* 2010) and mesopredator release (Crooks & Soulé 2010). Downstream extinctions of other species have been observed both in community models (Borrvall & Ebenman 2006) and in natural situations (Johnson *et al.* 2007; Letnic *et al.* 2009).

Top predators typically have large home ranges, which they require to fulfil their ecological needs (Schaller 1972). Therefore, they are often considered umbrella species, following the notion that protection of species with large ranges indirectly encompasses populations of co-occurring species (Noss 1990). In addition, large predators are especially sensitive to human activity and have been actively persecuted in most regions of the world, since their ecological role often conflicts with that of local people (Woodroffe 2000; Treves & Karanth 2003; Patterson *et al.* 2004; Woodroffe *et al.* 2005; Bauer *et al.* 2010; Sogbohossou *et al.* 2011; Yirga *et al.* 2012; Tumenta *et al.* 2013). Because of these characteristics, they are more strongly affected by edge effects, often occurring at the borders of protected areas (Woodroffe 1998). Due to the ecological importance and the demanding requirements of large carnivores, it is generally advocated that ensuring that an ecosystem can sustain populations of top predators, this is likely to also benefit other species. This makes carnivores suitable model species for defining and testing conservation strategies.

The African continent is home to a diverse assemblage of carnivores. The African large carnivore guild is made up of seven species: African wild dog (*Lycaon pictus*), spotted hyena (*Crocuta crocuta*), striped hyena (*Hyaena hyaena*), brown hyena (*Hyaena brunnea*), cheetah (*Acinonyx jubatus*), leopard (*Panthera pardus*) and lion (*Panthera leo*) (Dalerum *et al.* 2008). All seven species show decreasing trends and only the spotted hyena is attributed with a Least Concern status according to the IUCN global Red List (IUCN 2014). Despite the fact that all these species fulfil their role as a top predator, they exhibit a high functional diversity as a result of different ecological preferences and different prey spectra (Hayward & Kerley 2008). A study using data from the Serengeti ecosystem, collected over 40 years, has shown that predation on populations is not only affected by the abundance, but also by the diversity of predators, concluding that the loss of this diversity could disrupt important interactions (Sinclair *et al.* 2003). The conservation of ecological interactions has been stressed since the introduction of the key-stone species concept, according to which certain species have a disproportionately large effect on their environment relative to their abundance (Paine 1966, 1969). Targeting species-rich assemblages of large carnivores, which cover high phylogenetic and high functional diversity (Dalerum 2013), follows and expands this notion. The largest species generally represents an important aspect of unique functional diversity, and the lion in particular shows the largest contribution to functional diversity in the global assemblage (Dalerum 2013). These results provide justification for prioritizing this species for conservation efforts.

There are few studies with empirical data on the effect of extirpation of lions on the rest of the ecosystem. Long term data from six protected areas in Ghana show that after extinction of lions and leopards in three of the included areas, the landscape was opened up for olive baboons (*Papio anubis*), which strongly increased in number and range (Brashares *et al.* 2010). Although baboons are only moderately predated by lions and leopards, the absence of an apex predator had a large influence on both the behaviour and population numbers of olive baboons. This negatively affected population numbers of small primates and ungulates, and diet analysis showed that in the absence of a top predator, baboons shift towards a more carnivorous diet (Brashares *et al.* 2010). It illustrates that the presence of a predator, even though it might occur in low densities, may have a profound effect on an ecosystem. The potentially far-reaching effects of large predators was previously shown for wolves (*Canis lupus*) in Yellowstone National Park, also including behavioural changes in prey species, termed “the ecology of fear” (Brown *et al.* 1999). Similar forces may act on the African

savannah ecosystem, although monitoring studies in a range of lion habitats could provide more detailed data on the effect of losing this species from the system.
Above mentioned arguments illustrate the need for conservation actions for top predators in general, and the suitability of the lion as a model species in particular. The lions continent-wide distribution and its extension into Asia provides a framework which can be used to study phylogenetic relationships on a large geographic scale. Finally, the lion, generally perceived as the "king of the beasts", is a true cultural flagship for carnivore conservation.

The lion

The lion (*Panthera leo* (Linnaeus, 1758)) is one of the five big or "roaring" cats in the genus *Panthera*. In the current taxonomy, two subspecies are officially recognized: the African lion *Panthera leo leo* (Linnaeus, 1758) and the Asiatic lion *Panthera leo persica* (Meyer, 1826). The species is classified as 'Vulnerable' on the basis of criterion A2abcd on the IUCN Red List of Threatened Species (Bauer *et al.* 2012), defined as "Population reduction observed, estimated, inferred, or suspected in the past where the causes of reduction may not have ceased OR may not be understood OR may not be reversible." (IUCN 2012). For distinct geographic regions other categories are more suitable, e.g. West Africa where lions are categorized as 'Regionally Endangered' (Bauer & Nowell 2004) with the recent suggestion to uplist the West African lion as 'Critically Endangered' (Henschel *et al.* 2014). The Asiatic subspecies is categorized as 'Endangered', based on criterion D (number of mature individuals < 250) (Breitenmoser *et al.* 2008).

Although extinct subspecies of the lion ranged as far as North America and the northern part of South America during the Middle and Late Pleistocene, the range of the modern lion was restricted to Africa, near Asia and the Southern part of Europe (Turner & Antón 1997; Yamaguchi *et al.* 2004; Schnitzler 2011). Lions went extinct in Europe between 3000 and 1000 BC, likely to be the result of human persecution. Later, they disappeared in the Near East, Arabian Peninsula, Trans-Caucasia and the north of Afghanistan around the 12th and 13th century AD, and in all North African countries and the Middle East between the end of the 19th century and the first part of the 20th century (Guggisberg 1961; Schnitzler 2011). The last remaining population outside of Africa is located in the Gir forest National Park in India, being the sole representatives of the Asiatic subspecies (Breitenmoser *et al.* 2008; Bauer *et al.* 2012).

Several studies have aimed to estimate the number of remaining African lions, using a scope of different methods (Table 1). IUCN/SSC Cat Specialist Group members made "guesstimates" of 30,000 – 100,000 wild African lions in the early 1990s (Nowell & Jackson 1996a). The African Lion Working Group (ALWG), which is also affiliated with the IUCN/SSC Cat Specialist Group, conducted a mail survey, which resulted in an estimation of 23,000 lions in protected areas in Africa, with a range of 16,500–30,000 (Bauer & Van der Merwe 2004). Chardonnet (2002) based his estimates on the extrapolation of known populations estimates into areas where lion status was unknown, resulting in an estimation of 39,000 lions, with a range of 29,000–47,000 (Chardonnet 2002). Other studies used a GIS-based model to predict the range and numbers of the African lion (Ferreras & Cousins 1996; Riggio *et al.* 2012). Following the most recent estimate, it is expected that there are 32,000 - 35,000 free-ranging African lions, in 67 areas (Riggio *et al.* 2012). Because of methodological differences between the estimates, a direct comparison of these figures is bothersome. However, a group exercise led by the IUCN/SSC Cat Specialist Group estimated a decline in lion numbers for 42% of the major lion populations (IUCN SSC Cat Specialist Group 2006a; b). Also according to the IUCN Red List data the African lion is indicated as having a decreasing population trend (Bauer *et al.* 2012). The Asiatic populations numbers around 350 individuals and the populations trend is stable (Breitenmoser *et al.* 2008).

Table 1. Studies estimating the number of free-ranging African lions.

Year of prediction	Authors	Number of African lions	Method
1980	Ferreras and Cousins (1996)	75,800	GIS-based model
Early 1990s	Nowell and Jackson (1996a)	30,000-100,000	"Guesstimate"
2002	Chardonnet (2002)	39,000 (range: 29,000-47,000)	Extrapolation from known populations
2004	Bauer & Van der Merwe (2004)	23,000* (range: 16,500-30,000)	Mail survey
2012	Riggio *et al.* (2012)	32,000-35,000	GIS-based model

* protected areas only

The decrease in lion numbers is paralleled by the decline of suitable lion habitat, as is the case for many other species (Riggio *et al.* 2012; Ripple *et al.* 2014). The current range of the African lion is estimated to comprise only 17% of its historical range (Ray *et al.* 2005a; Ripple *et al.* 2014). Comparing the current state to the situation in 1960, using human population density and land conversion data, it was calculated that, at best, 3,390,821 km^2 or about 25% of the original savannah area remain as potential range of free-ranging lions (Riggio *et al.* 2012). The sub-Saharan human population is estimated to have increased nearly four-fold, from 229 million in 1960 to 863 million in 2010 (CIESIN & CIAT 2005). Human population projections further predict this number to double to 1.753 billion in 2050 (CIESIN & CIAT 2005), suggesting that human encroachment of habitat suitable for the lion and other wildlife will be a major challenge conservationists are facing. Apart from land conversion, decline of natural prey base and increased conflict between carnivores and humans are the main threats for the lion (Ferreras & Cousins 1996; Chardonnet 2002; Ray *et al.* 2005b; Winterbach *et al.* 2012). A modelling study has shown that human population and reserve size together accounted for 98% in the variation of extinction of 41 mammal species between reserves in West Africa (Brashares *et al.* 2001). Extinctions rates in six reserves were 14-307 times higher than those predicted by species-area models and reserve size alone, indicating the impact of human demography on local wildlife (Brashares *et al.* 2001). Particularly alarming is that data from the Masai Mara Reserve and its surroundings show that land conversion outside the protected area has resulted in approximately equal declines in wildlife both inside and outside the reserve (Ottichilo *et al.* 2000; Newmark 2008).

A total of 86 Lion Conservation Units (LCUs) was determined on the African continent (IUCN SSC Cat Specialist Group 2006a; b) by expert-defined criteria which classified these areas as important for lion conservation, following an approach previously applied to the jaguar (*Panthera onca*) (Sanderson *et al.* 2002) (Figure 1). The total area of these LCUs was calculated as being 3,163,260 km^2 and is in line with the estimate of suitable lion habitat, being 3,390,821 km^2 (Riggio *et al.* 2012). An estimated 588,000 km^2 are designated as hunting zones, where lions can be harvested as trophies (Lindsey *et al.* 2013). Although hunting zones may act as corridors and avert land conversions into e.g. agricultural land, they have also shown disturbed lion population structure, possibly extending the effect into the adjacent protected area (Loveridge *et al.* 2007; Croes *et al.* 2011; Lindsey *et al.* 2013). In addition to the decline in suitable habitat, remaining areas are often poorly protected and increasingly fragmented (Newmark 2008). This fragmentation and human activities, such as hunting, may lead to the creation of sinks (i.e. sites with unusually high mortality rates for wildlife) in the human-dominated matrix surrounding the protected areas, possibly affecting the populations within the protected area (Loveridge *et al.* 2007; Newmark 2008).

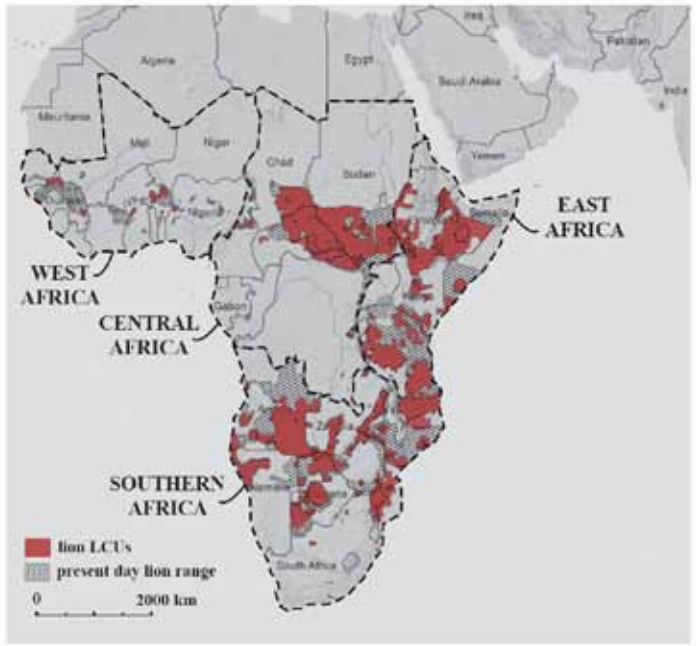


Figure 1.
Map showing delineation of the regions according to the Lion Conservation Strategies (IUCN SSC Cat Specialist Group 2006a;b) and the recognized LCUs in red (Panthera). Lion range data are from IUCN (2014).

It must be noted that only a relatively small part of the total lion range is located in West and (the western part of) Central Africa. The habitat in this region is generally smaller and more isolated, compared to East and Southern Africa. Also the declines in habitat are the strongest in West (39%) and Central Africa (18%), compared to East (16%) and Southern Africa (9%) (Chardonnet 2002). This is paralleled with exceptionally strong declines in wildlife in West Africa (Craigie *et al.* 2010) and some parts of Central Africa (Bouché *et al.* 2012). After recent surveys, resident lion populations could only be confirmed in four out of 16 LCUs in West Africa (Henschel *et al.* 2014) and were absent in all three Central African LCUs included in an earlier study (Henschel *et al.* 2010). This means that, at least for this region, the indicated lion range in Figure 1 may be an overestimation. On the African continent Riggio *et al.* (2012) identified ten lion strongholds, meeting the necessary requirements for long-term viability of the lion population. This entails that 1) the area contains at least 500 individuals, 2) the area is located within protected areas or designated hunting areas, and 3) the numbers of lions is stable or increasing as assessed by the IUCN Cat Specialist Group (Riggio *et al.* 2012). Of these strongholds four are located in East Africa, six in Southern Africa and none have been identified in West or Central Africa (Riggio *et al.* 2012).

Despite the recent and ongoing contraction of lion habitat, the species still exhibits a vast range, covering almost the entire sub-Saharan continent (with the exception of dense rain forest and dry desert), and one location in India. For many large mammals with a similar range, numerous subspecies are recognized and for several phylogenetic data are available (for an overview see Chapter 4). Distributions of diversity, either described in the taxonomy or as a phylogeographic pattern, are often congruent (Hewitt 2004; Lorenzen *et al.* 2012). This leads to the notion that current lion nomenclature, considering all African lion populations to belong to a single subspecies, should be questioned, as it may not reflect the full underlying diversity.

Phylogeography

The term phylogeography was defined in 1987 by Avise *et al.* (1987), aiming to bridge the gap between population genetics and systematics. Demographic events, like population expansions and bottlenecks, as well as migration patterns leave their traces in the genetic makeup of species and populations. Environmentally driven evolution entails that similar forces have determined the evolutionary histories and thereby shaped analogous phylogeographic patterns of multiple species, depending on their ecological requirements. Studying the patterns of co-distributed taxonomic groups, often termed “comparative phylogeography”, may lead to previously unrecognized biogeographic patterns and contribute to guiding conservation decisions (Bermingham & Moritz 1998; Moritz & Faith 1998; Arbogast & Kenagy 2001). Distinct genetic clades may be the result of retraction and subsequent expansion of populations into and from refugia during the cyclical climatic events (Hewitt 2000, 2004). Combining genetic data with climatic data and ecological niche modelling provides us with insights into evolutionary forces acting upon the species. This type of information may also contribute to recognizing risks for long term conservation of a species or population.

Although the species is the general unit for conservation practices (Mace 2004), the general aim to protect the full diversity embedded in the species leads to the notion that it is important to also include units below the species level. This is particularly important for species with a large range. Conserving the maximum of genetic diversity within a species, also entails that the evolutionary potential is maintained. Safeguarding the adaptability of species is of utmost importance in a changing environment, for example as a result of shifting climate conditions (Visser 2008). However, subspecies are generally delimited as geographic variants and do not necessarily reflect information on adaptive genetic diversity.
Phylogeographic data can be used to define such sub-specific units of importance for species conservation, such as Evolutionary Significant Unit (ESU) and Management Units (MU). The term Evolutionary Significant Unit was first proposed by Ryder (1986) and aimed at finding a rational basis for prioritizing units “possessing genetic attributes significant for present and future generations”. Although the precise definition of ESU is disputed (for reviews see: De Guia & Saitoh, 2007; Fraser

& Bernatchez, 2001), the concept has widely been used for conservation studies (Randi *et al.* 2003; Alpers *et al.* 2004; Duriez *et al.* 2006; Hansen *et al.* 2009; Hu *et al.* 2011; Höglund *et al.* 2013) and applied in legal and management contexts (Waples 1991). In general, it is suggested that ESUs ought to be geographically discrete and should display concordant divergence in both molecular and non-molecular traits (Ryder 1986; Dizon *et al.* 1992; Vogler & DeSalle 1994). Adaptive variation as well as divergence, as a reflection of the evolutionary history, are typically included in the definition of ESU, but genetic criteria range from significant divergence of allele frequencies (Waples 1991), to some level of genetic distance (Ryder 1986), to congruently structured phylogenies of genes (Avise & Ball 1990). Common criticism is that the concept is subjective, by aiming to conserve a group of populations which is e.g. "substantially reproductively isolated from other conspecific populations", and "represents an important component in the evolutionary legacy of the species" (Waples 1991), which are guidelines that are difficult to implement in practice.

A later, and commonly used, definition included that an ESU contains a group of populations that is "reciprocally monophyletic for mtDNA alleles" and "shows significant divergence of allele frequencies at nuclear loci" (Moritz 1994). To avoid an overly restrictive definition and to be able to address the cases with less separation than reciprocal monophyly, Moritz (1994) also proposed the term Management Unit (MU), defined as "populations with significant divergence of allele frequencies at nuclear or mitochondrial loci, regardless of the phylogenetic distinctiveness of the alleles.". These populations represent groups that are connected with such low levels of gene flow that they are functionally independent and although they show divergence in their allele frequencies, it is not necessarily the case that there is a phylogenetic distinctiveness of the alleles. Following these general definitions, ESUs are concerned with historical population structure, phylogeny and long-term conservation needs. In contrast, MUs address current population structure, allele frequencies and short-term management issues.

Despite numerous definitions and attempts to create a unifying concept (e.g. Fraser & Bernatchez, 2001), the role of neutral genetic markers versus adaptive diversity, concordance with other sources of information such as ecological, morphological and physiological data, and the applicability to differentiate units along the evolutionary continuum of populations, are still under debate (Crandall *et al.* 2000). All definitions that have been proposed are characterized by strong points on one hand, and inherent uncertainties on the other. The applicability of these concepts should be assessed at a case-by-case basis, and conservationists should aim to apply a flexible and integrative approach (Fraser & Bernatchez 2001).

As was previously mentioned, the African lion (*P. l. leo*) and the Asiatic lion (*P. l. persica*) are the only officially recognized subspecies according to the IUCN (Bauer *et al.* 2012). However, differentiation within Africa, distinguishing between lions from East/Southern Africa and West/Central Africa, and the close relationship of the latter to North Africa/Asia, was noted based on a range of morphometric data (Hemmer 1974). Results of a more recent study on craniometric data corroborated the close relationship of West African, North African and Asiatic lions (Mazák 2010). Historically, up to eight lion "subspecies" have been recognized based on intraspecific morphological variation and up to 24 synonyms circulate (Haas *et al.* 2005; Dubach *et al.* 2005; ISIS 2014). Initial descriptions of subspecies have not always adequately treated age- and sex-related variation, and conclusions should therefore be treated with caution (Dubach *et al.* 2005). Two of these eight subspecies , *P. l. leo* and *P. l. melanochaita*, are considered to be either extinct or a con(sub)specific with and extant subspecies (Barnett *et al.* 2006a). Additionally, Hemmer (1974) suggests to combine the four subspecies traditionally recognized in East/Southern Africa into two subspecies, representing East and Southern Africa. This classification leads to three main "subspecies" in Africa: *P. l. senegalensis* (West and Central Africa), *P .l nubica* (East Africa) and *P. l. melanochaita* / *krugeri* (Southern Africa), in addition to the Asiatic subspecies *P. l. persica* (also see Chapter 6).

Several phylogeographic studies have been conducted over the years, aiming to gain insight into the level of distinctiveness between populations and corroboration for the subspecies status of the Asiatic lion. The phylogenetic studies support the single-African-origin model, as was originally proposed based on morphology, distribution and parietal art (Yamaguchi *et al.* 2004). Long branches and the position of the outgroups point the evolutionary cradle of the lion towards East and Southern Africa (Barnett *et al.* 2006b; Antunes *et al.* 2008), which is further supported by the oldest fossil evidences that were found in this region (Petter 1973; Lewis 1997; Werdelin *et al.* 2010). The distinctiveness of the Asiatic subspecies was addressed by O'Brien *et al.* (1987), using allozyme separation data, although genetic distances between the subspecies was low and led to the conclusion that African and Asiatic lions shared a relatively recent common ancestor. Moreover, studies from which the subspecies distinction could be inferred, only included lions from India and from East and Southern Africa and therefore suffered from incomplete sampling of the full species' diversity (O'Brien *et al.* 1987; Burger *et al.* 2004). On the African continent, three major refugial areas have been proposed, deduced from current phylogeographic patterns in large mammals: West/Central Africa, East Africa, and Southern Africa (Hewitt 2004; Lorenzen *et al.* 2012). This may imply that lions from West/Central Africa also represent a different genetic clade, compared to their East and Southern African counterparts. Due to the knowledge gap for West and Central African lions, and the urgent conservation need in this region, these populations are of particular importance (Bauer 2003).

Most studies done so far have been based on mitochondrial DNA (mtDNA). Initially, within East and Southern Africa four distinct groups were distinguished: 1) South-West, 2) East of the Rift Valley, 3) West of the Rift Valley and 4) the Sabi Sands population (Dubach *et al.*, 2005). These four groups are discussed in the light of future conservation interventions, such as translocations to mimic gene flow, and are therefore proposed as distinct ESUs sensu Crandall (2000) (Dubach *et al.* 2005). Barnett and colleagues used a different part of the mtDNA for their phylogenetic studies and included several populations from West and Central Africa, as well as currently extinct populations from North Africa and the Middle East (Barnett *et al.* 2006b). They concluded that five main clades can be distinguished in the modern lion: 1) North Africa–Asia, 2)West Africa, 3) Eastern Sahel (steppe/savannah areas immediately south of the Sahara), 4) Eastern–Southern Africa, and 5) South- Western Africa. According to the authors, these groups could be interpreted both as ESUs or as MUs (sensu Moritz, 1994), although the authors also stress the scale-dependency of identifying units crucial for conservation. Barnett *et al.* (2006b) proposed basal geographic distinctions between lions from 1) North Africa-Asia, 2) Middle Africa and 3) Southern Africa. This pattern may have been shaped and maintained by natural barriers for lion dispersal, such as the Sahara desert, dense tropical rain forest and the Rift Valley (Nowell & Jackson 1996b; Burger *et al.* 2004; Dubach *et al.* 2005), as well as historic events, such as climatological changes (Barnett *et al.* 2006b, 2014). Later phylogeographic studies by these groups continuously added data to the existing dataset and were able to assess the intraspecific genetic diversity of the lion on a finer scale and with higher certainty. Dubach *et al.* (2013) confirmed

the distinct position of West & Central African lions, and their close genetic relationship with the Asiatic subspecies. However, they were not able to fully resolve the phylogenetic relationships in this clade, and branches obtained low support in the more diverse East and Southern African branches (Dubach *et al.* 2013). Although Dubach *et al.* (2013) also include microsatellite data for all sampled populations, these were analysed to detect admixture and gene flow, rather than phylogeographic origin (based on the number of identified clusters). Barnett *et al.* (2014), again including samples from currently extinct populations, were the first to describe a basal split in the lion phylogeny, although the branch of the Southern clade does not show significant support and the position of the Asiatic subspecies remains unresolved. Five ESUs (sensu Moritz, 1994) are proposed in absence of conflict with the limited morphological or nuclear DNA data: 1) North Africa/Asia, 2) West Africa, 3) Central Africa, 4) South Africa and 5) East-South Africa. Antunes *et al.* (2008) also included nuclear genetic markers and seroprevalence data of Feline Immunodeficiency Virus (FIV_{Ple}) to infer the evolutionary dynamics of the lion. In this study no free-ranging populations from West and Central Africa had been included and the distinct position of the Asiatic subspecies receives limited support. Interestingly, in the nuclear data there is a clear split observed between East and Southern African lions, whereas in the mtDNA haplotypes this split is less distinct. Although Antunes *et al.* (2008) do not extensively address the conservation implications of these results and the applicability of ESUs or MUs, they do state that due to the differentiation within the African lion "a bottom-up perspective that prioritizes populations, rather than large-scale units (e.g. all African lions), might preserve and maintain lion diversity and evolutionary processes most efficiently."

The consensus of the phylogeographic studies is that intraspecific diversity of the African lion is greater than the current taxonomy implies, and that different lineages on the African continent can be recognized. In addition the distinct status of the Asiatic subspecies is questioned, regarding the partially unresolved phylogeny of the northern lion populations. The phylogenetic position of the West and Central African lion is of particular interest, notably because of the vulnerable position of populations in this region. As several of the studies discussed in this section were published during the course of this project, the relationship between these studies and the datasets that were generated for this thesis are discussed in the individual chapters.

Conservation genetics and Inbreeding

In the field of conservation genetics, genetic methods are applied to guide conservation and restoration practices. Compared to phylogeography, conservation genetics is typically applied on a lower geographic scale, namely on that of a (meta)population. Levels of genetic diversity on this geographic scale are not so much related to evolutionary potential, but rather to direct fitness measures.

The loss of genetic variability is strongest in small, isolated populations, where genetic drift plays a relatively large role and deleterious mutations accumulate faster as a result of increasing levels of relatedness. Reduction in population size may ultimately lead to inbreeding with consequent fitness effects, such as lower fecundity and higher mortality, termed inbreeding depression (Wright 1977). The effect of inbreeding has been illustrated in several laboratory settings and showed that inbred populations display decreased fitness and are particularly vulnerable to environmental stress, thereby suffering from an increased extinction risk (Bijlsma *et al.* 2000; Reed & Frankham 2003; Armbruster & Reed 2005). In addition, the phenomenon has been studied in domestic and captive-bred wild populations (Ralls & Ballou 1986; Ralls *et al.* 1988; Lacy *et al.* 1993). Although these situations cannot be directly compared to free-ranging, wild populations, Ralls *et al.* (1988) conclude that "the total costs of inbreeding in natural populations are probably considerably higher than our estimates", which would make the effect of genetic diversity or inbreeding an important driver in natural populations.

Despite of this, the role of genetic factors in the process of extinction in wild populations has long been questioned, following the argument that environmental and demographic stochastic processes, as well as catastrophes, would drive small populations to extinction before the effects of inbreeding would become visible (Caughley 1994; Caro & Laurenson 1994). An extensive review of wild populations of mammals, birds, poikiloterms (snakes, fish and snails) and plants compared 169 estimates of inbreeding depression for 137 traits, and concluded that wild populations under natural conditions frequently exhibit moderate to high levels of inbreeding depression (Crnokrak & Roff 1999). More publications stressed the detrimental effects of low genetic diversity in natural populations and the risk of genetic erosion to long term survival (Keller & Waller 2002; Frankham 2005, 2010). The fact that the extinction risk significantly increased with decreasing heterozygosity, an indication of inbreeding, was shown in a field study on butterflies (Saccheri *et al.* 1998). A comparison of genetic diversity between 170 threatened and taxonomically related, non-threatened species showed that average heterozygosities were lower in the threatened species for 77% of comparisons (Spielman *et al.* 2004). Since reduced fitness often reveals itself as impaired fertility (Keller & Waller 2002), sperm abnormalities and motility were assessed in 20 mammal species (Fitzpatrick & Evans 2009). Species with reduced mean heterozygosity also showed impaired ejaculate quality, with a stronger positive correlation for endangered populations/(sub)species (Fitzpatrick & Evans 2009).

In the fertility study by Fitzpatrick & Evans (2009), three lion populations were assessed, including two populations well-known for their low levels of genetic diversity: the lions of the Ngorongoro crater in Tanzania and the lions of the Gir forest in India. Both inbred lion populations, and the outbred population from the Serengeti plains, had been subjected to a physiological study, which showed that reduced genetic diversity correlates well with an increased incidence of abnormal sperm and with decreased testosterone levels, a hormone crucial for normal spermatogenesis (Wildt *et al.* 1987). Other studies have shown that lion populations with reduced genetic diversity are more susceptible to disease, making them more vulnerable to extinction (Kissui & Packer 2004; Trinkel *et al.* 2011).

The well monitored lions of the Ngorongoro crater in Tanzania illustrate how fast inbreeding can play a role in a natural population. An outbreak of the biting fly *Stomoxys calcitrans* in the early sixties reduced the population from around 70 to 10 individuals, consisting of nine females and one male (Packer *et al.* 1991). Seven males immigrated into the crater in the late 50s and early 60s, but apart from this event, no immigration occurred in the following 25 years. The population recovered, and in 1975 a number of 75-125 lions was estimated to be in the Ngorongoro crater. However, all animals were descendants of only 15 founder individuals (seven females and eight males) and genetic diversity in this population proved to be much lower than in the neighbouring Serengeti ecosystem (O'Brien *et al.* 1987; Yuhki & O'Brien 1990; Packer *et al.* 1991), as well as above mentioned reproductive measures (Wildt *et al.* 1987).

The isolated lion population in the Gir forest, India, the sole representative of the Asiatic subspecies, is also characterized by extremely low heterozygosity levels. The Gir forest population may have suffered from isolation even before the extinction of populations connecting the Asiatic to the African subspecies (Driscoll *et al.* 2002). In addition, the population underwent a severe bottleneck as a result of sport hunting and habitat encroachment and as few as 20 individuals may have been the only survivors in the early twentieth century (see Driscoll *et al.*, 2002, and references herein). Several genetic markers, including allozymes (O'Brien *et al.* 1987), MHC-RFLP (Yuhki & O'Brien 1990), minisatellites (Gilbert *et al.* 1990) and microsatellites (Driscoll *et al.* 2002), confirmed that the Gir forest lions are genetically further impoverished than the Ngorongoro crater lions and even unrelated individuals showed identical minisatellite DNA fingerprints (Gilbert *et al.* 1990). This strong reduction in genetic variability can be explained by long-term geographic isolation and bottlenecks, exacerbated recently by human encroachment (Driscoll *et al.* 2002).

Apart from these two case studies, inbreeding and its potentially detrimental fitness effects are mainly known for lions in captivity. The occurrence of disturbed behaviour and malformations in cubs, possibly resulting from inbreeding depression, has been witnessed in captive lion populations, but according to our knowledge the prevalence was never assessed (Wensing, pers. comm.). A semi-captive setting is represented by the numerous, small and fenced reserves with intensively managed lion populations, in which changes in genetic diversity may be monitored and demographic parameters can be studied. During the past decades lions have been reintroduced into over 40 small (<1000 km^2), fenced reserves in RSA, currently comprising over 500 individuals (Hayward *et al.* 2007; Trinkel *et al.* 2010; Miller & Funston 2014). The lion population of the small and fenced Hluhluwe-iMfolozi Park, RSA, has originated from five founder individuals in the 1960s (Trinkel *et al.* 2008). Subsequent translocations have been conducted to reinforce the population and counteract inbreeding depression. Litter size and cub survival was higher for pairings including a translocated parent, compared to pairings between native lions (Trinkel *et al.* 2008). One of the two sources for the Hluhluwe/iMfolozi population, is the lion population in the Madikwe Game Reserve. These lions were monitored and results further illustrate how fast inbreeding coefficients rise if a small population is completely closed to gene flow (Trinkel *et al.* 2010). This leads to the notion that continuous supplementation of existing small and isolated populations may be needed and that a meta-population based management plan should be implemented for small reserves with no or reduced gene flow (Trinkel *et al.* 2010; Miller & Funston 2014). This is an important aspect that needs to be taken into account in the current debate about fencing of populations for conservation (Hayward & Kerley 2009; Packer *et al.* 2013a; b; Creel *et al.* 2013; Woodroffe *et al.* 2014).

Above mentioned cases have illustrated how genetic data can contribute to the understanding of the distribution of the genetic diversity in a species, and how this information can be applied in the field of conservation. With new technical and computational developments, genomic data are becoming readily available to conservation geneticists. Recently developed Next Generation Sequencing (NGS) techniques open up possibilities to target new magnitudes of genome-wide genetic markers, also in non-model organisms (Ekblom & Galindo 2011; Gayral *et al.* 2013). The versatile character of such datasets allow to tackle diverse questions of importance for conservation practices, ranging from the identification of management units, to insights into demographic histories (McCormack *et al.* 2013; McMahon *et al.* 2014). The entire genome of a single individual already provides enough data for many kinds of analyses, which may lead to a shift from "one gene, many individuals" to "few individuals, several genes", as is suggested by McMahon *et al.* (2014). A shift from genetics to genomics may therefore open new possibilities in the field of conservation biology.

Research questions and thesis outline

In this thesis the intraspecific genetic diversity of the lion is explored. Current taxonomy, only distinguishing African lion (*P. l. leo*) and Asiatic lion (*P. l. persica*), does not reflect the diversity within the African subspecies. Considering the indications for a unique position of populations in West and Central Africa, the urgent conservation needs and the knowledge gap in this region, this research aims to clarify the position and status of the West and Central Africa lion.

A number of approaches is applied, targeting different genetic markers, to solve the following questions:

1) How is the intraspecific genetic diversity in the lion distributed, and what is the phylogenetic position of the West and Central African lion in particular?
2) How can the phylogeographic pattern of the lion give insight into the evolutionary history of the species, and how does that relate to the phylogeographic patterns of other large African savannah mammals?
3) How does the genetic diversity in West and Central African lions compare with Southern and Eastern African lions, and Asiatic lions? Is there evidence that populations went through bottlenecks?
4) How do different genetic markers with different modes of inheritance and different coalescence times contribute to an insight into the genomic complexity underlying the intraspecific genetic diversity in the lion?
5) Is a revision of lion taxonomy justified and advisable? What are the implications for management of the species, especially with respect to defining ESUs and MUs?

This thesis is structured as follows: Chapter 1 is a general introduction, providing a background on the status of the world's carnivores in general, and the lion in particular. Further it frames the field of phylogeography and conservation genetics. Most relevant studies on lion phylogenetics and population genetics are shortly reviewed. In Chapter 2 data of the two main phylogeographic studies on lions (Dubach *et al.* 2005; Barnett *et al.* 2006b) are combined and further extended by addition of new populations, mainly from West and Central Africa. Since mtDNA data may not represent the true underlying structure in a biogeographic context, nuclear markers were assessed for a number of representative populations in Chapter 3. In addition, genetic diversity measures were calculated to gain insight into the levels of genetic diversity on a population level, and detect traces of bottleneck events. Since microsatellite data, presented in Chapter 3, did not contradict the main genetic clades suggested in Chapter 2, the mtDNA dataset was expanded by more sampling locations to obtain a more fine scale picture of phylogeographic groups. Chapter 4 presents this dataset, including fourteen complete mitochondrial genomes from representative populations and aDNA data, along with newly published data from Barnett *et al*. (2006+2014). The phylogeographic pattern of the lion is compared to a range of large savannah species and its evolutionary history is assessed. Chapter 5 describes the development of a new lion-specific genetic marker by SNP discovery from whole

genome data of ten lions. These data are analysed in a phylogeographic framework and compared to previously described scenarios in Chapter 2, 3 and 4. Chapter 6 is a general discussion, aiming to get a complete overview of available lion data and unraveling the phylogenetic relationships of the different lion clades and their evolutionary histories. It sums up with a look into the future and general recommendations on how to apply these data in the field of lion conservation.

References

Alpers DL, Van Vuuren BJ, Arctander P, Robinson TJ (2004) Population genetics of the roan antelope (Hippotragus equinus) with suggestions for conservation. *Molecular ecology*, **13**, 1771–84.

Antunes A, Troyer JL, Roelke ME *et al.* (2008) The evolutionary dynamics of the lion *Panthera leo* revealed by host and viral population genomics. *PLoS genetics*, **4**, e1000251.

Arbogast B, Kenagy GJ (2001) Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography*, **28**, 819–825.

Armbruster P, Reed DH (2005) Inbreeding depression in benign and stressful environments. *Heredity*, **95**, 235–42.

Avise JC, Ball RM (1990) Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surveys in Evolutionary Biology*, **7**, 45–67.

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). *Proceedings of the Royal Society B: Biological Sciences*, **273**, 2119–2125.

Barnett R, Yamaguchi N, Shapiro B *et al.* (2014) Revealing the maternal demographic history of Panthera leo using ancient DNA and a spatially explicit genealogical analysis. *BMC Evolutionary Biology*, **14**, 70.

Bauer H (2003) Research needs for lion conservation in West and Central Africa. *Comptes Rendus Biologies*, **326**, 112–118.

Bauer H, Iongh HD, Sogbohossou E (2010) Assessment and mitigation of human-lion conflict in West and Central Africa. *Mammalia*, **74**, 363–367.

Bauer H, Van der Merwe S (2004) Inventory of free-ranging lions Panthera leo in Africa. *Oryx*, **38**, 26–31.

Bauer H, Nowell K (2004) Endangered Classification for West African Lions. *Cat News*, **41**, 35–36.

Bauer H, Nowell K, Packer C (2012) IUCN Red List of Threatened Species: Panthera leo. *Version 2013.2*.

Bermingham E, Moritz C (1998) Comparative phylogeography: concepts and applications. *Molecular ecology*, **7**, 367–369.

Beschta RL, Ripple WJ (2009) Large predators and trophic cascades in terrestrial ecosystems of the western United States. *Biological Conservation*, **142**, 2401–2414.

Bijlsma R, Bundgaard J, Boerema AC (2000) Does inbreeding affect the extinction risk of small populations?: predictions from Drosophila. *Journal of Evolutionary Biology*, **13**, 502–514.

Borrvall C, Ebenman B (2006) Early onset of secondary extinctions in ecological communities following the loss of top predators. *Ecology letters*, **9**, 435–42.

Bouché P, Nzapa Mbeti Mange R, Tankalet F *et al.* (2012) Game over! Wildlife collapse in northern Central African Republic. *Environmental monitoring and assessment*, **184**, 7001–11.

Bowen BW (1999) Preserving genes, species, or ecosystems? Healing the fractured foundations of conservation policy. *Molecular ecology*, **8**, S5–10.

Brashares JS, Arcese P, Sam MK (2001) Human demography and reserve size predict wildlife extinction in West Africa. *Proceedings. Biological sciences / The Royal Society*, **268**, 2473–8.

Brashares JS, Prugh PR, Stoner CJ, Epps CW (2010) Trophic Cascades: Predators, Prey, and the Changing Dynamics of Nature. In: *Trophic Cascades: Predators, Prey, and the Changing Dynamics of Nature* (eds Terborgh J, Estes JA), pp. 221–240. Island Press, Washington D.C., U.S.A.

Breitenmoser U, Mallon DP, Ahmad Khan J, Driscoll C (2008) Panthera leo persica. In: *IUCN 2013. IUCN Red List of Threatened Species. Version 2013.2.*

Brown JS, Laundre JW, Gurung M (1999) The ecology of fear: Optimal foraging, game theory, and trophic interactions. *Journal of Mammalogy*, **80**, 385–399.

Bruno JF, Cardinale BJ (2008) Cascading effects of predator richness. *Frontiers in Ecology and the Environment*, **6**, 539–546.

Burger J, Rosendahl W, Loreille O *et al.* (2004) Molecular phylogeny of the extinct cave lion Panthera leo spelaea. *Molecular phylogenetics and evolution*, **30**, 841–849.
Butchart SHM, Walpole M, Collen B *et al.* (2010) Global Biodiversity : Indicators of recent Declines. *Science*, **328**, 1164–1168.
CIESIN, CIAT (2005) Gridded Population of the World Version 3 (GPWv3): Population Density Grids.
Caro TM, Laurenson MK (1994) Ecological and genetic factors in conservation: a cautionary tale. *Science (New York, N.Y.)*, **263**, 485–6.
Caughley G (1994) Directions in conservation biology. *Journal of Animal Ecology*, **63**, 215–244.
Ceballos G, Ehrlich PR (2002) Mammal population losses and the extinction crisis. *Science (New York, N.Y.)*, **296**, 904–7.
Chardonnet P (2002) *Conservation of the African Lion : Contribution to a Status Survey*. International Foundation for the Conservation of Wildlife, France and Conservation Force, U.S.A., Paris.
Craigie ID, Baillie JEM, Balmford A *et al.* (2010) Large mammal population declines in Africa's protected areas. *Biological Conservation*, **143**, 2221–2228.
Crandall KA, Bininda-emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Science*, **15**, 290–295.
Creel S, Becker MS, Durant SM *et al.* (2013) Conserving large populations of lions - the argument for fences has holes. *Ecology letters*.
Crnokrak P, Roff DA (1999) Inbreeding depression in the wild. *Heredity*, **83 (Pt 3)**, 260–70.
Croes BM, Funston PJ, Rasmussen G *et al.* (2011) The impact of trophy hunting on lions (Panthera leo) and other large carnivores in the Bénoué Complex, northern Cameroon. *Biological Conservation*, **144**, 3064–3072.
Crooks KR, Soulé ME (2010) Mesopredator release and avifaunal extinctions in a fragmented system. *Nature*, **400**, 563–566.
Dalerum F (2013) Phylogenetic and functional diversity in large carnivore assemblages. *Proceedings. Biological sciences / The Royal Society*, **280**, 20130049.
Dalerum F, Somers MJ, Kunkel KE, Cameron EZ (2008) The potential for large carnivores to act as biodiversity surrogates in southern Africa. *Biodiversity and Conservation*, **17**, 2939–2949.
Dirzo R, Young HS, Galetti M *et al.* (2014) Defaunation in the Anthropocene. *Science (New York, N.Y.)*, **345**, 401–6.
Dizon AE, Lockyer C, Perrin W, Demaster D, Sisson J (1992) Rethinking the Stock Concept: A Phylogeographic Approach. *Conservation Biology*, **6**, 24–36.
Driscoll CA, Menotti-Raymond M, Nelson G, Goldstein D, O'Brien SJ (2002) Genomic microsatellites as evolutionary chronometers: a test in wild cats. *Genome research*, **12**, 414–423.
Dubach JM, Briggs MB, White PA, Ament BA, Patterson BD (2013) Genetic perspectives on "Lion Conservation Units" in Eastern and Southern Africa. *Conservation Genetics*, **1942**.
Dubach J, Patterson BD, Briggs MB *et al.* (2005) Molecular genetic variation across the southern and eastern geographic ranges of the African lion, Panthera leo. *Conservation Genetics*, **6**, 15–24.
Duriez O, Sachet J-M, Ménoni E *et al.* (2006) Phylogeography of the capercaillie in Eurasia: what is the conservation status in the Pyrenees and Cantabrian Mounts? *Conservation Genetics*, **8**, 513–526.
Ekblom R, Galindo J (2011) Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, **107**, 1–15.
Estes JA, Terborgh J, Brashares JS *et al.* (2011) Trophic downgrading of planet Earth. *Science (New York, N.Y.)*, **333**, 301–6.
Ferreras P, Cousins SH (1996) *The use of a Delphi technique with GIS for estimating the global abundance of top predators*. Cranfield, UK.
Fitzpatrick JL, Evans JP (2009) Reduced heterozygosity impairs sperm quality in endangered mammals. *Biology letters*, **5**, 320–3.
Frankham R (2005) Genetics and extinction. *Biological Conservation*, **126**, 131–140.

Frankham R (2010) Inbreeding in the wild really does matter. *Heredity*, **104**, 124.
Fraser DJ, Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular ecology*, **10**, 2741–2752.
Fretwell SD (1987) Food chain dynamics: The central theory of ecology? *Oikos*, **50**, 291–301.
Gayral P, Melo-Ferreira J, Glémin S *et al.* (2013) Reference-free population genomics from next-generation transcriptome data and the vertebrate-invertebrate gap. *PLoS genetics*, **9**, e1003457.
Gilbert DA, Packer C, Pusey AE, Stephens JC, O'Brien SJ (1990) Analytical DNA fingerprinting in lions: parentage, genetic diversity, and kinship. *The Journal of heredity*, **82**, 378–86.
Guggisberg CAW (1961) *Simba, the life of the lion*. Howard Timmins, Capetown.
De Guia APO, Saitoh T (2007) The gap between the concept and definitions in the Evolutionarily Significant Unit: the need to integrate neutral genetic variation and adaptive variation. *Ecological Research*, **22**, 604–612.
Haas SK, Hayssen V, Krausman PR (2005) Mammalian species - Panthera leo. *Mammalian Species*, 1–11.
Hairston NG, Smith FE, Slobodkin LB (2010) The American Society of Naturalists Community Structure , Population Control , and Competition. *The American Naturalist*, **94**, 421–425.
Hansen BD, Harley DKP, Lindenmayer DB, Taylor AC (2009) Population genetic analysis reveals a long-term decline of a threatened endemic Australian marsupial. *Molecular ecology*, **18**, 3346–3362.
Hayward MW, Adendorff J, O'Brien J *et al.* (2007) Practical Considerations for the Reintroduction of Large, Terrestrial, Mammalian Predators Based on Reintroductions to South Africa's Eastern Cape Province. *The Open Conservation Biology Journal*, **1**, 1–11.
Hayward MW, Kerley GIH (2008) Prey preferences and dietary overlap amongst Africa ' s large predators. *South African Journal of Wildlife Research*, **38**, 93–108.
Hayward MW, Kerley GIH (2009) Fencing for conservation: Restriction of evolutionary potential or a riposte to threatening processes? *Biological Conservation*, **142**, 1–13.
Hemmer H (1974) *Untersuchungen zur Stammesgeschichte der Pantherkatzen (Pantherinae) 3 Zur Artgeschichte des Löwen Panthera (Panthera) leo (Linnaeus 1758).* Veröffentlichungen der Zoologischen Staatssammlung München, München, Germany.
Henschel P, Azani D, Burton C *et al.* (2010) Lion status updates from five range countries in West and Central Africa. *CATnews*, **52**, 34–39.
Henschel P, Coad L, Burton C *et al.* (2014) The Lion in West Africa Is Critically Endangered (M Hayward, Ed,). *PLoS ONE*, **9**, e83500.
Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–13.
Hewitt GM (2004) The structure of biodiversity - insights from molecular phylogeography. *Frontiers in zoology*, **1**, 4.
Hu Y, Guo Y, Qi D *et al.* (2011) Genetic structuring and recent demographic history of red pandas (Ailurus fulgens) inferred from microsatellite and mitochondrial DNA. *Molecular ecology*, **20**, 2662–75.
Hunter LTB, Pretorius K, Carlisle LC *et al.* (2007) Restoring lions Panthera leo to northern KwaZulu-Natal, South Africa: short-term biological and technical success but equivocal long-term conservation. *Oryx*, **41**, 196.
Höglund J, Wang B, Axelsson T, Quintela M (2013) Phylogeography of willow grouse (Lagopus lagopus) in the Arctic: taxonomic discordance as inferred from molecular data. *Biological Journal of the Linnean Society*, **110**, 77–90.
ISIS (2014) Species Holdings - Panthera leo.
IUCN (2012) IUCN Red List Categories and Criteria: Version 3.1. Second edition.
IUCN (2014) The IUCN Red List of Threatened Species. Version 2014.1.
IUCN SSC Cat Specialist Group (2006a) *Conservation strategy for the lion in West and Central Africa*. IUCN, Gland, Switzerland.
IUCN SSC Cat Specialist Group (2006b) *Conservation strategy for the lion (Panthera leo) in Eastern and Southern Africa*. IUCN, Gland, Switzerland.

Johnson CN, Isaac JL, Fisher DO (2007) Rarity of a top predator triggers continent-wide collapse of mammal prey: dingoes and marsupials in Australia. *Proceedings. Biological sciences / The Royal Society*, **274**, 341–6.

Keller LF, Waller DM (2002) Inbreeding effects in wild populations. *Trends in ecology & evolution*, **17**, 19–23.

Kissui BM, Packer C (2004) Top-down population regulation of a top predator: lions in the Ngorongoro Crater. *Proceedings. Biological sciences / The Royal Society*, **271**, 1867–74.

Lacy RC, Petric A, Warneke M (1993) Inbreeding and Outbreeding in Captive Populations of Wild Animals. pdf. In: *The Natural History of Inbreeding and Outbreeding: Theoretical and Empirical Perspectives* (ed Thornhill NW), pp. 352–374. Chicago University Press, Chicago, U.S.A.

Letnic M, Koch F, Gordon C, Crowther MS, Dickman CR (2009) Keystone effects of an alien top-predator stem extinctions of native mammals. *Proceedings. Biological sciences / The Royal Society*, **276**, 3249–56.

Letnic M, Ritchie EG, Dickman CR (2012) Top predators as biodiversity regulators: the dingo Canis lupus dingo as a case study. *Biological reviews of the Cambridge Philosophical Society*, **87**, 390–413.

Lewis ME (1997) Carnivoran paleoguilds of Africa : implications for hominid food. *Journal of Human Evolution*, **32**, 257–288.

Lindsey PA, Balme GA, Funston P *et al.* (2013) The Trophy Hunting of African Lions: Scale, Current Management Practices and Factors Undermining Sustainability (M Hayward, Ed,). *PLoS ONE*, **8**, e73808.

Lorenzen ED, Heller R, Siegismund HR (2012) Comparative phylogeography of African savannah ungulates. *Molecular ecology*, **21**, 3656–70.

Loveridge AJ, Searle AW, Murindagomo F, Macdonald DW (2007) The impact of sport-hunting on the population dynamics of an African lion population in a protected area. *Biological Conservation*, **134**, 548–558.

MacArthur R (2008) Fluctuations of Animal Populations and a Measure of Community Stability. *Ecology*, **36**, 533–536.

Mace GM (2004) The role of taxonomy in species conservation. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **359**, 711–9.

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

McCann KS (2000) The diversity-stability debate. *Nature*, **405**, 228–33.

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular phylogenetics and evolution*, **66**, 526–38.

McMahon BJ, Teeling EC, Höglund J (2014) How and why should we implement genomics into conservation? *Evolutionary Applications*, n/a–n/a.

McNaughton SJ (1977) Diversity and Stability of Ecological Communities : A Comment on the Role of Empiricism in Ecology. *The American Naturalist*, **111**, 515–525.

Millenium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being*.

Miller B, Dugelby B, Foreman D *et al.* (2001) The Importance of Large Carnivores to Healthy Ecosystems. *Endangered Species*, **18**, 202–210.

Miller SM, Funston PJ (2014) Rapid growth rates of lion (Panthera leo) populations in small , fenced reserves in South Africa: a management dilemma. *South African Journal of Wildlife Research*, **44**, 43–55.

Moritz C (1994) Defining "Evolutionarily Significant Units" for conservation. *Trends in ecology & evolution (Personal edition)*, **9**, 373–375.

Moritz C, Faith DP (1998) Comparative phylogeography and the identification of genetically divergent areas for conservation. *Molecular ecology*, **7**, 419–429.

Morrison JC, Sechrest W, Dinerstein E, Wilcove DS, Lamoreux JF (2007) Persistence of Large Mammal Faunas as Indicators of Global Human Impacts. *Journal of Mammalogy*, **88**, 1363–1380.

Newmark WD (2008) Isolation of African protected areas. *Frontiers in Ecology and the Environment*, **6**, 321–328.

Noss R (1990) Indicators for monitoring biodiversity: a hierarchical approach. *Conservation Biology*, **4**, 355–364.

Nowell K, Jackson P (Eds.) (1996a) *Wild Cats, Status Survey and Conservation Action Plan*. IUCN, Gland.

Nowell K, Jackson P (1996b) *Wild Cats*. IUCN, Gland, Switzerland.

Ottichilo WK, Leeuw JD, Skidmore AK *et al.* (2000) Population trends of large non-migratory wild herbivores and livestock in the Masai Mara. *African Journal of Ecology*, **38**, 202–216.

O'Brien SJ, Martenson JS, Packer C *et al.* (1987) Biochemical Genetic Variation in Geographic Isolates of African and Asiatic Lions. *National Geographic Research*, **3**, 114–124.

Packer C, Loveridge a., Canney S *et al.* (2013a) Conserving large carnivores: dollars and fence (N Haddad, Ed,). *Ecology Letters*, n/a–n/a.

Packer C, Pusey AE, Rowley H *et al.* (1991) Case Study of Population Bottleneck: Lions of Ngorongoro. *Biological Conservation*, **5**, 219–230.

Packer C, Swanson a., Canney S *et al.* (2013b) The case for fencing remains intact (J Chase, Ed,). *Ecology Letters*, n/a–n/a.

Paine RT (1966) Food Web Complexity and Species Diversity. *The American Naturalist*, **100**, 65–75.

Paine RT (1969) A Note on Trophic Complexity and Community Stability. *The American Naturalist*, **103**, 91–93.

Patterson BD, Kasiki SM, Selempo E, Kays RW (2004) Livestock predation by lions (Panthera leo) and other carnivores on ranches neighboring Tsavo National ParkS, Kenya. *Biological Conservation*, **119**, 507–516.

Peterson G, Allen CR, Holling CS (1998) Ecological Resilience , Biodiversity , and Scale. *Ecosystems*, **1**, 6–18.

Petter G (1973) *Fossil vertebrates of Africa: Carnivores Pléistocènes du Ravin d'Olduvai* (LSB Leakey, RJG Savage, SC Coryndon, Eds,). Academic Press, London, UK.

Pimm SL (1984) The complexity and stability of ecosystems. *Nature*, **307**, 321–326.

Pimm SL, Jenkins CN, Abell R *et al.* (2014) The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, **344**, 1246752–1246752.

Pimm SL, Raven P (2000) Extinction by numbers. *Nature*, **403**, 843–845.

Pimm SL, Russell GJ, Gittleman JL, Brooks TM (1995) The future of biodiversity. *Science (New York, N.Y.)*, **269**, 347–50.

Ralls K, Ballou J (1986) Captive breeding programs for populations with a small number of founders. *Trends in ecology & evolution*, **1**, 19–22.

Ralls K, Ballou JD, Templeton A (1988) Estimates of Lethal Equivalents and the Cost of Inbreeding in Mammals. *Conservation Biology*, **2**, 185–193.

Randi E, Tabarroni C, Rimondi S, Lucchini V, Sfougaris a. (2003) Phylogeography of the rock partridge (Alectoris graeca). *Molecular Ecology*, **12**, 2201–2214.

Ray JC, Hunter L, Zigouris J (2005a) *Setting conservation and research priorities for larger African carnivores*. New York, U.S.A.

Ray J, Redford . H, Steneck R, Berger J (2005b) *Large Carnivores and the Conservation of Biodiversity*. Island Press, Washington D.C., U.S.A.

Reed DH, Frankham R (2003) Correlation between Fitness and Genetic Diversity. *Conservation Biology*, **17**, 230–237.

Riggio J, Jacobson A, Dollar L *et al.* (2012) The size of savannah Africa: a lion's (Panthera leo) view. *Biodiversity and Conservation*, **22**, 17–35.

Ripple WJ, Estes JA, Beschta RL *et al.* (2014) Status and Ecological Effects of the World's Largest Carnivores. *Science*, **343**, 1241484–1241484.

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. *Trends in ecology & evolution*, **1**, 9–10.

Saccheri I, Kuussaari M, Kankare M, Vikman P, Hanski I (1998) Inbreeding and extinction in a butterfly metapopulation. *Nature*, **45**, 1996–1999.

Sala E (2006) Top predators provide insurance against climate change. *Trends in ecology & evolution*, **21**, 479–80.

Sanderson EW, Redford KH, Chetkiewicz C-LB *et al.* (2002) Planning to Save a Species: the Jaguar as a Model. *Conservation Biology*, **16**, 58–72.

Schaller GB (1972) *The Serengeti Lion - A study of predator-prey relations*. University of Chicago Press, Chicago, U.S.A.

Schmitz OJ, Hawlena D, Trussell GC (2010) Predator control of ecosystem nutrient dynamics. *Ecology letters*, **13**, 1199–209.

Schnitzler AE (2011) Past and present distribution of the North African-Asian lion subgroup: a review. *Mammal Review*.

Sinclair ARE, Mduma S, Brashares JS (2003) Patterns of predation in a diverse predator – prey system. *Nature*, **425**, 288–290.

Sogbohossou EA, de Iongh HH, Sinsin B, de Snoo GR, Funston PJ (2011) Human–carnivore conflict around Pendjari Biosphere Reserve, northern Benin. *Oryx*, **45**, 569–578.

Spielman D, Brook BW, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. *PNAS*, **101**, 15261–15264.

Strong DR, Frank KT (2010) Human Involvement in Food Webs. *Annual Review of Environment and Resources*, **35**, 1–23.

Terborgh J, Lopez L, Nuñez P *et al.* (2001) Ecological meltdown in predator-free forest fragments. *Science (New York, N.Y.)*, **294**, 1923–6.

Tilman D, Downing JA (1994) Biodiversity and stability in grasslands. *Nature*, **367**, 363–365.

Treves A, Karanth KU (2003) Human-Carnivore Conflict and Perspectives on Carnivore Management Worldwide. *Conservation Biology*, **17**, 1491–1499.

Trinkel M, Cooper D, Packer C, Slotow R (2011) Inbreeding depression increases susceptibility to bovine tuberculosis in lions: an experimental test using an inbred-outbred contrast through translocation. *Journal of wildlife diseases*, **47**, 494–500.

Trinkel M, Ferguson N, Reid a. *et al.* (2008) Translocating lions into an inbred lion population in the Hluhluwe-iMfolozi Park, South Africa. *Animal Conservation*, **11**, 138–143.

Trinkel M, Funston P, Hofmeyr M *et al.* (2010) Inbreeding and density-dependent population growth in a small, isolated lion population. *Animal Conservation*, **13**, 374–382.

Tumenta PN, Visser HD, van Rijssel J *et al.* (2013) Lion predation on livestock and native wildlife in Waza National Park, northern Cameroon. *Mammalia*, **77**, 247–251.

Turner A, Antón M (1997) *The big cats and their fossil relatives*. Columbia University Press, New York.

UNEP-CBD (2006) Convention on Biological Diversity, Framework for monitoring implementation of the achievement of the 2010 target and integration of targets into the thematic programmes of work, COP8 Decision VIII/15.

Visser ME (2008) Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings. Biological sciences / The Royal Society*, **275**, 649–659.

Vogler AP, DeSalle R (1994) Diagnosing Units of Conservation Managment. *Conservation Biology*, **6**, 170–178.

Waples RS (1991) Pacific Salmon, Oncorhynchus spp ., and the Definition of "Species" Under the Endangered Species Act. *Marine Fisheries Reviews*, **53**, 11–22.

Werdelin L, Yamaguchi N, Johnson WE, O'Brien SJ (2010) Biology and Conservation of Wild Felids. In: *Biology and Conservation of Wild Felids* (eds Macdonald DW, Loveridge AJ), pp. 59–82. Oxford University Press, Oxford.

Wildt DE, Bush M, Goodrowe KL *et al.* (1987) Reproductive and genetic consequences of founding isolated lion populations. *Nature*, **329**, 328–331.

Winterbach HEK, Winterbach CW, Somers MJ, Hayward MW (2012) Key factors and related principles in the conservation of large African carnivores. *Mammal Review*, **43**, 89–110.

Woodroffe R (1998) Edge Effects and the Extinction of Populations Inside Protected Areas. *Science*, **280**, 2126–2128.

Woodroffe R (2000) Predators and people: using human densities to interpret declines of large carnivores. *Animal Conservation*, **3**, 165–173.

Woodroffe R, Hedges S, Durant SM (2014) To Fence or Not to Fence. , **344**, 46–48.

Woodroffe R, Thirgood S, Rabinowitz A (2005) *People and Wildlife, Conflict or Co-existence*. Cambridge University Press, Cambridge, U.K.

Wright S (1977) *Evolution and the Genetics of Populations: Volume 3 - Experimental Results and Evolutionary Deductions*. Chicago University Press, Chicago, U.S.A.

Yamaguchi N, Cooper A, Werdelin L, Macdonald DW (2004) Evolution of the mane and group-living in the lion (Panthera leo): a review. *Journal of Zoology*, **263**, 329–342.

Yirga G, Iongh HHD, Leirs H *et al.* (2012) The ecology of large carnivores in the highlands of northern Ethiopia. *African Journal of Ecology*, **51**, 78–86.

Yuhki N, O'Brien SJ (1990) DNA variation of the mammalian major histocompatibility complex reflects genomic diversity and population history. *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 836–840.

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.

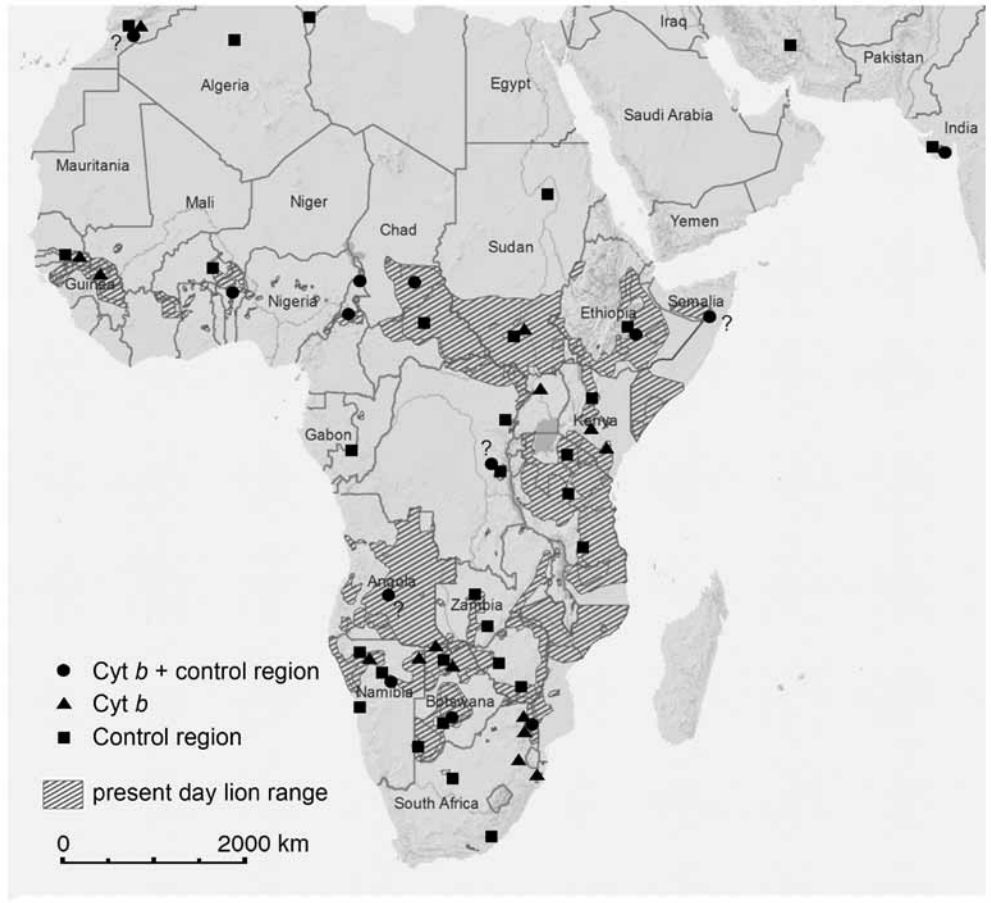


Figure 1. Map showing the origin of the lion (*Panthera leo*) samples that were used for the phylogenetic analyses. Dots indicate the samples from which cytochrome *b* (cyt *b*), tRNAThr, tRNAPro and the left domain of the control region sequences are known, triangles are cytochrome *b* sequences, and 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.

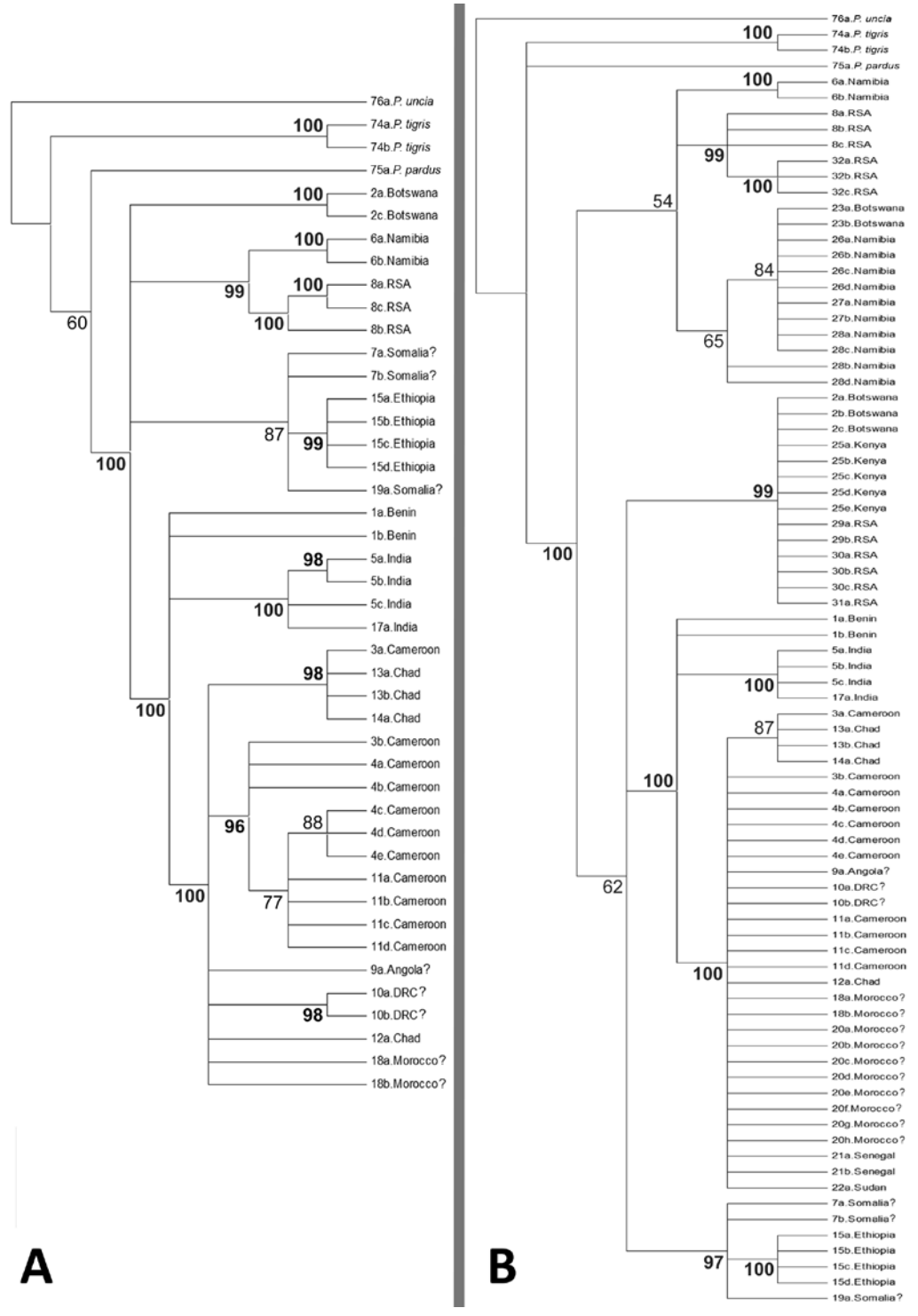


Figure 2. Phylogenetic trees resulting from Bayesian analysis of two sets of lion (*Panthera leo*) sequences: (A) cytochrome *b* + control region, (B) cytochrome *b* alone. The numbers represent the percentages for Bayesian posterior probability (PP). DRC = Democratic Republic of the Congo, RSA = Republic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

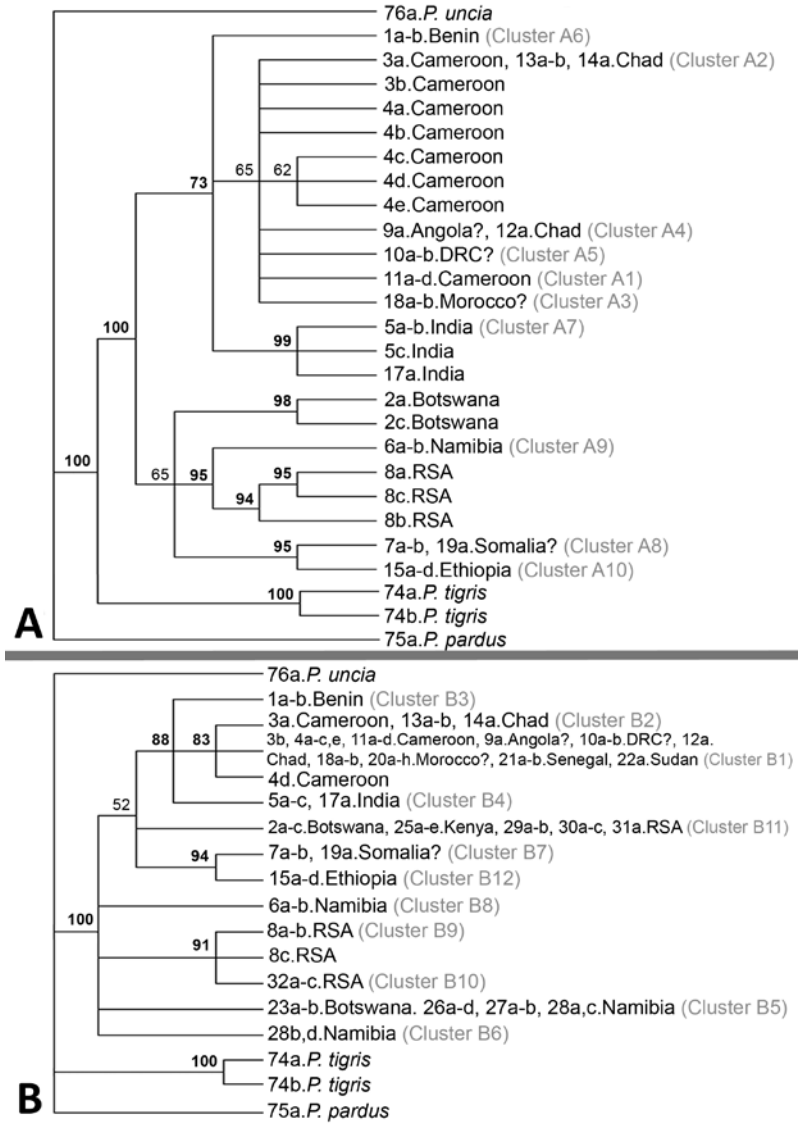


Figure 3. Phylogenetic trees resulting from maximum likelihood (ML) analysis of two sets of lion (*Panthera leo*) sequences: (A) cytochrome *b* + control region, and (B) cytochrome *b* alone. The numbers indicate the percentage for bootstrap support. Identical sequences were pooled. These clusters correspond 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, $P < 0.0001$).

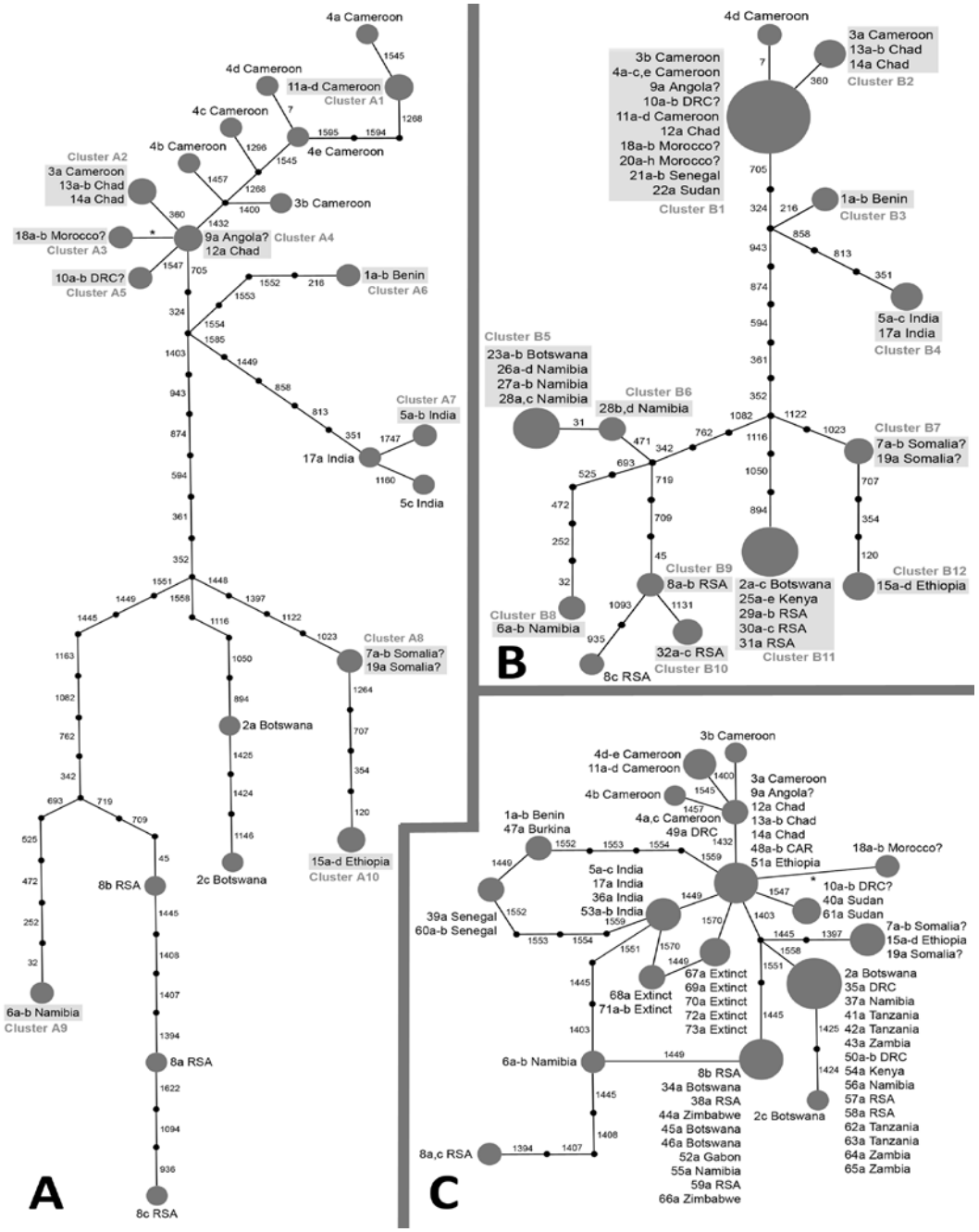


Figure 4. Haplotype networks for each of the analysed sets of lion (*Panthera leo*) sequences: (A) cytochrome *b* + control region, (B) cytochrome *b* alone, and (C) control region alone. The numbers indicate the location of each mutation. In (a) and (c), one of these mutations is indicated by *, representing 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.

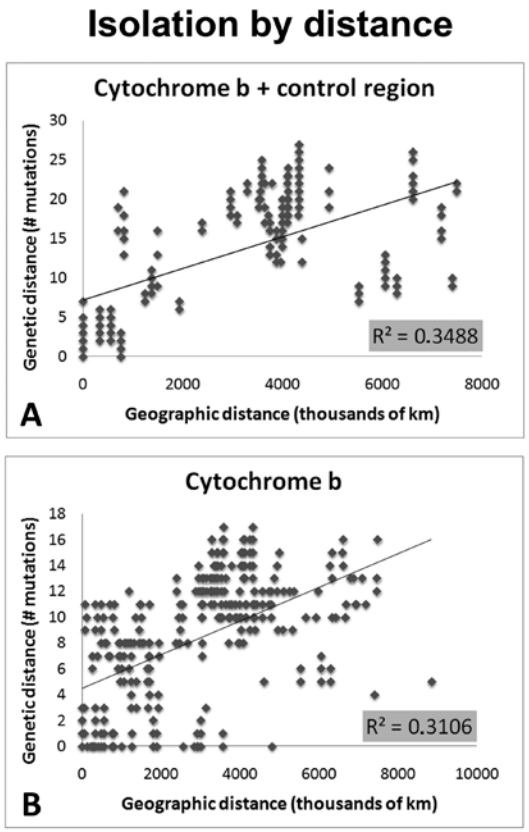
Irwin, D.M., Kocher, T.D. & Wilson, A.C. (1991) Evolution of the cytochrome *b* gene of mammals. *Journal of Molecular Evolution*, **32**, 128-144.

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.

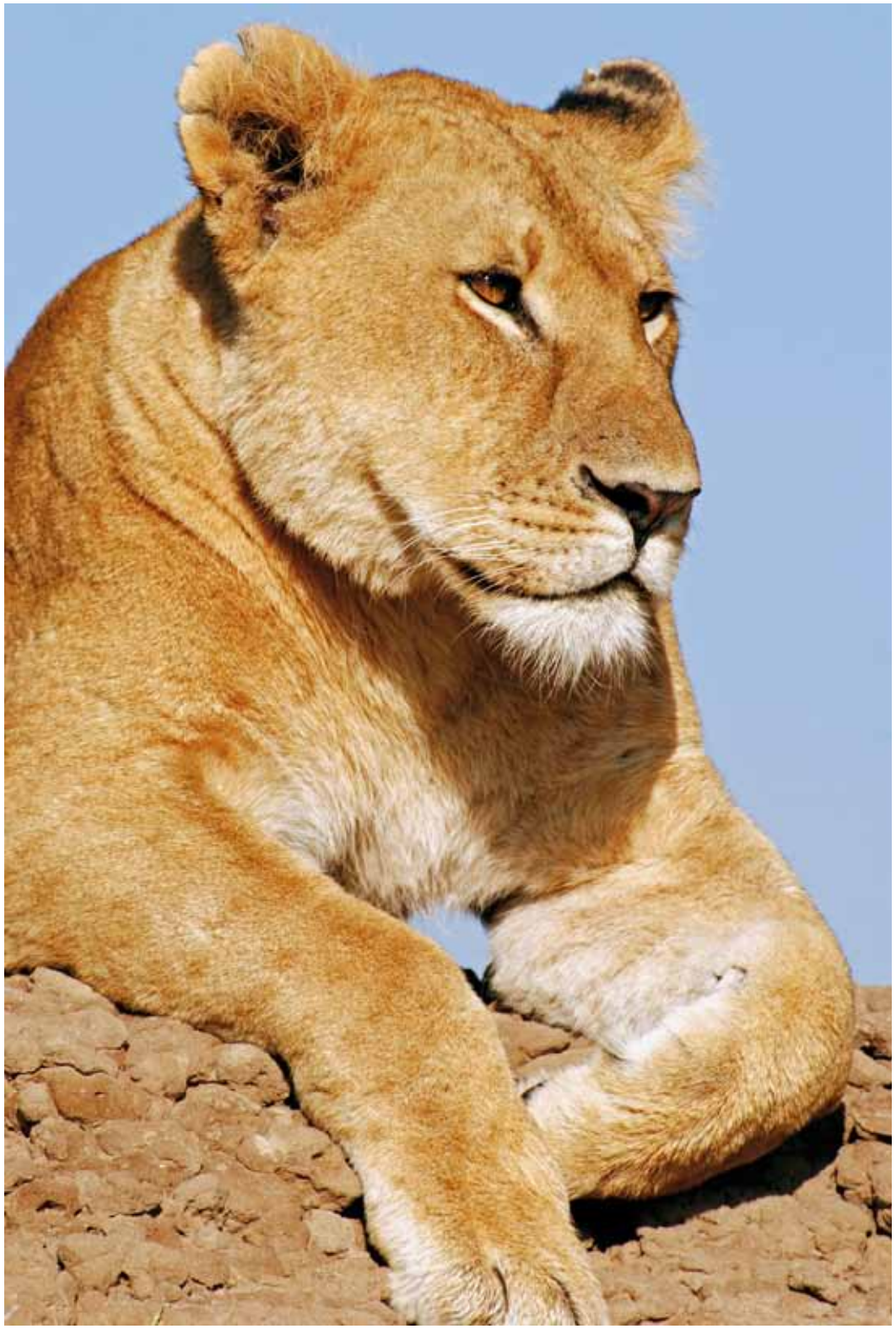
Kim, J.H., Antunes, A., Luo, S.J., Menninger, J., Nash, W.G., O'Brien, S.J. & Johnson, W.E. (2006) Evolutionary analysis of a large mtDNA translocation (*numt*) into the nuclear genome of the *Panthera* genus species. *Gene,* **366**, 292-302.

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.



Chapter 3

Autosomal and mtDNA markers reveal concordant phylogenetic patterns of lion populations over the entire geographic range

(under review)

L. D. Bertola, L. Tensen, P. van Hooft, P. A. White, C. A. Driscoll, P. Henschel, A. Caragiulo, I. Dias-Freedman, E. A. Sogbohossou, P. N. Tumenta, T. H. Jirmo, G. R. de Snoo, H. H. de Iongh and K. Vrieling

Abstract

The evolutionary history of a species is key for understanding the taxonomy and for the design of effective management strategies for species conservation. The knowledge about the evolutionary history of the lion (*Panthera leo*) is largely based on mitochondrial markers. Here, we investigate whether autosomal markers are concordant with previously described phylogeographic patterns. Special emphasis is placed on the lion in West/Central Africa, as previous studies using only mitochondrial markers have shown this region to hold a distinct evolutionary lineage. In addition, anthropogenic factors have led to a strong decline in West/Central African lion numbers in recent history, thus, the conservation value of these populations is particularly high. Analysis of 20 microsatellites and 1,454 bp of the mitochondrial DNA in 16 lion populations representing the entire geographic range of the species, found congruence in both types of markers, identifying four clusters: 1) West/Central Africa, 2) East Africa, 3) Southern Africa and 4) the Asiatic subspecies. This is not in line with the current taxonomy, which only recognizes an African and an Asiatic subspecies. There are no indications that genetic diversity in West/Central Africa lions is lower than in either East or Southern Africa, however, given this genetic distinction and the recent declines of lion numbers in this region, we strongly recommend prioritization of conservation projects in West/Central Africa. As the current taxonomic nomenclature does not reflect the evolutionary history of the lion, we suggest that a taxonomic revision of the lion is warranted.

Keywords: African climate history, lion (*Panthera leo*), mitochondrial genome, phylogeography, savannah mammals, West and Central Africa

Introduction

Identifying and describing patterns of mitochondrial (mtDNA) and nuclear genetic variation is a crucial component to fully understanding the evolutionary history of a species. High quality phylogeographic data that represent the underlying genetic complexity are important for taxonomy and contribute to designing effective conservation strategies. This is of particular importance for species such as the lion (*Panthera leo*) that occupy large geographic ranges within which disjunct populations may not allow for natural dispersal and gene flow. Increasing habitat fragmentation and variable anthropogenic factors have created a growing need to manage lions at the population level (Riggio *et al.* 2012). In addition, several recent publications have sparked the discussion whether the current taxonomic nomenclature for the lion is justified (Bertola *et al.* 2011a; Dubach *et al.* 2013; Barnett *et al.* 2014).

Two subspecies of lion are currently recognized by the IUCN: the African lion (*Panthera leo leo*), ranging throughout sub-Saharan Africa with the exception of dense rain forest, and the Asiatic lion (*Panthera leo persica*), which exists as a single population in the Gir forest, India. Although all African lion populations are considered as belonging to the African subspecies, distinct genetic subgroups have been recognized (Dubach *et al.* 2005, 2013; Barnett *et al.* 2006a; b, 2014; Antunes *et al.* 2008; Bertola *et al.* 2011a; Bruche *et al.* 2012). Based on phylogenetic analysis of mitochondrial haplotypes only, lions in West/Central Africa were described as a genetically distinct group (Bertola *et al.* 2011a; Dubach *et al.* 2013; Barnett *et al.* 2014) (region definitions from IUCN SSC Cat Specialist Group 2006a; b, see Figure 1). The genetic dichotomy that separates the West/Central African lion populations from East and Southern African populations, has also been found in other large mammal species and is often reflected in their taxonomy, including African buffalo (*Syncerus caffer*) (Van Hooft *et al.* 2002; Smitz *et al.* 2013), roan antelope (*Hippotragus equinus*) (Alpers *et al.* 2004), hartebeest (*Alcelaphus buselaphus*) (Arctander *et al.* 1999; Flagstad *et al.* 2001), giraffe (*Giraffa camelopardalis*) (Brown *et al.* 2007; Hassanin *et al.* 2007) and cheetah (*Acinonyx jubatus*) (Freeman *et al.* 2001; Charruau *et al.* 2011). However, mtDNA is a single, non recombining locus in the maternal lineage and does not permit the detection of admixture events and sorting at multiple loci, as may occur in autosomal markers. Therefore, the observed pattern in mtDNA data may not adequately depict the underlying genetic complexity.

Because the Asiatic lion subspecies occupies a nested position in the mtDNA based phylogenetic tree within the West/Central Africa clade, the current taxonomic division is challenged (Bertola *et al.* 2011a; Dubach *et al.* 2013; Barnett *et al.* 2014). The dichotomy within Africa has previously been described based on a range of morphometric data (Hemmer 1974) and the close relationship of West/Central African populations to the Asiatic subspecies is further supported by craniometric data (Mazák 2010). However, autosomal data are needed to assess how well the mtDNA tree represents the phylogeographic complexity in the lion, since conflicting patterns between phylogenies based on mtDNA and phylogenies based on autosomal markers have been described in several other species (Sota & Vogler 2001; DeBruyne 2005; Pinho *et al.* 2007; Roca *et al.* 2007; Zink & Barrowclough 2008; Rato *et al.* 2010). Most commonly a monophyletic pattern is detected in the mtDNA, but is not supported, or is contradicted, by phylogenies based on autosomal loci. This is often explained by incomplete lineage sorting, as coalescence time in mtDNA is four times shorter than in autosomal markers. Since lineage sorting during the process of coalescence has a random nature, this could also lead to an 'incorrect' gene tree by mtDNA markers if populations divergences were closely spaced in time. Female philopatry is another strong contributing factor in mtDNA trees. As gene flow in lions is biased towards the male sex (Pusey *et al.* 1987; Spong *et al.* 2002), gene trees based on autosomal markers may show less discrete groups. This argument has been used by Antunes *et al.* (2008) to explain incongruent patterns in their lion data based on mtDNA and autosomal markers. Taxonomic revisions have potentially far-reaching ramifications with regard to management (e.g., CITES, USFWS, IUCN), and therefore, should be approached cautiously. Ideally proposed revisions should be supported by a combination of biogeographic, mtDNA and autosomal DNA, and morphological data.

In this study, we analyzed 20 microsatellite loci for lions from thirteen wild populations, one of which is located in West Africa (Benin) and four in Central Africa (Chad, DRC and two from Cameroon). Furthermore, we included microsatellite data from another West African population in Senegal, and from two distinct zoo populations of Ethiopian lions. To compare the phylogenetic clusters derived from the microsatellite data and to check for congruence with previously published patterns, we included data from 1,454 base pairs (bp) of the mitochondrial DNA for each sampling location. Using this approach we are aiming to contribute in the ongoing discussion about lion taxonomy, by answering four questions: 1) Do autosomal data support previously described phylogenetic groupings in the lion in general and the distinct position of the West/Central African lion in particular? 2) Can an effect of sex-biased gene flow be detected? 3) How genetically distinct are the sampled populations, at both the continental and regional scales, and how do levels of genetic diversity compare amongst regional subdivisions? 4) Are there signs for reduced genetic diversity in particular lion populations with an emphasis on West/Central Africa? Our study is the first to include multiple lion populations from West/Central Africa, using both autosomal and mtDNA markers in a phylogenetic context covering the entire current geographic range of the lion.

Materials and Methods

We processed a total of 48 samples from eight populations, including one population from West Africa (Benin), four populations from Central Africa (two from Cameroon, one from Chad and one from DRC), two populations from East Africa (Ethiopia2 (captive) and Kenya) and one population from Southern Africa (Zambia). Except for Ethiopia2, all included samples originated from free-ranging lions, with no known history of anthropogenic introductions of lions from other populations. Samples were collected in full compliance with specific permits (CITES and permits related to national legislation in the countries of origin). Details on permits, sample storage, DNA extraction, polymerase chain reaction (PCR) amplification, fragment analysis and quality control are given in Supplemental Information S1. See Supplemental Table S1 and S2 for used loci and primer information. All microsatellite allele length data are given in Supplemental Table S3.

Generated microsatellite data were supplemented by published data for the same 20 loci from another six populations (Driscoll *et al.* 2002), together summarized as Dataset 1. Dataset 2 consists of all 15 samples from Ethiopia1 (captive) with ten analyzed loci (Bruche *et al.* 2012), of which six are overlapping with our dataset. For two samples from Ethiopia1, all 20 microsatellites were analyzed and added to Dataset 1. Dataset 3 (Panthera/AMNH) contains microsatellite data from 12 loci for

seven lions from Senegal, which could not be resized to Dataset 1 and were therefore only included for calculation of diversity indices and bottleneck statistics (for details on permits and the processing of Senegal samples, see Supplemental Information S2). An overview of datasets used in each analysis is provided in Figure 1 and Table 1.

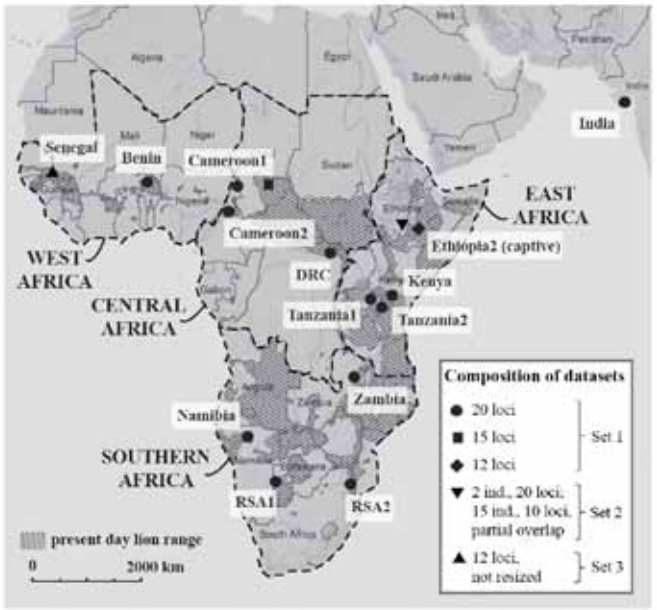


Figure 1. Map showing the location of the 16 lion populations included in the analysis.
In the legend the composition of the datasets and the number of included microsatellite loci is indicated. Lion range data from IUCN (2014). Region definitions from IUCN SSC Cat Specialist Group (2006a; b).

Table 1. Overview of included lion populations in this study. PopSize: population size according to the most recent estimate in Riggio *et al.* (2012) for the African populations, except for Zambia: Paula White (personal communication); estimate for the Indian population from (Singh & Gibson 2011); N msat: number of sampled individuals for microsatellite analysis; N mtDNA: number of sampled individuals for mtDNA analysis.

Set	Population	Area	Geographic Region	PopSize	N msat	N mtDNA	Source msat data
1	**Benin**	Pendjari NP	West Africa	100	5	5	this dataset
	Cameroon1	Waza NP	Central Africa	20	9	9	this dataset
	Cameroon2	Bénoué Ecosystem	Central Africa	200	3	3	this dataset
	Chad	Zakouma NP	Central Africa	140	4	4	this dataset
	DRC	Garamba NP	Central Africa	175	7	6	this dataset
	Ethiopia2	Yemen Zoo	East Africa	(captive)	4	4	this dataset
	Kenya	Amboseli NP	East Africa	60	7	7	this dataset
	Tanzania1	Serengeti NP	East Africa	3465	10	3	Driscoll et al., 2002
	Tanzania2	Ngorongoro CA	East Africa	53	10	1	Driscoll et al., 2002
	Zambia	Luangwa Valley	Southern Africa	750	9	9	Driscoll et al., 2002
	Namibia	Etosha NP	Southern Africa	455	10	2	Driscoll et al., 2002
	RSA1	Kalahari-Gemsbok NP	Southern Africa	350	10	2	Driscoll et al., 2002
	RSA2	Kruger NP	Southern Africa	1684	10	10*	Driscoll et al., 2002
	India	Gir forest NP	India	411	10	6	Driscoll et al., 2002
2	**Ethiopia1**	Addis Ababa Zoo	East Africa	(captive)	15	5	Bruche et al., 2012
3	**Senegal**	Niokolo Koba NP	West Africa	15	7	7	Panthera/AMNH

* mtDNA and microsatellite data are not from the same samples

STRUCTURE 2.3.3 (Pritchard *et al.* 2000) was used for assessing population structure in Dataset 1 with unknown loci scored as missing data. Simulations were run assuming the admixture model with correlated allele frequencies. Ten runs were performed for K=1 to K=11, using 1,000,000 permutations and a burn-in period of 100,000. To check the assignment of Ethiopia1 to any of the clusters identified by STRUCTURE, we included the two Ethiopian samples for all 20 microsatellites. The true value of K was determined using Structure Harvester (Evanno *et al.* 2005). CLUMPP was used to combine replicate runs and avoid label switching (Jakobsson & Rosenberg 2007). Clustering of individuals was further assessed by performing Principal Component Analysis (PCA) in GenAlEx 6.501 (Peakall & Smouse 2012). A neighbour-joining tree was created based on D_A distance in POPTREE2 using 1,000 bootstraps (Takezaki *et al.* 2010).

For each sampling location, a mitochondrial region of 1,454 bp that encompassed cytochrome b (cytB), tRNAThr, tRNAPro and part of the control region was included for a number of individuals (Table 1). Details on polymerase chain reaction (PCR) amplification and sequencing are given in Supplemental Information S1. Sequences were deposited in GenBank and supplemented by sequences previously published by Bertola *et al.* (2011) (see Supplemental Table S4 for accession numbers). Variable sites and nucleotide diversity were calculated using ARLEQUIN 3.5 (Excoffier *et al.* 2005). For phylogenetic analysis, a haplotype network was created using the median-joining algorithm in Network 4.6.1.1 (www.fluxus-engineering.com). A repeat region of cytosines of variable length was excluded due to unknown homology (bp 1382-1393) and all remaining characters were included with equal weighting.

For AMOVA of Dataset 1, individuals for which all 20 loci were analyzed were included as either 1) without an indicated substructure (as all 1 group), 2) following IUCN classification (Africa; Asia), 3) following a North/South division as was indicated from the haplotype network or 4) using the four groups identified by STRUCTURE (West/Central Africa; East Africa; Southern Africa; India). Isolation By Distance (IBD) was assessed by correlating geographic to genetic distances and using a Mantel's permutation test with 999 permutations, as implemented in GenAlEx 6.501 (Peakall & Smouse 2012). In addition, AMOVA and IBD analysis were performed on a regional level, using the regions as indicated above (Africa; North; South; West/Central Africa; East Africa; Southern Africa). Pairwise Fst and Nei's genetic distances were computed with GenAlEx 6.501 (Peakall & Smouse 2012) for microsatellite data and with ARLEEQUIN 3.5 for mtDNA data (Excoffier *et al.* 2005).

The average number of alleles per locus (Na) was calculated using ARLEQUIN 3.5 (Excoffier *et al.* 2005). Private allelic richness (NaPr) was calculated with HP-Rare 1.1 (Kalinowski 2005) including statistical rarefaction to compensate for different sample sizes. GenAlEx 6.501 (Peakall & Smouse 2012) was used to calculate observed (Ho) and unbiased expected heterozygosity (uHe) (Nei 1978). Fis per population was calculated in FSTAT (Goudet 2001). The occurrence of recent bottlenecks was evaluated by testing for heterozygosity excess using the program Bottleneck (Cornuet & Luikart 1996; Piry *et al.* 1999). The program was run for 10,000 iterations, using the stepwise mutation model (SMM). Significant (<0.05) results from the Wilcoxon signed-rank test were scored, as this test proved to be the most powerful and robust when used with few (<20) polymorphic loci (Piry *et al.* 1999).

Results

Based on the STRUCTURE results of Dataset 1, Structure Harvester identified that the observed genetic structure is best described by four clusters representing the following geographic areas: 1) West/Central Africa, 2) East Africa, 3) Southern Africa and 4) India (Figure 2). Individuals from Chad are part of the West/Central Africa cluster. The Ethiopian lions show affiliation either to West/Central Africa, admixed with Southern Africa (Ethiopia1) or to East Africa, admixed with Southern Africa (Ethiopia2). The Zambia population shows a substructure as a result of admixture. All Zambian individuals are partially assigned to the Southern Africa cluster, and depending on the individual, either to West/Central Africa, or to East Africa. The admixed signal of the Zambia population is also visible by the central position in the plot of the first two axes of the PCA when India is excluded (Figure 3B). Since PCA illustrated the effect of India (Figure 3A) and it is known that STRUCTURE may find fewer clusters than is expected based on known evolutionary history when one cluster is strongly deviating (Kalinowski 2010), STRUCTURE runs were repeated excluding Indian genotypes. These analyses did not lead to a difference in clustering of the remaining individuals and the same three groups were identified within Africa (data not shown).

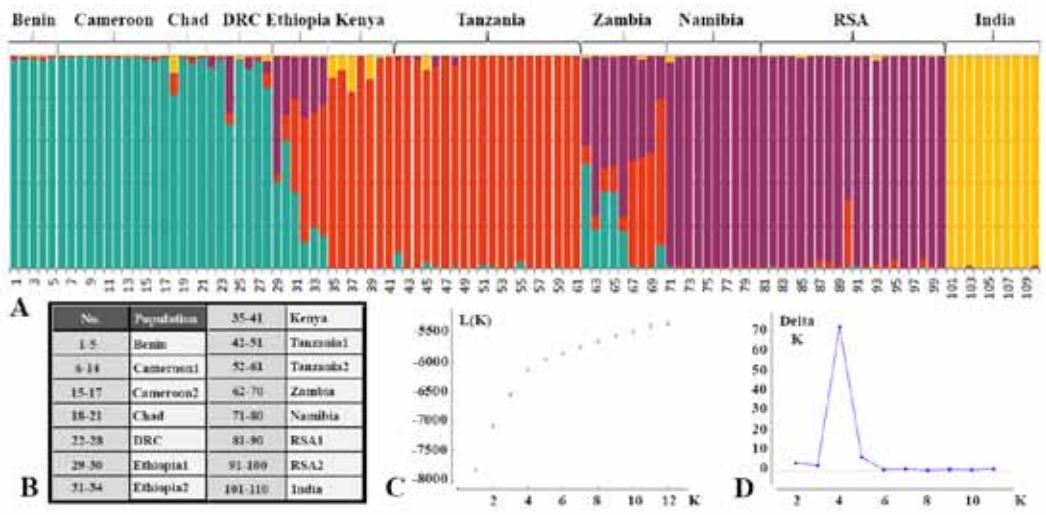


Figure 2. Results of STRUCTURE analysis, based on 20 microsatellite loci of 15 lion populations (Dataset 1 + 2 individuals from Ethiopia1). A: representation of assignment values found by STRUCTURE, using K=4; B: overview of included populations; C: plot indicating mean log likelihood Ln (P(X|K); D: plot indicating DeltaK values as a function of the number of genetic clusters (K), in which DeltaK=mean(|L''(K)|)/sd(L(K)).

A total of 87 sequences of 1,454 bp were analyzed. Nucleotide diversity (π) was 0.102. Based on 43 polymorphic sites, 15 different haplotypes were distinguished. The haplotype network (Figure 4A) and the neighbour-joining tree (Figure 4B) based on the microsatellite data show a similar topology, in which West/Central African lions are grouped together on a significantly supported branch (bootstrap value >70) and East and Southern African lions are represented on two different significantly supported branches (Figure 4B). A basal split into a North group (West/Central Africa and India) and a South group (East Africa and Southern Africa) is most clearly visible in the haplotype network, as the clustering of East Africa and Southern Africa on a South branch in the phenetic tree has only limited support. Furthermore, Kenya and India both have a basal and unresolved position in the tree.

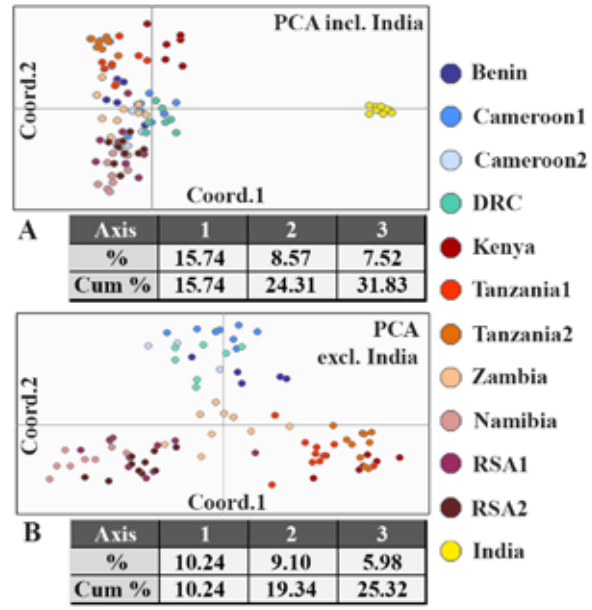


Figure 3. Results of PCA, based on 20 microsatellite loci of lion populations.
A: results PCA of 12 populations (Dataset 1, excluding Chad and Ethiopia2), shown in a two-dimensional plot and a table indicating the percentage and the cumulative percentage explained by the first three axes; B: results of PCA of 11 populations, excluding India.

Results from AMOVA of the microsatellite data show that following the clusters identified by STRUCTURE to assign populations to four groups, resulted in a relative high percentage of the molecular variance being attributed to among-groups for microsatellite data (17.4%) and mtDNA data (52.3%) (Supplemental Table S5). While in the microsatellite data the highest percentage (29.6%) of molecular variance in among-groups variance is attributed to the split between Africa and Asia, i.e. between the two subspecies, no molecular variance among-groups for the Africa/Asia division is found in the mtDNA data. In addition, following the basal split in a North group and a South group, AMOVA attributes 54.6% of molecular variance to among-groups variance for mtDNA data, but only finds 7.5% in among-groups variance when using microsatellite data. Absolute percentages may be misleading, as within-population variance is very different amongst the used markers.

Mantel tests showed that the effect of isolation by distance is evident, both on the continental and the regional scale (summary and graphs in Supplemental Table S6). In regional analyses, the highest values for among-groups variance according to AMOVA and the highest numbers for the slope of the trend line in IBD are found in West/Central Africa (compared to the South group, East Africa or Southern Africa) suggesting strong isolation between these populations. Pairwise Fst values ranged from 0.064 to 0.736 and were significant for all pairwise comparisons (50,000 permutations, $P<0.05$) (Supplemental Table S7). Within Africa pairwise Fst values ranged from 0.064 to 0.396. Nei's genetic distance ranged from 0.196 to 2.193 for all lion populations and within Africa it ranged from 0.196 to 2.018 (Supplemental Table S7).

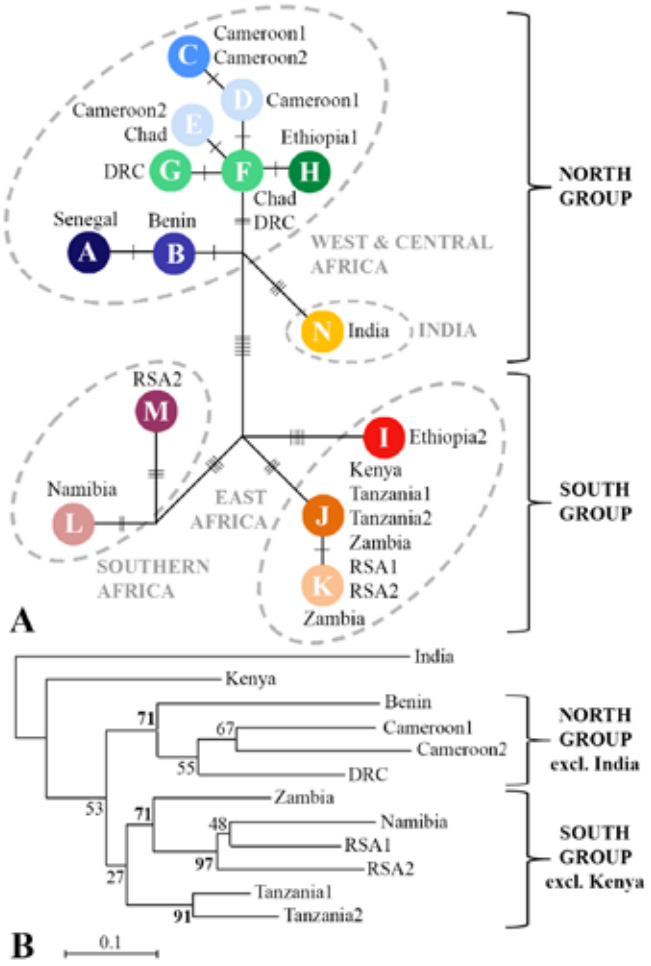


Figure 4. Relationship between populations of lions based on mtDNA data and on 20 microsatellite loci. A: Haplotype network based on median-joining algorithm in Network; B: Phenetic tree based on Da genetic distance of microsatellite data of 12 lion populations.

Diversity indices (Supplemental Table S8) show that the Indian population comprises the lowest number of microsatellite alleles per locus, smallest allelic range and the highest number of fixed alleles. In the Indian population 75% of the loci are fixed while in all other populations at maximum 17% of the loci are fixed. Diversity indices were found to be relatively constant across the African populations; surprisingly West/Central Africa showed no clear signs of loss of genetic diversity. Four out of seven populations in West/Central Africa contained more than one haplotype (Cameroon1, Cameroon2, Chad, DRC), whereas this was only observed for two out of eight populations in East and Southern Africa (Zambia and RSA2). Observed and expected heterozygosity values further confirmed the low genetic diversity of the Indian population. Fis values illustrated a significant heterozygosity excess in Benin ($P<0.01$) and Cameroon1 ($P<0.01$) and a significant heterozygosity deficiency in Zambia ($P<0.01$), RSA1 ($P<0.05$) and Ethiopia1 ($P<0.05$). Results of the bottleneck analysis showed that there was a significant excess of heterozygotes found in Cameroon1 ($P<0.01$), Kenya ($P<0.05$) and Ethiopia1 ($P<0.05$), possibly indicating a recent reduction in population size.

Discussion

Here we present an analysis of microsatellite and mtDNA datasets in lions sampled across their current geographic range. We included autosomal markers because this method had not been previously applied to investigate the genetic dichotomy between lion populations in West/Central Africa and those in East and Southern Africa. Moreover, we assessed levels of genetic diversity across different geographic scales to detect signs of low genetic diversity.

Analysis of microsatellite data (STRUCTURE) identified three clusters in the African lion: 1) West/Central Africa, 2) East Africa, and 3) Southern Africa, in addition to a cluster comprising the Asiatic subspecies. Although the high level of fixation of alleles in the Asiatic lion is likely to contribute to the identification of this population as a distinct cluster, genetic structure is found within the African subspecies. This supports the genetically distinct position of lions from West/Central Africa reported previously (Barnett *et al.* 2006b, 2014; Bertola *et al.* 2011a) and found again here based on mtDNA data. In addition, STRUCTURE also indicates divergence within the East and Southern African lions. The observed split between East and Southern Africa, as was previously shown by Bruche *et al.* (2012), remained after inclusion of a population from Zambia, geographically intermediate between Tanzania and RSA. Bruche *et al.* (2012) included lions from the Ethiopia1 population in a STRUCTURE analysis with data from Driscoll *et al.* (2002) and found a distinction between India, East Africa, Southern Africa and Ethiopia1. From this the authors concluded that the Ethiopia1 individuals form a unique group within the African lion. Including two individuals of Ethiopia1 in our microsatellite dataset, we find strong admixture with West/Central Africa. This is further confirmed by the mitochondrial haplotype of these lions, which is closely related to haplotypes found in Chad and DRC. The position of the Ethiopia1 lions in this study leads to the conclusion that these individuals do not form a unique group, but are instead assigned for a substantial part to a cluster that was not represented in the work by Bruche *et al.* (2012). Although the origin of the Ethiopia1 founder lions is disputed, it is claimed that they originate from the south-western part of Ethiopia (Tefera 2003), west of the Rift Valley, which has previously been suggested as a barrier for lion dispersal (Pitra *et al.* 2002; Burger *et al.* 2004; Dubach *et al.* 2005, 2013; Barnett *et al.* 2006b). The other captive Ethiopian population, Ethiopia2, contains a haplotype that clusters within the East Africa group. Assessment of the microsatellite data showed that Ethiopia2 individuals indeed contained a stronger signal from East Africa, compared to Ethiopia1. STRUCTURE analysis detected admixture in both captive Ethiopian lion populations that may be explained by the geographical location of Ethiopia. However, human-mediated translocations of lions between regions is not uncommon in zoo settings and may have contributed to the observed pattern. More data from free-ranging Ethiopian lions are required to determine if this pattern of admixture is accurately representative for that region. In Zambia, a substructure in the population is induced due to the two detected types of admixture: the Southern Africa cluster is admixed either with the West/Central Africa cluster, or with the East Africa cluster. These findings are parsimonious with the geographic isolation representative of Zambia's Luangwa Valley which is an offshoot of the Rift Valley System. We found no clear admixture between West/Central Africa with East Africa, possibly due to the Rift valley. The absence of a mitochondrial haplotype from outside the East Africa cluster in the Zambian individuals that were analyzed indicates that the pattern of admixture is likely due to male-mediated gene flow.

The mtDNA haplotype network shows the same four groups as identified in the STRUCTURE analysis: 1) West/Central Africa, 2) East Africa, 3) Southern Africa and 4) India. These groups have

also been proposed based on morphometric data, with a more basal clustering into a northern and a southern clade (Hemmer 1974). This coincides with the deepest split in the haplotype network which distinguishes a North group consisting of the West/Central African lion together with the Asiatic subspecies, and a South group consisting of lions from East and Southern Africa. Only a single or two closely related haplotypes are found in a single country, with two exceptions where more divergent haplotypes are present: 1) Ethiopia, which could be explained by the geographic location of the country as previously noted, and 2) RSA2, likely due to past translocations to and amongst small reserves in RSA (Miller *et al.* 2013). The geographical boundaries between the identified groups based on mtDNA and microsatellite data differ in the southern part of their range. STRUCTURE and PCA plots show that all populations from Namibia and RSA are assigned to Southern Africa, with a more central position for the admixed Zambia population, while East African haplotypes are found in RSA. The same discrepancy was previously described by Antunes *et al.* (2008) and attributed to sex-biased gene flow. The neighbour-joining tree, based on microsatellite data also shows a distinction between lions from West/Central Africa, and populations from East and Southern Africa. The basal position of the Indian and Kenyan lions probably results from the lower genetic diversity in these populations, as is indicated by the relatively high number of monomorphic loci. Elongation of branch length resulting from a population size reduction has been previously described, especially for DA as a measure of genetic distance (Kalinowski 2002). Despite of this DA is commonly accepted as the most suitable measure for inferring phylogenetic relationships, and therefore has been used in our analyses (Takezaki & Nei 1996, 2008). To further assess congruence between mtDNA and autosomal markers, a Mantel test was performed based on corrected Nei's genetic distances for both datasets (Supplemental Table S9). This illustrates a significant relationship (999 permutations, $P<0.01$) between both measures, which increases further after the exclusion of India.

Strongest congruence in AMOVA results between the autosomal and mtDNA data are found when using the groups identified by STRUCTURE. Microsatellite AMOVA show that the largest part of molecular variance is attributed to among-groups variance according to the IUCN classification, i.e. distinguishing an African and an Asiatic subspecies, which is also congruent with PCA results. This result is likely attributable to low genetic diversity and the high number of monomorphic loci in the Indian population rather than to long evolutionary distance. This is further confirmed by the fact that the Africa/Asia distinction leads to an exceptionally low percentage for among-groups variance when haplotype data are used. Haplotype data give the highest percentage for among-groups variance when following a distinction between a North group (West/Central Africa/India) and a South group (East Africa/Southern Africa), which is only moderately supported by microsatellite based among-group variance. Following the four groups identified by STRUCTURE shows a relative high among-group variance in both datasets, indicating a robust phylogenetic pattern that is reflected both in mtDNA and in autosomal DNA.

IBD explains the genetic distances on a continental scale and on a regional scale. The strong slope of the trend line in IBD analysis for West/Central Africa, compared to Southern and East Africa, is suggestive of near complete isolation between populations in the West/Central region. This is also supported by the high among-groups variance in the AMOVA. Based on the genetic distances (pairwise Fst and Nei's genetic distance), we conclude that all sampled populations are significantly differentiated from each other.

It was hypothesized that lion populations in West Africa and parts of Central Africa were especially vulnerable to declining levels of genetic diversity since fragmentation of the habitat is particularly severe in this region. However, we did not find significant heterozygotic deficiencies, reduced number of alleles or fixed loci in any of the six sampled populations in this region. The significantly negative Fis values (excess of heterozygotes) may be explained by the mating system as was also shown for prides in Selous GR (Spong *et al.* 2002), however we acknowledge the possible effect of a small sample size in our study. The strongly significant heterozygote deficiency observed in the Zambia lion population is likely to be the result of substructure in the population (Wahlund effect), which was consistent with the results from the STRUCTURE analysis. The significantly positive Fis value found in RSA1 is congruent with previous findings (Dubach *et al.* 2013): Dubach *et al.* (2013) reports a relative high value for the Kalahari population (RSA1), although it was tested as non significant. A high Fis value in the Ethiopia1 lions can be explained by the breeding history of the population, which was founded by five males and two females in 1948 (Bruche *et al.* 2012). Bottleneck analysis indicated that both RSA1 and Ethiopia1 have gone through recent population reductions. Similarly, Cameroon1 and Kenya appear to have experienced bottlenecks, which is consistent with observations obtained from monitoring studies (Iongh *et al.* 2009; Riggio *et al.* 2012). Since the excess of heterozygotes as a result of a bottleneck is transient, the Bottleneck approach only detects recent reductions in population size, which explains why historically documented bottlenecks i.e., Tanzania2 and India, were not detected.

The unexpected high levels of genetic diversity found in West/Central Africa lions could be explained by the fact that the range contraction and the decline of lion numbers is too recent to show clear signs of genetic erosion at this point. However, as genetic diversity is rapidly lost in small populations as a result of genetic drift and inbreeding, keeping the population at a genetically healthy level may require urgent management decisions to safeguard against these effects. Monitoring of an intensively managed lion population showed that drift and inbreeding were noticeable within five years after reintroduction of eleven founders from four genetic lineages (Trinkel *et al.* 2010). By showing a congruent phylogeographic pattern in both mtDNA and autosomal markers, our data illustrate which populations belong to the same evolutionary lineage and may contribute importantly to conservation decisions e.g., identifying suitable candidates for translocations or population augmentation.

Our study is the first to confirm that autosomal markers support the distinct genetic position of West/Central African lions within the African subspecies. The phylogenetic split between West/Central Africa and East/Southern Africa found in other species is reiterated in lions. Based on results derived from mtDNA data and from autosomal microsatellites, we recommend recognition and consideration of these four groups for management decisions: 1) West/Central Africa, 2) East Africa, 3) Southern Africa and 4) India. In consideration of genetic distinctions coupled with anthropogenic factors that are accelerating decline of wildlife in West and Central Africa, this region is of particular and urgent conservation importance. We support a revision of the taxonomic nomenclature as has been proposed by Barnett *et al.* (2014), following the deepest ancestral split found in the haplotype network, recognizing a North group and a South group. In addition, there may be arguments to warrant the Asiatic population its subspecies status (Dubach *et al.* 2013), although this would lead to a paraphyletic status of (one of) the other subspecies due to the close genetic relationship between lions from West/Central Africa and India. Primarily, as mtDNA, autosomal markers and morphological data show a congruent pattern, we believe that it is enough support for a taxonomic split within the African subspecies of the lion.

Acknowledgements

Samples were kindly provided by N. Vanherle (CURESS, Chad), S. Adam and selected members of the West/Central African lion network (ROCAL), R. Buij and B. Croes (Cameroon), ICCN, Garamba NP (DRC), South Africa National Parks (SANParks) (RSA2). We also thank Direction des Parcs Nationaux (DPN - Senegal) for facilitating sample collection and export. Further, we thank F. Oberwemmer (Leipzig Zoo) and IZW Leibnitz Institute for Zoo and Wildlife Research (Ethiopia1: Addis Ababa Zoo), and Sanaa Zoo, Yemen (Ethiopia2) for access to samples of two captive populations. We are grateful for the permit of the Ethiopian Wildlife Conservation Authority (formerly Wildlife Conservation Department) to obtain and export blood samples of the lions from Addis Ababa Zoo (Ethiopia1) through Zoo Leipzig, Germany. We would like to thank S. Bruche and coauthors for enabling us to include their data (Dataset 2) and S. Rabinowitz and S. Loss (American Museum of Natural History) for their contribution to the Senegal data (Dataset 3). We further thank H. Visser for contributing to processing of the samples. The investigations were supported by the Division for Earth and Life Sciences (ALW) with financial aid from the Netherlands Organization for Scientific Research (NWO) (project no. 820.01.002).

References

Alpers DL, Van Vuuren BJ, Arctander P, Robinson TJ (2004) Population genetics of the roan antelope (Hippotragus equinus) with suggestions for conservation. *Molecular ecology*, **13**, 1771–84.

Antunes A, Troyer JL, Roelke ME *et al.* (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 MA (1999) Phylogeography of three closely related African bovids (tribe Alcelaphini). *Molecular biology and evolution*, **16**, 1724–39.

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). *Proceedings of the Royal Society B: Biological Sciences*, **273**, 2119–2125.

Barnett R, Yamaguchi N, Shapiro B *et al.* (2014) Revealing the maternal demographic history of Panthera leo using ancient DNA and a spatially explicit genealogical analysis. *BMC Evolutionary Biology*, **14**, 70.

Bertola LD, van Hooft WF, Vrieling K *et al.* (2011) Genetic diversity, evolutionary history and implications for conservation of the lion (Panthera leo) in West and Central Africa. *Journal of Biogeography*, **38**, 1356–1367.

Brown DM, Brenneman R a, Koepfli K-P *et al.* (2007) Extensive population genetic structure in the giraffe. *BMC biology*, **5**, 57.

Bruche S, Gusset M, Lippold S *et al.* (2012) A genetically distinct lion (Panthera leo) population from Ethiopia. *European Journal of Wildlife Research*.

Burger J, Rosendahl W, Loreille O *et al.* (2004) Molecular phylogeny of the extinct cave lion Panthera leo spelaea. *Molecular phylogenetics and evolution*, **30**, 841–849.

Charruau P, Fernandes C, Orozco-Terwengel P *et al.* (2011) Phylogeography, genetic structure and population divergence time of cheetahs in Africa and Asia: evidence for long-term geographic isolates. *Molecular ecology*, **15**, 367–371.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–14.

DeBruyne R (2005) A case study of apparent conflict between molecular phylogenies: the interrelationships of African elephants. *Cladistics*, **21**, 31–50.

Driscoll CA, Menotti-Raymond M, Nelson G, Goldstein D, O'Brien SJ (2002) Genomic microsatellites as evolutionary chronometers: a test in wild cats. *Genome research*, **12**, 414–423.

Dubach JM, Briggs MB, White PA, Ament BA, Patterson BD (2013) Genetic perspectives on "Lion Conservation Units" in Eastern and Southern Africa. *Conservation Genetics*, **1942**.

Dubach J, Patterson BD, Briggs MB *et al.* (2005) Molecular genetic variation across the southern and eastern geographic ranges of the African lion, Panthera leo. *Conservation Genetics*, **6**, 15–24.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, **14**, 2611–20.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary bioinformatics online*, **1**, 47–50.

Flagstad O, Syvertsen PO, Stenseth NC, Jakobsen KS (2001) Environmental change and rates of evolution: the phylogeographic pattern within the hartebeest complex as related to climatic variation. *Proceedings. Biological sciences / The Royal Society*, **268**, 667–77.

Freeman AR, MacHugh DE, McKeown S *et al.* (2001) Sequence variation in the mitochondrial DNA control region of wild African cheetahs (Acinonyx jubatus). *Heredity*, **86**, 355–62.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and xation indices (version 2.9.3). , 215–225.

Hassanin A, Ropiquet A, Gourmand A-L, Chardonnet B, Rigoulet J (2007) Mitochondrial DNA variability in Giraffa camelopardalis: consequences for taxonomy, phylogeography and conservation of giraffes in West and central Africa. *Comptes rendus biologies*, **330**, 265–74.

Hemmer H (1974) *Untersuchungen zur Stammesgeschichte der Pantherkatzen (Pantherinae) 3 Zur Artgeschichte des Löwen Panthera (Panthera) leo (Linnaeus 1758).* Veröffentlichungen der Zoologischen Staatssammlung München, München, Germany.

Van Hooft WF, Groen AF, Prins HHT (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**, 267–79.

IUCN SSC Cat Specialist Group (2006a) *Conservation strategy for the lion in West and Central Africa*. IUCN, Gland, Switzerland.

IUCN SSC Cat Specialist Group (2006b) *Conservation strategy for the lion (Panthera leo) in Eastern and Southern Africa*. IUCN, Gland, Switzerland.

Iongh H de, Tumenta P, Croes B *et al.* (2009) Threat of a lion population extinction in Waza National Park, North Cameroon. *CATnews*, **50**, 26–27.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics (Oxford, England)*, **23**, 1801–6.

Kalinowski ST (2002) Evolutionary and statistical properties of three genetic distances. *Molecular ecology*, **11**, 1263–73.

Kalinowski ST (2005) Hp-Rare 1.0: a Computer Program for Performing Rarefaction on Measures of Allelic Richness. *Molecular Ecology Notes*, **5**, 187–189.

Kalinowski ST (2010) The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. *Heredity*, **106**, 625–32.

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

Miller SM, Bissett C, Burger A *et al.* (2013) Management of reintroduced lions in small , fenced reserves in South Africa : an assessment and guidelines. *South African Journal of Wildlife Research*, **43**, 138–154.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics (Oxford, England)*, **28**, 2537–9.

Pinho C, Harris DJ, Ferrand N (2007) Comparing patterns of nuclear and mitochondrial divergence in a cryptic species complex: the case of Iberian and North African wall lizards (Podarcis, Lacertidae). *Biological Journal of the Linnean Society*, **91**, 121–133.

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, **90**, 502–503.

Pitra C, Hansen AJ, Lieckefeldt D, Arctander P (2002) An exceptional case of historical outbreeding in African sable antelope populations. *Molecular ecology*, **11**, 1197–1208.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–59.

Pusey AE, Packer C, Erhoff-Mulder MB (1987) The Evolution os Sex-biased Dispersal in Lions. *Behaviour*, **101**, 275–310.

Rato C, Carranza S, Perera a, Carretero M a, Harris DJ (2010) Conflicting patterns of nucleotide diversity between mtDNA and nDNA in the Moorish gecko, Tarentola mauritanica. *Molecular phylogenetics and evolution*, **56**, 962–71.

Riggio J, Jacobson A, Dollar L *et al.* (2012) The size of savannah Africa: a lion's (Panthera leo) view. *Biodiversity and Conservation*, **22**, 17–35.

Roca AL, Georgiadis N, O'Brien SJ (2007) Cyto-nuclear genomic dissociation and the African elephant species question. *Quaternary international : the journal of the International Union for Quaternary Research*, **169-170**, 4–16.

Singh HS, Gibson L (2011) A conservation success story in the otherwise dire megafauna extinction crisis: The Asiatic lion (Panthera leo persica) of Gir forest. *Biological Conservation*, **144**, 1753–1757.

Smitz N, Berthouly C, Cornélis D *et al.* (2013) Pan-African Genetic Structure in the African Buffalo (Syncerus caffer): Investigating Intraspecific Divergence. *PloS one*, **8**, e56235.

Sota T, Vogler AP (2001) Incongruence of mitochondrial and nuclear gene trees in the Carabid beetles Ohomopterus. *Systematic biology*, **50**, 39–59.

Spong G, Stone J, Creel S, Björklund M (2002) Genetic structure of lions (Panthera leo L.) in the Selous Game Reserve: implications for the evolution of sociality. *Journal of Evolutionary Biology*, **15**, 945–953.

Takezaki N, Nei M (1996) Genetic Distances and Reconstruction of Phylogenetics Trees from Microsatellite DNA. *Genetics*, **144**, 389–399.

Takezaki N, Nei M (2008) Empirical tests of the reliability of phylogenetic trees constructed with microsatellite DNA. *Genetics*, **178**, 385–92.

Takezaki N, Nei M, Tamura K (2010) POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. *Molecular biology and evolution*, **27**, 747–52.

Tefera M (2003) Phenotypic and reproductive characteristics of lions (Panthera leo) at Addis Ababa Zoo. *Biodiversity and Conservation*, **12**, 1629–1639.

Trinkel M, Funston P, Hofmeyr M *et al.* (2010) Inbreeding and density-dependent population growth in a small, isolated lion population. *Animal Conservation*, **13**, 374–382.

Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular ecology*, **17**, 2107–2121.

Data accessibility

Microsatellite data are accessible in Supplemental Table S3. All sequence data generated in this study have been submitted to GenBank. Accession numbers are listed in Supplemental Table S4.

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 microsatellite loci used in the different lion populations. No. loci: number of included loci.

Set	Population	No. loci	Microsatellite loci
1	**Benin**	20	FCA026, FCA032, FCA057, FCA075, FCA085, FCA091, FCA094, FCA097, FCA126, FCA136, FCA144, FCA161, FCA178, FCA191, FCA205, FCA208, FCA211, FCA224, FCA247, FCA275
	Cameroon1	20	FCA026, FCA032, FCA057, FCA075, FCA085, FCA091, FCA094, FCA097, FCA126, FCA136, FCA144, FCA161, FCA178, FCA191, FCA205, FCA208, FCA211, FCA224, FCA247, FCA275
	Cameroon2	20	FCA026, FCA032, FCA057, FCA075, FCA085, FCA091, FCA094, FCA097, FCA126, FCA136, FCA144, FCA161, FCA178, FCA191, FCA205, FCA208, FCA211, FCA224, FCA247, FCA275
	Chad	15	FCA032, FCA057, FCA075, FCA091, FCA094, FCA097, FCA126, FCA136, FCA144, FCA161, FCA205, FCA208, FCA224, FCA247, FCA275
	DRC	20	FCA026, FCA032, FCA057, FCA075, FCA085, FCA091, FCA094, FCA097, FCA126, FCA136, FCA144, FCA161, FCA178, FCA191, FCA205, FCA208, FCA211, FCA224, FCA247, FCA275
	Ethiopia2	12	FCA032, FCA075, FCA091, FCA094, FCA097, FCA126, FCA144, FCA205, FCA208, FCA224, FCA247, FCA275
	Kenya	20	FCA026, FCA032, FCA057, FCA075, FCA085, FCA091, FCA094, FCA097, FCA126, FCA136, FCA144, FCA161, FCA178, FCA191, FCA205, FCA208, FCA211, FCA224, FCA247, FCA275
	Tanzania1	20	FCA026, FCA032, FCA057, FCA075, FCA085, FCA091, FCA094, FCA097, FCA126, FCA136, FCA144, FCA161, FCA178, FCA191, FCA205, FCA208, FCA211, FCA224, FCA247, FCA275
	Tanzania2	20	FCA026, FCA032, FCA057, FCA075, FCA085, FCA091, FCA094, FCA097, FCA126, FCA136, FCA144, FCA161, FCA178, FCA191, FCA205, FCA208, FCA211, FCA224, FCA247, FCA275
	Zambia	20	FCA026, FCA032, FCA057, FCA075, FCA085, FCA091, FCA094, FCA097, FCA126, FCA136, FCA144, FCA161, FCA178, FCA191, FCA205, FCA208, FCA211, FCA224, FCA247, FCA275
	Namibia	20	FCA026, FCA032, FCA057, FCA075, FCA085, FCA091, FCA094, FCA097, FCA126, FCA136, FCA144, FCA161, FCA178, FCA191, FCA205, FCA208, FCA211, FCA224, FCA247, FCA275
	RSA1	20	FCA026, FCA032, FCA057, FCA075, FCA085, FCA091, FCA094, FCA097, FCA126, FCA136, FCA144, FCA161, FCA178, FCA191, FCA205, FCA208, FCA211, FCA224, FCA247, FCA275
	RSA2	20	FCA026, FCA032, FCA057, FCA075, FCA085, FCA091, FCA094, FCA097, FCA126, FCA136, FCA144, FCA161, FCA178, FCA191, FCA205, FCA208, FCA211, FCA224, FCA247, FCA275
	India	20	FCA026, FCA032, FCA057, FCA075, FCA085, FCA091, FCA094, FCA097, FCA126, FCA136, FCA144, FCA161, FCA178, FCA191, FCA205, FCA208, FCA211, FCA224, FCA247, FCA275
2	**Ethiopia1**	10*	FCA006, FCA082, FCA097, FCA136, FCA161, FCA178, FCA191, FCA200, FCA211, FCA249
3	**Senegal**	12**	FCA032 , FCA075 , FCA096 , FCA100 , FCA124 , FCA126 , FCA132 , FCA208 , FCA212 , FCA225 , FCA229 , FCA275

* for two samples all twenty loci were analysed, for the others ten loci were analyzed of which six overlap with Dataset 1

** not resized

Supplemental Table S2. Primers used for amplification of microsatellites and mtDNA.
Location: targeted genetic marker, for microsatellite loci the chromosomal location in cat (Felis catus) is indicated; Annealing Temp: annealing temperature in ˚C.

Type	Location	Primer name	Annealing Temp (˚C)	Sequence (5'-3')	Origin
mtDNA	Cytochrome b, tRNAPro, tRNAThr, partial control region	F: 1F	51	CGTTGTACTTCAACTATAAGAACTT	own design
		R: 1R		ATGGGATTGCTGATAGGAGATTAG	own design
		F: 2F	53	GTGGGGCCAAATATCCTTTT	own design
		R: 2R internal*		GAAGGCCTAGGATATCTTTGATTG	own design
		F: 3F internal*	51	GACTCAGATAAAATTCCATTCCA	own design
		R: 3R internal*		CATTATTCCTCGCTGTTTGG	own design
		F: 4F internal*	51	CAATTATCCCTGCCCTCCA	own design
		R: 4R		TTTTTGGTTTACAAGACCAAGGTA	own design
		F: 5F	54	AAATCGCCTCCTCAAATGAA	own design
		R: 5R		AATATTCATGGGAGGGCAGTC	own design
Microsatellite	Chromosome D3 (Cat)	F: FCA026F	51	GGAGCCCTTAGAGCTATGCA	Menotti-Raymond et al., 1999
		R: FCA026R		TGTACACGCACCAAAAACAA	Menotti-Raymond et al., 1999
	Chromsome A2 (Cat)	F: FCA032F	51	GGCAATTCATGGTAGAGAAAAA	Menotti-Raymond et al., 1999
		R: FCA032R		CAAGAGTGCATTGGGCAGTA	Menotti-Raymond et al., 1999
	Chromsome C1 (Cat)	F: FCA057F	51	AAGTGTGGGATTGGGTGAAA	Menotti-Raymond et al., 1999
		R:FCA057R		CCATAAGAGGCTCTTAAAAACTGA	Menotti-Raymond et al., 1999
	Chromsome A2 (Cat)	F: FCA075F	51	ATGCTAATCAGTGGCATTTGG	Menotti-Raymond et al., 1999
		R:FCA075R		GAACAAAAATTCCAGACGTGC	Menotti-Raymond et al., 1999
	Chromsome A2 (Cat)	F: FCA085F	51	CTGTACATTTCTCTTCCCATTGC	Menotti-Raymond et al., 1999
		R: FCA085R		CCCCTACTGGGTGCACTG	Menotti-Raymond et al., 1999
	Chromsome B4 (Cat)	F: FCA091F	51	TGAGAACAAGCCATTAATAGCA	Menotti-Raymond et al., 1999
		R: FCA091R		CCCAAACATAAGGCTGCATT	Menotti-Raymond et al., 1999
	Chromsome F2 (Cat)	F: FCA094F	51	TCAAGCCCCATTTTACCTTC	Menotti-Raymond et al., 1999
		R: FCA094R		CACCTGAGCCAAAGGCTATC	Menotti-Raymond et al., 1999
	Chromsome B1 (Cat)	F: FCA097F	51	TAATGTTCAACTTGAATTGCTTCC	Menotti-Raymond et al., 1999
		R: FCA097R		GAACAGTAGTTTGCCCATACAGG	Menotti-Raymond et al., 1999
	Chromsome B1 (Cat)	F: FCA126F	51	GCCCCTGATACCCTGAATG	Menotti-Raymond et al., 1999
		R: FCA126R		CTATCCTTGCTGGCTGAAGG	Menotti-Raymond et al., 1999
	Chromsome F2 (Cat)	F: FCA136F	51	GAATGACATCGCCAATGAAA	Menotti-Raymond et al., 1999
		R: FCA136R		CCCCCCCAAAACTGATACTT	Menotti-Raymond et al., 1999
	Chromsome D1 (Cat)	F: FCA144F	51	GGAAATCCTGGAAACTTCTGC	Menotti-Raymond et al., 1999
		R: FCA144R		CCCGGCAAAATTATGAAGG	Menotti-Raymond et al., 1999
	Chromsome A3 (Cat)	F: FCA161F	51	TTACCGATACACACCTGCCA	Menotti-Raymond et al., 1999
		R: FCA161R		CACAGACGTGCTCTAGCCAA	Menotti-Raymond et al., 1999
	Chromsome A1 (Cat)	F: FCA178F	51	GTGCCCCATGAATCTCACTT	Menotti-Raymond et al., 1999
		R: FCA178R		TACAACTCAGGGGTCGTATGG	Menotti-Raymond et al., 1999
	Chromsome C1 (Cat)	F: FCA191F	51	TCCTGTTCCTATTCACCCTACA	Menotti-Raymond et al., 1999
		R: FCA191R		GCATGGCACTTTTGTTGAGA	Menotti-Raymond et al., 1999
	Chromsome B3 (Cat)	F: FCA205F	51	CCTGCTCTCAAGGAGCTCC	Menotti-Raymond et al., 1999
		R: FCA205R		CCCATTTCTCCTACCAGTTCC	Menotti-Raymond et al., 1999
	Chromsome A3 (Cat)	F: FCA208F	51	TCAGGGTTCAAAAAAAGAAAAA	Menotti-Raymond et al., 1999
		R: FCA208R		CAAAGCACCAGCTTAGAAGTCA	Menotti-Raymond et al., 1999
	Chromsome B1 (Cat)	F: FCA211F	51	TGTAGAACATAATGCCTCAGCC	Menotti-Raymond et al., 1999
		R: FCA211R		TCTTGAACCTATTTCCCCACA	Menotti-Raymond et al., 1999
	Chromsome A3 (Cat)	F: FCA224F	51	CTGGGTGCTGACAGCATAGA	Menotti-Raymond et al., 1999
		R: FCA224R		TGCCAGAGTTGTATGAAAGGG	Menotti-Raymond et al., 1999
	Chromsome C1 (Cat)	F: FCA247F	51	GGAAATTAGGAGCTCTGCCA	Menotti-Raymond et al., 1999
		R: FCA247R		AAGATTTACCCAGTTGCCCC	Menotti-Raymond et al., 1999
	Chromsome B2 (Cat)	F: FCA275F	51	TTGGCTGCCCAGTTTTAGTT	Menotti-Raymond et al., 1999
		R: FCA275R		ACGAAGGGGCAGGACTATCT	Menotti-Raymond et al., 1999

* internal primers were only used when amplification with primerpair 2F-4R failed

Supplemental Table S3. (online only) Microsatellite data for 20 loci in 16 lion populations.

Supplemental Table S4. Identified haplotypes and accompanying accession numbers from Genbank.

Haplotype	Country	Genbank Accession	Source
A	Senegal	KJ652247	this publication
B	Benin	GU131164 - GU131165	this publication
C	Cameroon1 + Cameroon2	GU131174 - GU131175, AY781202 - AY781205	Bertola et al., 2011
D	Cameroon1	GU131170 - GU131173	Bertola et al., 2011
E	Cameroon2 + Chad	GU131169, AY781198 - AY781199, AY781197	Bertola et al., 2011
F	Chad+DRC	AY781200, DQ018993 - DQ018994	this publication
G	DRC	KJ652248	this publication
H	Ethiopia1	KJ652249	this publication
I	Ethiopia2	AY781207 - AY781210	Bertola et al., 2011
J	Kenya + Tanzania1 + Tanzania2 + Zambia + RSA1 + RSA2	GU131166 - GU131168	this publication
K	Zambia	KJ652250	this publication
L	Namibia	KJ652251	this publication
M	RSA2	GU131183 - GU131185	this publication
N	India	GU131176 - GU131178, AY781206	this publication

Unique point mutations (i.e. observed in a single sample) were checked by an independent PCR and sequencing. This resulted in correcting the following previously published sequences by Bertola et al. (2011) on Genbank (September 2013): GU131167-GU131168, GU131170, GU131172-GU131175, GU131178, GU131183, GU131185. Conclusions based on the uncorrected sequences as published in Bertola et al. (2011) still hold.

Supplemental Table S5. Results of an AMOVA for a microsatellite dataset of 12 lion populations and a mtDNA dataset of 16 lion populations.

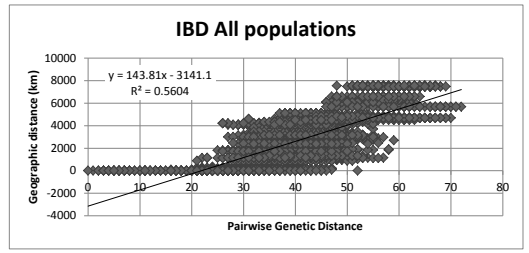
Four different divisions were tested: no substructure (all in one group), following IUCN categorisation (2 groups), following haplotype structure (2 groups) and following STRUCTURE results (4 groups); Variance component AG: Among Groups; AP: Among Populations; AI: Among Individuals; Within Groups; WP: Within Populaions; WI: Within Individuals; df: degrees of freedom; % Total: percentage of total molecular variance explained on the different levels

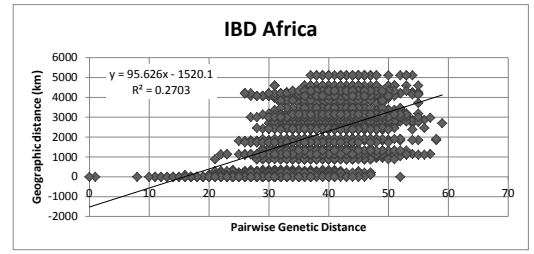
Number of groups	Division	Variance component	microsatellites				mtDNA			
			df	Sum of squares	Variants of components	% Total	df	Sum o f squares	Variants of components	% Total
1	**No division**	**AP**	11	494.445	2.37882	29.9	15	406.281	5.29171	98.1
		AI/WP	88	489.720	-0.0225	-0.3	67	6.972	0.10406	1.9
		WI	100	561.000	5.61000	70.4	-	-	-	-
2	**Africa; Asia**	**AG**	1	151.848	3.09325	29.6	1	32552	0.04718	0.9
		AP/WG	10	342.597	1.76592	16.9	14	373729	5.28484	97.2
		WP	188	1050.720	5.58894	53.5	67	6972	0.10406	1.9
2	**North (West/Central Africa + India); South (East & Southern Africa)**	**AG**	1	96.294	0.62456	7.5	1	183037	3.93253	54.6
		AP/WG	10	398.151	2.06955	25.0	14	223244	3.17305	44.0
		WP	188	1050.720	5.58894	67.5	67	6972	0.10406	1.4
4	**West/Central Africa; East Africa; Southern Africa; India**	**AG**	3	289.637	1.43925	17.4	3	237368	3.3016	52.3
		AP/WG	8	204.808	1.25075	15.1	12	168913	2.91068	46.1
		WP	188	1050.720	5.58894	67.5	67	6972	0.10406	1.7

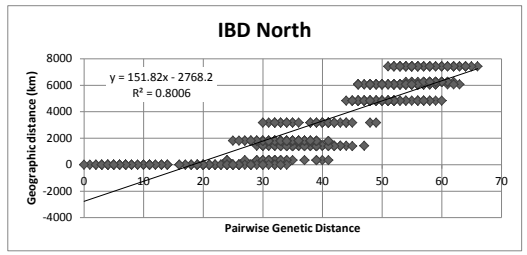
Supplemental Table S6. Results of the Mantel tests indicating IBD effects in lion populations on continental and regional scale.

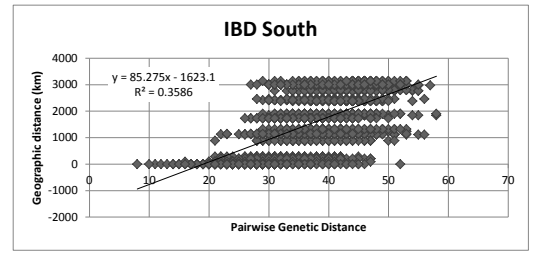
NPop: number of included populations; Pair. comparisons: number of pairwise comparisons; Rxy: correlation coefficient (P(Rxy_random≥Rxy_data), one-tailed, 999 permutations); x: slope of regression line (95% confidence interval (CI)); AG: Among group variance according to regional AMOVA.

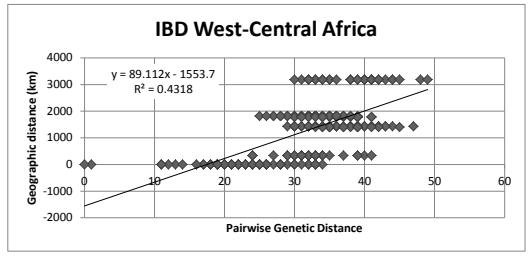
Groups	NPop	Pair. comparisons	Rxy	x (95% CI)	AG
All populations	12	4950	0.749 (P≤0.001)	143.81 (140.26-147.36)	29.86
Africa	11	4005	0.520 (P≤0.001)	95.63 (90.76-100.50)	22.10
North (West/Central Africa; India)	5	561	0.895 (P≤0.001)	151.82 (145.53-158.12)	42.18
South (East Africa; Southern Africa)	7	2145	0.599 (P≤0.001)	85.28 (80.44-90.11)	19.47
West/Central Africa	4	276	0.657 (P≤0.001)	89.11 (76.95-101.27)	20.48
East Africa	3	351	0.495 (P≤0.001)	7.94 (6.47-9.40)	13.77
Southern Africa	4	741	0.555 (P≤0.001)	54.1 (48.25-59.96)	15.18

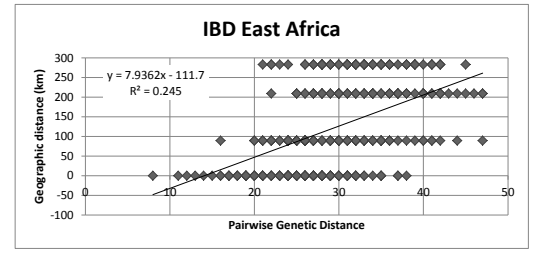


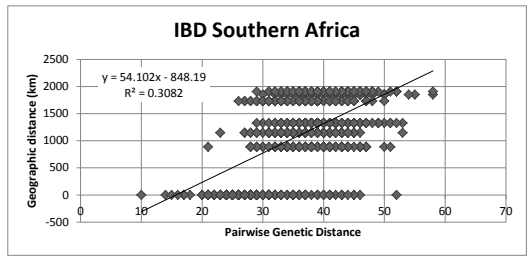












Supplemental Table S7. Pairwise Fst (below diagonal) and Nei's genetic distances (above diagonal) based on 20 microsatellite loci from 14 lion populations.

	Benin	Cameroon1	Cameroon2	Chad	DRC	Ethiopia2	Kenya	Tanzania1	Tanzania2	Zambia	Namibia	RSA1	RSA2	India
Benin	-	0.789	0.801	0.716	0.902	1.900	1.091	0.872	0.659	0.630	0.920	0.997	1.289	1.990
Cameroon1	0.242	-	0.491	0.619	0.493	1.794	0.957	0.805	0.903	0.772	1.129	0.991	1.027	1.493
Cameroon2	0.268	0.134	-	0.600	0.571	1.311	1.362	0.927	1.014	0.839	1.022	0.804	1.050	1.918
Chad	0.169	0.160	0.104	-	0.718	1.089	1.085	1.042	0.963	0.959	1.081	0.998	1.261	1.789
DRC	0.225	0.147	0.140	0.146	-	1.344	1.175	0.783	0.820	0.629	0.989	0.874	1.115	1.228
Ethiopia2	0.396	0.337	0.301	0.269	0.239	-	1.225	1.353	2.018	1.539	1.488	1.217	1.458	2.193
Kenya	0.336	0.337	0.348	0.267	0.264	0.296	-	0.373	0.463	0.729	1.065	0.882	0.837	1.065
Tanzania1	0.265	0.264	0.247	0.218	0.170	0.263	0.163	-	0.196	0.358	0.834	0.493	0.612	1.589
Tanzania2	0.271	0.321	0.318	0.260	0.231	0.391	0.229	0.064	-	0.518	0.932	0.768	0.888	2.098
Zambia	0.189	0.225	0.225	0.189	0.134	0.267	0.234	0.087	0.183	-	0.631	0.523	0.601	1.724
Namibia	0.289	0.300	0.259	0.248	0.218	0.307	0.334	0.245	0.304	0.206	-	0.506	0.511	1.922
RSA1	0.253	0.262	0.211	0.180	0.188	0.210	0.246	0.133	0.234	0.138	0.156	-	0.467	1.683
RSA2	0.268	0.262	0.214	0.228	0.198	0.246	0.258	0.170	0.253	0.139	0.166	0.118	-	1.775
India	0.699	0.602	0.720	0.661	0.552	0.736	0.589	0.577	0.643	0.575	0.601	0.551	0.562	-

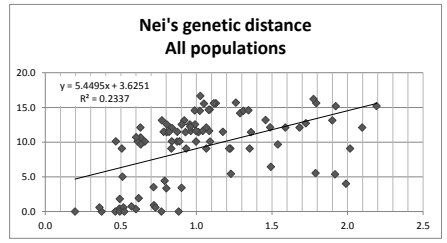
Supplemental Table S8. Genetic variation in microsatellite loci and mtDNA among 16 lion populations.

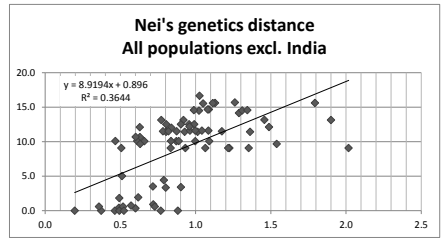
No.loci: number of amplified loci; Np: number of polymorphic loci (% of total amplified loci); Na: average number of alleles per locus; Allelic range: average over allelic ranges calculated as number of repeats per locus; NaPr: private allelic richness, calculated over the complete dataset; Haplo.: haplotypes detected in this population, as referred to in Figure 4; Ho: observed heterozygosity; uHe: unbiased expected heterozygosity; Fis: Fis index, P indicated between brackets for samples which tested significant; Bottlen.: indicated are significant results from the Wilcoxon signed-rank test in the program Bottleneck. Standard errors (SE) are presented in parentheses.

Set	Population	No. loci	Np (%)	Na (±SE)	Allelic range (±SE)	NaPr	Haplo.	Ho (±SE)	uHe (±SE)	Fis (P)	Bottlen.
1	**Benin**	20	19 (95%)	2.95 (±1.00)	5.10 (±4.81)	0.05	B	0.65 (±0.07)	0.55 (±0.05)	-0.204 (P<0.01)	
	Cameroon1	20	20 (100%)	3.20 (±0.77)	4.65 (±3.82)	0.00	C+D	0.68 (±0.05)	0.61 (±0.02)	-0.129 (P<0.01)	P<0.05
	Cameroon2	20	19 (95%)	2.85 (±0.88)	4.75 (±3.25)	0.05	C+E	0.58 (±0.07)	0.61 (±0.05)	0.060	
	Chad	15	15 (100%)	2.60 (±1.82)	7.40 (±5.77)	0.20	E+F	0.6 (±0.07)	0.56 (±0.05)	0.085	
	DRC	20	20 (100%)	4.65 (±1.87)	6.40 (±6.37)	0.10	F+G	0.74 (±0.04)	0.7 (±0.03)	-0.066	
	Ethiopia2	12	10 (83.3%)	1.25 (±1.16)	3.33 (±3.62)	0.00	I	0.44 (±0.1)	0.41 (±0.07)	-0.068	
	Kenya	20	17 (85%)	2.65 (±1.04)	5.10 (±4.20)	0.00	J	0.51 (±0.06)	0.5 (±0.05)	-0.025	P<0.01
	Tanzania1	20	20 (100%)	4.65 (±1.57)	7.50 (±4.66)	0.00	J	0.64 (±0.04)	0.65 (±0.03)	0.019	
	Tanzania2	20	20 (100%)	3.65 (±1.18)	5.65 (±3.69)	0.00	J	0.56 (±0.05)	0.57 (±0.04)	0.018	
	Zambia	20	20 (100%)	4.75 (±1.74)	6.30 (±4.11)	0.15	J+K	0.57 (±0.05)	0.69 (±0.03)	0.182 (P<0.01)	
	Namibia	20	20 (100%)	3.45 (±1.00)	6.40 (±4.32)	0.05	L	0.56 (±0.04)	0.57 (±0.03)	0.011	
	RSA1	20	20 (100%)	4.00 (±1.08)	5.95 (±3.61)	0.05	J	0.61 (±0.06)	0.66 (±0.03)	0.082 (P<0.05)	P<0.01
	RSA2	20	20 (100%)	4.60 (±1.27)	6.65 (±4.39)	0.25	J+M	0.69 (±0.04)	0.69 (±0.03)	-0.002	
	India	20	5 (25%)	1.45 (±0.89)	0.95 (±0.84)	0.10	N	0.11 (±0.05)	0.13 (±0.05)	0.095	
2	**Ethiopia1**	10*	8 (80%)	3.00 (±1.56)	3.00 (±2.48)	0.05	H	0.45 (±0.11)	0.49 (±0.09)	0.165 (P<0.05)	P<0.01
3	**Senegal**	12**	12 (100%)	3.33 (±1.56)	5.42 (±3.32)	-	A	0.63 (±0.08)	0.54 (±0.07)	-0.079	

Supplemental Table S9. Nei's (corrected) genetic distances for microsatellite (below diagonal) and mtDNA data (above diagonal) of 14 lion populations, and results of Mantel tests, including all populations, and excluding India.

	Benin	Cameroon1	Cameroon2	Chad	DRC	Ethiopia2	Kenya	Tanzania1	Tanzania2	Zambia	Namibia	RSA1	RSA2	India
Benin	-	4.435	3.345	3.511	3.412	13.145	10.075	10.075	10.075	10.670	13.136	10.075	14.166	4.014
Cameroon1	0.789	-	0.418	1.922	1.822	15.615	12.531	12.531	12.531	13.129	15.604	12.531	16.643	6.451
Cameroon2	0.801	0.491	-	0.334	0.735	14.516	11.435	11.435	11.435	12.032	14.505	11.435	15.540	5.360
Chad	0.716	0.619	0.600	-	0.902	14.679	11.599	11.599	11.599	12.197	14.669	11.599	15.700	5.525
DRC	0.902	0.493	0.571	0.718	-	14.575	11.497	11.497	11.497	12.094	14.565	11.497	15.600	5.425
Ethiopia2	1.900	1.794	1.311	1.089	1.344	-	9.077	9.077	9.077	9.672	12.137	9.077	13.170	15.189
Kenya	1.091	0.957	1.362	1.085	1.175	1.225	-	0.000	0.000	0.584	9.071	0.000	10.099	12.107
Tanzania1	0.872	0.805	0.927	1.042	0.783	1.353	0.373	-	0.000	0.584	9.071	0.000	10.099	12.107
Tanzania2	0.659	0.903	1.014	0.963	0.820	2.018	0.463	0.196	-	0.584	9.071	0.000	10.099	12.107
Zambia	0.630	0.772	0.839	0.959	0.629	1.539	0.729	0.358	0.518	-	9.665	0.000	10.694	12.705
Namibia	0.920	1.129	1.022	1.081	0.989	1.488	1.065	0.834	0.932	0.631	-	9.071	5.028	15.179
RSA1	0.997	0.991	0.804	0.998	0.874	1.217	0.882	0.493	0.768	0.523	0.506	-	10.099	12.107
RSA2	1.289	1.027	1.050	1.261	1.115	1.458	0.837	0.612	0.888	0.601	0.511	0.467	-	16.201
India	1.990	1.493	1.918	1.789	1.228	2.193	1.065	1.589	2.098	1.724	1.922	1.683	1.775	-





Supplemental Information S1. Details on sample storage, DNA extraction, PCR, fragment analysis and sequencing.

Permits

No samples were collected specifically for this study, and all included samples had been collected during previous fieldwork and stored in biobanks. More detailed information about permits and issuing authorities for each included population is given below:

Benin: Direction Generale des Forets et Ressources Naturelles – DGFRN (National Directorate for Forests and Natural Resources), local park authorities Pendjari Biosphere Reserve: samples were collected during previous fitting of GPS collars on lions (publication in prep. by Sogbohossou *et al.*), no individuals were handled for this study.

Cameroon: Ministry of Environment and Forestry, local park authorities Waza National Park and Bénoué Ecosystem: samples were collected during previous fitting of GPS collars on lions (publication by Tumenta *et al.* (2009)), no individuals were handled for this study.

Chad: Ministry of Environment, Water, & Fisheries, local park authorities Zakouma National Park: samples were collected during previous studies, no individuals were handled for this study.

DRC: Institut Congolais pour la Conservation de la Nature – African Parks Network, Garamba National Park management: samples were collected during previous fitting of GPS collars on lions, no individuals were handled for this study.

Zambia: Zambia Wildlife Authority (ZAWA): samples were collected during previous studies (publication by Dubach *et al.* (2013)), no individuals were handled for this study.

RSA: South African National Parks (SANParks): samples were collected during previous studies, no individuals were handled for this study.

Kenya: Kenya Wildlife Service (KWS): samples were collected during previous fitting of GPS collars on lions (publication in prep. by Jirmo *et al.*), no individuals were handled for this study.

Ethiopia (captive), Tanzania, Namibia, RSA: samples were obtained from the IZW Leibniz Institute for Zoo and Wildlife Research (Germany). Samples were collected during previous studies (publications by Driscoll *et al.* (2002) and Bruche *et al.* (2012)), no individuals were handled for this study.

India: samples were obtained from National Cancer Institute (NCI) (U.S.A.): samples were collected during previous studies (publication by Driscoll *et al.*, 2002), no individuals were handled for this study.

Sample storage and DNA extraction

Blood and tissue samples were preserved dried (Zambia) or in buffer (0.15 M NaCl, 0.05 M Tris-HCl, 0.001 M EDTA, pH = 7.5) and stored at -20°C. DNA was extracted using the Dneasy Blood & Tissue kit (Qiagen) following the manufacturer's protocol.

Microsatellite analysis

Twenty microsatellite loci, originally developed for domestic cat (Menotti-Raymond *et al.* 1999), were selected from a set that had previously been used in lion (Driscoll *et al.* 2002) (see Supplemental Table S2). A selection was made based on high variability, high amplification success and no apparent presence of null alleles. To enable resizing to already published datasets from Driscoll *et al.* (2002) and Bruche *et al.* (2012), four samples of these studies (Tanzania10 and RSA10; Ethiopia12 and Ethiopia13) were included for all 20 microsatellites. During PCR the products were fluorescently labelled (HEX, TAMRA and FAM) by adding M13 tails to the 5' end (Schuelke 2000). PCR reactions contained 0.75 mM $MgCl_2$, 0.4 mg/ml bovine serum albumin, 10x PCR buffer, 200 µM dNTPs, 0.1 U/µl Taq polymerase, 0.4 µM of both amplification primers and the M13 fluorescently labeled primer, and 1 µl of DNA template in a total volume of 15 µl. The PCR reaction was performed using an initial denaturation step of 94°C for 4min, followed by 40 cycles of 94°C for 20s, 51°C for 1 min, 72°C for 1 min and a final extension step of 72°C for 10 min. PCR products with different labels and non-overlapping fragment sizes (min. 20 bp difference between longest and shortest allele documented) were pooled and run on a MegaBACE sequencer (GE Health Care, Eindhoven, The Netherlands) or ABI3730XL (Macrogen Inc., Amsterdam, The Netherlands) with ET-ROX400 or ROX400 as an internal standard. To enable comparisons between runs and machines, we included a minimum of two known samples for every locus in each run. Allele lengths were scored using MegaBace Fragment Profiler version 1.2 (Amersham Biosciences, 2003) or Peak Scanner Software v1.0 (Life Technologies). Samples with a weak or distorted signal were re-amplified and were included in a subsequent run.

Microsatellite data were checked for potential null alleles and allelic dropout using Microchecker (Van Oosterhout *et al.* 2004). The data were tested for linkage disequilibrium using the Fisher's exact test in GENEPOP 4.2.1 (Raymond & Rousset 1995), applying 10000 dememorisations, 100 batches and 5000 iterations per batch as Markov chain parameters.

mtDNA analysis

PCRs were performed with three primer pairs (Supplemental Table S2), designed with the web-based software Primer3v 0.4.0 (Rozen & Skaletsky 2000). PCR reactions contained 1 mM $MgCl_2$, 0.4 mg/ml bovine serum albumin, 10x PCR buffer, 200 µM dNTPs, 0.1 U/µl Taq polymerase, 0.4 µM of both amplification primers and 1 µl of DNA template in a total volume of 20 µl. The PCR program consisted of an initial denaturation step of 94°C for 4min, followed by 40 cycles of 94°C for 20s, annealing temperature ranging from 51°C to 54°C depending on the primer set, for 1 min, 72°C for 1 min and a final extension step of 72°C for 10 min. Sequencing was performed by Macrogen Inc., Amsterdam, The Netherlands.

Quality control

Out of a total of 2188 data points (Dataset 1: 104 samples * 20 loci + 4 samples (Chad) * 15 loci + 4 samples (Ethiopia2)* 12 loci), Dataset 1 included a total of 28 missing genotypes (1.28%). None of the individuals had missing values at more than two loci. Indications for stuttering errors or null-alleles, as is suggested by the general excess of homozygotes in Microchecker, were identified in three populations for one locus (FCA178 in Ethiopia1 and Namibia; FCA211 in RSA1) and in the Zambia population for six loci (FCA026, FCA057, FCA094, FCA208, FCA211 and FCA224). However, in the case of the Zambia population this is likely to be the result of genetic structure within the population (see results STRUCTURE analysis). Since none of the loci were consistently positive for more than two populations, we included all loci in downstream analyses. There was no indication of allelic dropout. Pairwise comparison of loci in each population identified significant linkage ($P<0.05$) in 37 cases in a total of 2850 comparisons (1.30%). No loci were consistently in linkage disequilibrium across populations and a pairwise comparison between loci on the entire dataset did not reveal significant linkage.

References

Driscoll CA, Menotti-Raymond M, Nelson G, Goldstein D, O'Brien SJ (2002) Genomic microsatellites as evolutionary chronometers: a test in wild cats. *Genome research*, **12**, 414–423.

Menotti-Raymond M, David V, Lyons LA *et al.* (1999) A genetic linkage map of microsatellites in the domestic cat (Felis catus). *Genomics*, **57**, 9–23.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-Checker: Software for Identifying and Correcting Genotyping Errors in Microsatellite Data. *Molecular Ecology Notes*, **4**, 535–538.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 248–249.

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds S. K, S. M), pp. 365–386. Humana Press, Totowa, NY.

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology*, **18**, 233–234.

Supplemental Information S2. Details on DNA extraction, PCR, fragment analysis for Dataset 3.

Permits

Senegal: No specific permits were required for the collection and export of samples from Senegal. The samples were collected as part of a lion survey carried out jointly with the Senegalese national park authority (Direction des Parcs Nationaux; DPN). DPN waived all permit requirements, to facilitate the timely assessment of lion population status in the park, including of its genetic makeup. Samples were collected non-invasively (scat), no individuals were handled for this study.

DNA extraction & species identification

Genomic DNA was extracted from scat samples using the QIAmp DNA Stool Mini Kit (QIAGEN, Valencia, California, USA) following the manufacturer's protocol for isolation from stool for human DNA analysis with some modifications to improve DNA quality and quantity. All instruments were cleaned with DNAaway™ (Molecular BioProducts, San Diego, California, USA) and placed in an ultraviolet cross-linker prior to use. Additionally, all extractions were performed on a dedicated lab bench and in a pre-PCR laboratory to avoid contamination. Two microcentrifuge tubes were prepared with approximately 180 – 200 mg per sample in each using shavings from the outer layer of the scat sample. Samples were incubated overnight (approximately 18 h) at 56 °C on a rotator (approximately 22 RPM) with 1.5 mL of ASL buffer. All of the supernatant for each unique sample was combined into two tubes, with one tube containing 1.5 mL of supernatant and the other tube having all remaining supernatant up to 1.5 mL. A proportionate amount of InhibitEx tablet was added to the second tube of supernatant if the volume was less than 1.5 mL. Cold ethanol was used and AE elution buffer was heated to 70°C prior to addition to the spin-column membrane. Elution of DNA from the spin-column membrane was carried out in three stages, with the addition of 60 µl of AE buffer each time and a 40 minute and two 15 minute incubations at room temperature prior to centrifugation, for a total elution volume of 180 µl.

All scat samples were screened for species identification using species-specific primers amplifying regions of four mitochondrial genes as described in Caragiulo *et al.* (2013) (Table S2-1). DNA amplifications from all extractions were carried out in 25 µl reaction volumes containing 22.3 µl of ultrapure water, 0.7 µl of forward primer (10 µM), 0.7 µl of reverse primer (10 µM), 0.3 µl bovine serum albumin (BSA), one illustra™ puReTaq Ready-To-Go PCR Bead (GE Healthcare, Piscataway, New Jersey, USA), and 1.0 µl of template DNA.

Table S2-1. Four mitochondrial gene regions spanning 1,140 bp were amplified using the primer sets described below. The Carnivorous primers amplify a region nested within the region amplified by the Canideos primers.

Type	Location	Primername	Annealing Temp (˚C)	Origin
mtDNA	Cytochrome b	Carnivorous F: H15149	50	Kocher *et al.* 1989
		Carnivorous R: Farrel-R		Farrell *et al.* 2000
		Canideous F: H15149	52	Kocher *et al.* 1989
		Canideos R: Canid-L1		Paxinos *et al.* 1997
	12S rRNA	L1085	57	Kitano *et al.* 2007
		H1259		Kitano *et al.* 2007
	16S rRNA	16Scp-F	52.5-51.5 (touchdown cycle)	Pomilla *et al*. 2009
		16Scp-R		Pomilla *et al*. 2009
		L2513	57	Kitano *et al*. 2007
		H2714		Kitano *et al.* 2007
	ATPase-6	ATP6-DF3	50	Chaves *et al.* 2012
		ATP6-DR2		Chaves *et al.* 2012

All samples were visualized on a 2.0% agarose gel in TBE buffer and PCR products were purified using the Agencourt AMPure XP PCR purification protocol (Beckman Coulter, Indianapolis, Indiana, USA). Purified PCR products were then sequenced using the Big Dyeterminator protocol (Applied Biosystems, Carlsbad, California, USA). Sequencing reactions were carried out in 8.0 µl reaction volumes containing 4.75 µl ultrapure water, 0.75 µl extension buffer, 0.5 µl Big Dye, 1.0 µl primer (1.6 µM), and 1.0 µl of purified PCR product. The sequencing reaction was carried out separately for the forward and reverse primers. The thermocycler profile for all sequencing reactions followed that of Platt *et al.* (2007). Sequencing amplifications were purified using the Agencourt CleanSEQ Dye Terminator Removal protocol (Beckman Coulter, Indianapolis, Indiana, USA) and sequenced in an ABI 3730xl DNA Analyzer (Applied Biosystems, Carlsbad, California, USA). Sequences were manually edited using Sequencher (version 5.0, Gene Codes Corporation, Ann Arbor, Michigan, USA) and Geneious (Drummond *et al.*, 2012), and compared to both an in-house database of carnivore mtDNA sequences and the NCBI nucleotide BLAST database to confirm species identification. All samples identified as lion were used in further analyses.

The 12 microsatellite loci used in this study (FCA032, FCA075, FCA096, FCA100,FCA124, FCA126, FCA132, FCA208, FCA212, FCA225, FCA229, FCA275) were adapted from a genetic map of the domestic cat (Menotti-Raymond *et al.*, 1999) and optimized for lions. PCR reactions were carried out in 20.0 µl multiplex reactions containing 5 µl of extracted DNA, 0.20 – 1.60 µl of each forward and reverse 10 µM primer, 10.0 µl Qiagen Mutliplex PCR Master Mix, 2.0 µl Q-solution, and the remaining volume was RNAse-Free water (Qiagen, Valencia, California, USA). Primers were grouped into five multiplex reactions based upon fluorescent tag and amplicon size (Table S2-2). Thermocycling conditions were the same for each multiplex, except for the touchdown and annealing temperature, and were as follows: 95 °C for 15 minutes, 13 cycles of 94 °C for 30 seconds, touchdown annealing temperature for 1.5 minutes, and 72 °C for 1 minute, followed by 32 cycles of 94 °C for 30 seconds, annealing temperature for 1.5 minutes, and 72 °C for 1 minute, followed by 30 minutes at the annealing temperature (Table S2-2). Samples were prepared for analysis by mixing 1 µl of PCR product with 9 µl of an 8.82 µl: 0.18 µl mixture of Hi-Di formamide: GeneScan 500 LIZ size standard (Applied Biosystems, Carlsbad, California, USA). Samples were heat-shocked for 3 minutes at 95 °C and genotypes were analyzed using an ABI 3730xl DNA analyzer (Applied Biosystems, Carlsbad, California, USA). Genotypes were scored with GeneMapper v. 4.0 software (Applied Biosystems, Carlsbad, California, USA) and individually verified by visual inspection.

Table S2-2. FCA primers were grouped into five multiplexes and thermocycling conditions were optimized for each multiplex.

Group	Included loci	Touchdown Cycle Annealing Temperature (˚C)	Annealing Temp (˚C)
Multiplex 1	FCA032	60.4 - 0.3	58
	FCA100		
	FCA124		
Multiplex 2	FCA126	62.4 - 0.3	60
	FCA212		
	FCA229		
Multiplex3	FCA096	59.4 - 0.3	57
	FCA132		
	FCA275		
Multiplex 4	FCA075	59.4 - 0.3	57
	FCA208		
Multiplex 5	FCA225	57.4 - 0.3	55

All microsatellite amplifications were performed at least four times using the multi-tube approach (Taberlet *et al.*, 1996) to identify possible allelic dropout. Allelic dropout and PCR success was quantified per locus using GIMLET (Valière, 2002). Consensus genotypes were defined for each sample by comparing results from both a consensus genotype inference method using GIMLET (Valière, 2002) and manual inspection. All samples that did not produce reliable consensus genotypes for at least 6 loci were excluded from further analyses.

References

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. & Wheeler, D.L. 2005 GenBank. *Nucleic Acids Research*, **33**, D34-D38.

Caragiulo, A., Dias-Freedman, I., Clark, J.A., Rabinowitz, S. & Amato, G. 2013 Mitochondrial DNA sequence variation and phylogeography of Neotropic pumas (Puma concolor). *Mitochondrial DNA*, **0**, 1-9.

Chaves P.B., Graeff V.G., Lion M.B., Oliveira L.R., Eizirik E. 2012 DNA barcoding meets molecular scatology: Short mtDNA sequences for standardized species assignment of carnivore noninvasive samples. *Molecular Ecology Resources*, **12**, 18–35.

Drummond, A., Ashton, B., Buxton, S., Cheung, M., Cooper, A., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Sturrock, S., Thierer, T. & Wilson, A. 2012 Geneious v5.6.

Farrell, L., Roman, J. & Sunquist, M.E. 2000 Dietary separation of sympatric carnivores identified by molecular analysis of scats. *Molecular Ecology*, **9**, 1583-1590.

Kitano T., Umetsu K., Tian W., Osawa M. 2007 Two universal primer sets for species identification among vertebrates. *International Journal of Legal Medicine*, **121**, 423–7.

Kocher, T.D., Thomas W.K., Meyer A. *et al.* 1989 Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences USA*, **86**, 6196-6200.

Menotti-Raymond, M., David, V., Lyons, L., Schäffer, A., Tomlin, J., Hutton, M. & O'Brien, S. 1999 A genetic linkage map of microsatellites in the domestic cat (*Felis catus*). *Genomics*, **57**, 9-23.

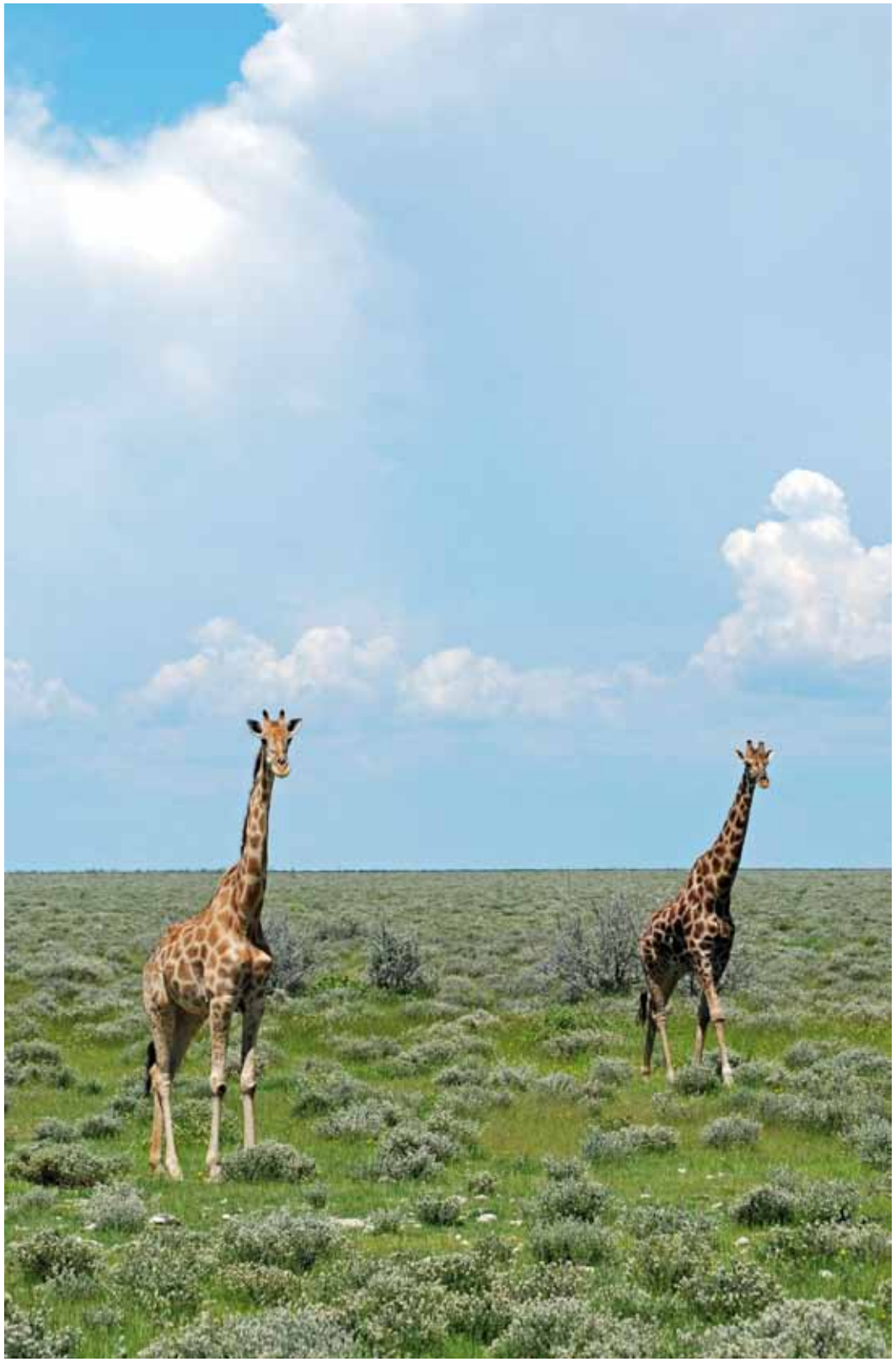
Paxinos E., McIntosh C., Ralls K. & Fleischer R. 1997 A noninvasive method for distinguishing among canid species: amplification and enzyme restriction of DNA from dung. *Molecular Ecology*, **6**, 483-486.

Pomilla C., Harmsen B.J., Foster R.J., Devlin A., Bartholomew C., Vynne C., Carillo-Percastegui S., *et al.* 2009 Large and fine scale genetic characterization of jaguars (Panthera onca) throughout their range. Oral Presentation at the 23rd International Congress for Conservation Biology, July, Beijing, China.

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L.P. & Bouvet, J. 1996 Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research*, **24**, 3189-3194.

Valière, N. 2002 GIMLET: a computer program for analysing genetic individual identification data. *Molecular Ecology Notes*, **2**, 377-379.

Wei, K., Zhang, Z., Zhang, W., Xu, X., Liang, X., He, G., Shen, F., Zhang, L., Hou, R. & Yue, B. 2008 PCR-CTPP: a rapid and reliable genotyping technique based on ZFX/ZFY alleles for sex identification of tiger (*Panthera tigris*) and four other endangered felids. *Conservation Genetics*, **9**, 225-228.



4 Chapter

Phylogeographic patterns in Africa and high resolution delineation of genetic clades in the African lion

(under review)

L. D. Bertola, H. Jongbloed, K.J. van der Gaag, P. de Knijff, N. Yamaguchi, H. Hooghiemstra,
H. Bauer, P. Henschel, P.A. White, C.A. Driscoll, T. Tende, U. Ottosson, Y. Saidu,
K. Vrieling and H. H. de Iongh

Abstract

Numerous African savannah mammals show a congruent phylogenetic pattern in which populations in West/Central Africa are distinct from populations in East/Southern Africa. However, for the lion, all African populations are currently classified as a single subspecies (*Panthera leo leo*), while the only remaining lion population in Asia is considered to be distinct (*Panthera leo persica*). In this study, we assess the phylogeography of the lion, by analysing mitochondrial DNA data of populations throughout the complete geographic range of the lion. This reveals six supported clades and a strongly supported ancestral dichotomy with populations from the northern part of the range (West Africa, Central Africa, North Africa/Asia) on one branch, and populations from the southern part of the range (North East Africa, East/Southern Africa and South West Africa) on the other. This phylogeography is congruent with patterns found in other savannah mammals and is addressed in relation to large scale environmental changes in Africa, driven by climate. The degree of divergence and the nested position of the Asiatic subspecies strongly support the revision of current lion taxonomy, and we propose to recognize a northern and a southern subspecies as this is more in line with the evolutionary history of the lion.

Keywords: phylogeography, Africa, lion (*Panthera leo*), mitochondrial genome, African climate history, savannah mammals

Introduction

Insight into the genetic lineages in a species is of importance, because phylogeographies contribute to the understanding of evolutionary histories as well as to the design of effective conservation strategies. On the African continent we observe strongly congruent phylogenetic patterns for savannah mammal species with a comparable, continent-wide distribution. The distribution of subspecies, and species within species complexes, tends to follow a north-south axis in sub-Saharan Africa, in which West and Central Africa is inhabited by other taxonomic groups than East and Southern Africa (regions defined following the lion conservation strategies (IUCN SSC Cat Specialist Group 2006a; b)) (Table 1 and Figure 1). This north-south dichotomy was further confirmed by genetic data from primates, elephant, rhinoceros, numerous other ungulates and carnivores (see Table 1 for accompanying references). For most species mentioned distinct subspecies are recognized, but not for the African lion (*Panthera leo leo*), which is reflected in the conservation strategy of the species (Bauer *et al.* 2012).

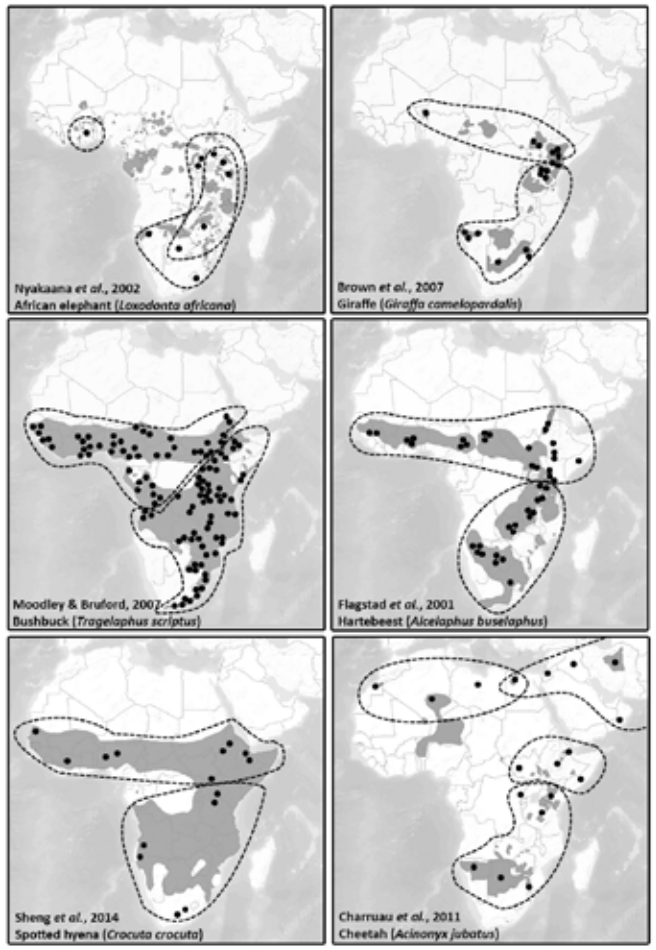


Figure 1. Examples from six species for which a dichotomy between West/Central African populations and populations in East/Southern Africa has been shown in phylogenetic data: African elephant (*Loxodonta africana*), giraffe (*Giraffa camelopardalis*), bushbuck (*Tragelaphus scriptus*), hartebeest (*Alcelaphus buselaphus*), spotted hyena (*Crocuta crocuta*) and cheetah (*Acinonyx jubatus*). Sample locations (black dots) are indications and are not necessarily proportional to the number of collected samples. The most basal phylogenetic groups identified are delineated. Range data from IUCN (IUCN 2014).

Table 1. Overview of African mammals for which a distinction between West/Central African populations and populations in East/Southern Africa has been described (Kingdon 2007; IUCN 2014).

Taxonomic distinction between West/Central Africa and East/Southern Africa				
Order	**Species (complex)**	**(sub)Species**	**Phylogeography references***	**Genetic marker**
Primates	Babboon complex (*Papio)*	5 species	Zinner et al. (2009)	mtDNA
	Green monkey complex (*Chlorocebus*)	6 species	Haus et al. (2013)	mtDNA
	Senegal galago (*Galago senegalensis*)	4 subspecies	-	-
Hyracoidea	Rock hyrax complex (*Procavia*)	5 species	-	-
Perissodactyla	Black rhino (*Diceros bicornis*)	4 subspecies	Harley et al. (2005)	msats
	White rhino (*Ceratotherium simum*)	2 subspecies	-	-
Artiodactyla	Giraffe (*Giraffa camelopardalis*)	9 subspecies	Brown et al. (2007); Hassanin et al. (2007); Bock et al. (2014)	mtDNA + msats
	African buffalo (*Syncerus caffer*)	3-4 subspecies	Van Hooft et al. (2002); Smitz et al. (2013)	mtDNA + Y chromosomal msats
	Bushbuck (*Tragelaphus scriptus*)	2 groups, numerous subspecies	Moodley & Bruford (2007)	mtDNA
	Greater kudu (*Tragelaphus strepsiceros*)	3 subspecies	-	-
	Eland complex (*Tragelaphys debianus / T. oryx*)	2 species	-	-
	Bush duiker (*Sylvicapra grimmia*)	8 groups, numerous subspecies	-	-
	Dwarf antelope complex *Neotragus*	3 species	-	-
	Oribi (*Ourebia ourebi*)	7-13 subspecies	-	-
	Reedbuck complex (*Redunca redunca / R. arundinum*)	2 species	-	-
	Mountain reedbuck (*Redunca fulvorufula*)	3 subspecies	-	-
	Kob / Puku complex (*Kobus kob / K. vardoni*)	2 species	Lorenzen et al. (2007)	mtDNA + msats
	Lechwe complex (*Kobus leche / K.megaceros*)	2 species	-	-
	Waterbuck (*Kobus ellipsiprymnus*)	2 subspecies	Lorenzen et al. (2006)	mtDNA + msats
	Red-fronted gazelle (*Eudorcas rufifrons*)	5 subspecies	-	-
	Grant's gazelle complex (*Nanger*)	3 species	-	-
	Topi (*Damaliscus lunatus*)	5-6 subspecies	-	-
	Hartebeest (*Alcelaphus buselaphusi*)	8 subspecies	Arctander et al. (1999); Flagstad et al. (2001)	mtDNA
	Roan antelope (*Hippotragus equinus*)	2 groups, 6 subspecies	Alpers et al. (2004); Matthee & Robinson (1999)	mtDNA + msats
	Oryx complex (*Oryx*)	3 species	-	-
Carnivora	Egyptian mongoose (*Herpestes ichneumon*)	up to 11 subspecies	Gaubert et al. (2011)	mtDNA
	Slender mongoose (*Herpestes sanguineus*)	up to 50 subspecies	-	-
	White-tailed mongoose (*Ichneumia albicauda*)	6 subspecies	Dehghani et al. (2008)	mtDNA
	Common genet (Genetta genetta)	3 groups, numerous subspecies	Gaubert et al. (2011); Delibes & Gaubert (unpub.)	mtDNA
	African civet (*Civettictis civetta*)	5 subspecies	-	-
	Wild cat (*Felis silvestris*)	5 subspecies	Driscoll et al. (2007)	mtDNA + msats
	Caracal (*Caracal caracal*)	8 subspecies	-	-
	Cheetah (*Acinonyx jubatusi*)	5 infra-specific taxa assessed	Freeman et al. (2001); Charruau et al. (2011)	mtDNA + msats
No taxonomic distinction				
Order	**Species (complex)**	**(sub)Species**	**Phylogeography references***	**Genetic marker**
Proboscidea	African (bush) elephant (*Loxodonta africana*)	-	Nyakaana et al. (2002)	mtDNA + msats
Pholidota	Ground pangolin (*Manis temmicnkii)*	-	-	-
Tubulidentata	Aardvark (*Orycteropus afer*)	-	-	-
Artiodactyla	Common warthog (*Phacochoerus africanus*)	-	Muwanika et al. (2003)	mtDNA
Carnivora	African wild dog (*Lycaon pictus*)	-	-	-
	Zorilla (*Ictonyx striatus*)	-	-	-
	Honey badger (*Mellivora capensis*)	-	-	-
	Banded mongoose (*Mungos mungo*)	-	-	-
	Marsh mongoose (*Atilax paludinosus*)	-	-	-
	Spotted hyena (*Crocuta crocuta*)	-	Rohland et al. (2005), Sheng et al. (2014)	mtDNA
	Serval (*Leptailurus serval*)	-	-	-
	African Leopard (*Panthera pardus pardus*)	(one subspecies in Africa)	-	-
	African Lion (*Panthera leo leo*)	(one subspecies in Africa)	Dubach et al. (2013); Barnett et al. (2014); Bertola et al. (2011); Bertola et al. (submitted)	mtDNA + msats

* Only references that cover the complete (sub)species's range on the African continent are listed. Publications focussing on a more regional level were excluded.

Phylogenetic data of lion populations indicate that current taxonomy does not sufficiently reflect the genetic diversity within the African lion (Dubach *et al.* 2005, 2013; Barnett *et al.* 2006a; b, 2014; Antunes *et al.* 2008; Bertola *et al.* 2011; Bruche *et al.* 2012). Notably, lion populations from West and Central Africa have a distinct phylogenetic position, with a nested position for the Asiatic subspecies (*Panthera leo persica*) (Barnett *et al.* 2006a; b, 2014; Bertola *et al.* 2011). The validity of the subspecies status of the Asiatic lion, nowadays confined to a single population in India, is thereby challenged. However, previous studies describing distinct genetic lineages within the African lion did not thoroughly cover the West and Central African region and based there results on relatively small sample sizes (Dubach *et al.* 2005, 2013; Barnett *et al.* 2006a; b, 2014; Bertola *et al.* 2011). The position of the these populations and their relation to the Asiatic subspecies in the phylogenetic tree remained largely unresolved (Barnett *et al.* 2006b, 2014; Bertola *et al.* 2011).

The present study aims to provide a more complete overview of genetic diversity within the African lion and compares this to phylogeographic patterns and taxonomy in a range of African savannah mammals. In addition, we estimate the dates of the major splits in the phylogenetic tree and aim to relate observed patterns to the dynamic climate history of Africa. Previous studies have shown that mitochondrial DNA (mtDNA) loci produce phylogenies that are not contradictory with phylogenies based on autosomal data (Antunes *et al.* 2008; Bertola *et al.* submitted), which indicates that mtDNA is an appropriate marker for making this case. For sixteen lions the complete mitochondrial genome was analyzed. In addition, 1454 base pairs (bp) of the mtDNA were analyzed for 178 lions throughout their complete geographic range (see Figure 2 and Supplemental Table S1 for sampling locations), including samples from each of the Lion Conservation Unit (LCU) in West and Central Africa that still contains a recently confirmed resident lion populations (Riggio *et al.* 2012; Henschel *et al.* 2014). To reconstruct the evolutionary history of the West and Central African lion, museum samples from extinct populations in North Africa and Asia, representing a historical connection between the African and the Asiatic subspecies, were obtained and an ancient DNA (aDNA) approach was used for processing. These museum samples also included lions from areas from which it was not possible to include modern samples of wild lions.

This is the first study in which a number of approaches, including Next Generation Sequencing (NGS) techniques, are applied on a dataset of 194 lions from 22 different countries throughout the complete geographic range and most of the historic range of the modern lion. These data contribute to a better understanding of evolutionary forces that shaped the phylogenetic patterns observed among numerous savannah mammals on the African continent. Results should be translated into recommendations for the management of diverse species and populations. In particular, we challenge current lion taxonomy that recognizes only the African and the Asiatic subspecies, and we investigate options for a revision that is more parsimonious with the recently improved understanding of the evolutionary history of the lion.

Materials and Methods

In total, 194 samples from lions of 22 different countries were analyzed, including samples previously described in Bertola *et al.* (2011), Barnett *et al.* (2014) and Bertola *et al.* (submitted) (Supplemental Table S1 and Figure 2 for samples locations). Blood, tissue or scat samples were collected from free-ranging individuals or captive lions with proper documentation of their breeding history. A total of 16 museum specimens, collection dates ranging from 1831 to 1967, was added to the dataset. Maxilloturbinal bone was sampled, unless another sample was more readily available. Samples were collected in full compliance with specific legally required permits (CITES and permits related to national legislation in the countries of origin). For details on sample storage and processing, see Supplemental Information S1, Supplemental Tables S2 and S3.

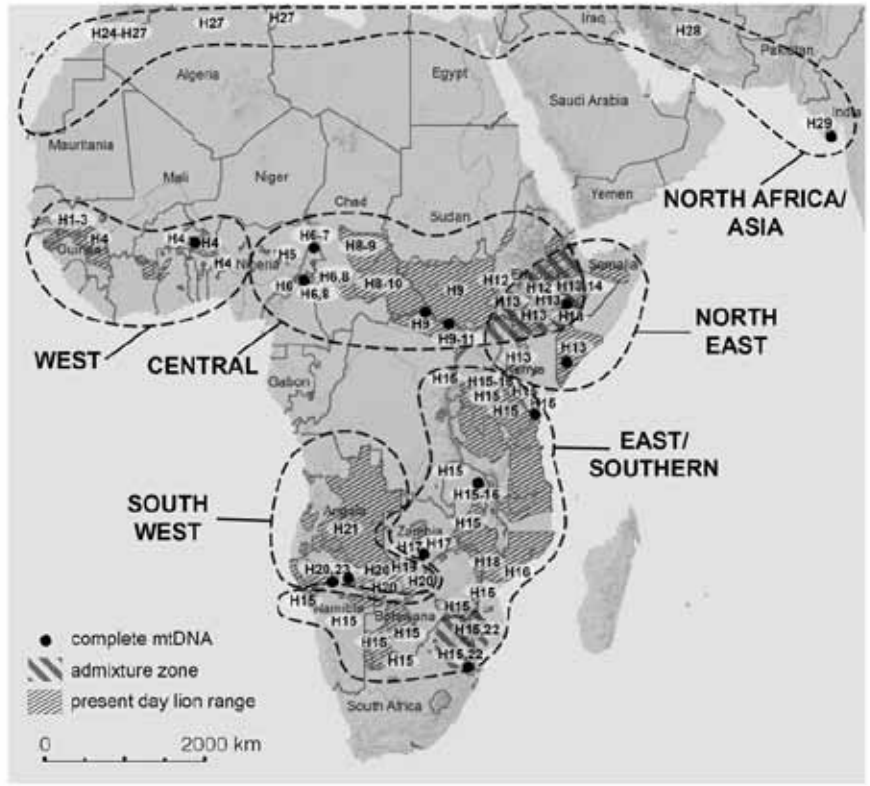


Figure 2. Locations of lion samples and haplotype numbers included in this study. Proposed phylogenetic lineages are delineated. Admixture zones in which haplotypes from different phylogenetic lineages are found are indicated by shading. Lion range data from IUCN (2014).

For all available samples, analyses were performed on alignments consisting of cytochrome b, tRNAThr, tRNAPro and the left domain of the control region (hereafter referred to as cytB+ctrl reg.) (1454 bp, 202 sequences), the complete mitogenome (16756 bp, excluding RS-2 and RS-3, 23 sequences) and an alignment including all sequence data, where ambiguous nucleotides were added to create sequences of equal length. Bayesian analysis was performed using MrBayes v.3.1.2

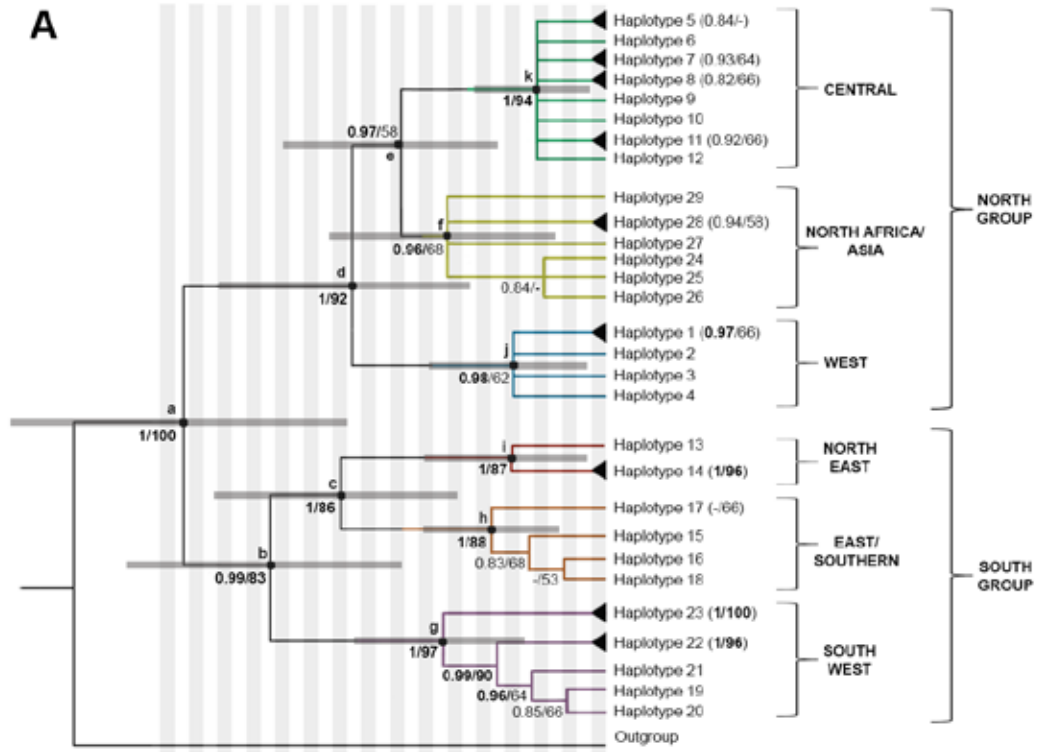
(Huelsenbeck & Ronquist 2001; Ronquist *et al.* 2012), using parameters determined by MrModeltest2 (v.2.3) (Nylander 2004) and Maximum Likelihood (ML) analysis was done in Garli (Zwickl 2006). Branches receiving >0.95 PP in Bayesian analysis and/or 70 bootstrap support in ML analysis are considered to be significantly supported. A haplotype network was created using the median-joining algorithm in Network 4.6.1.1 (available from www.fluxus-engineering.com) with equal weighing of all characters.

BEAST v.1.7.5 (Drummond & Rambaut 2007) was used to obtain estimated values for the time to most recent common ancestor (TMRCA) to date splits in the lion tree. Five independent runs of 50 million iterations were performed, discarding the first ten percent of each run as burnin, and using the same model as was used for Bayesian analysis and relaxed molecular clock setting. Fossil evidence for the origin of *Panthera*, the *P.leo-P.pardus* group and *P.leo* (including *P. leo spelaea*) was used for calibration and set to 3.8 million years before present (Ma), 1.6 Ma and 0.55 Ma respectively (Kurten & Anderson 1980; Janczewski *et al.* 1995; Burger *et al.* 2004; Johnson *et al.* 2006; Davis *et al.* 2010). Convergence of the runs was assessed in Tracer. Logcombiner, Treeannotator and Figtree (available from http://tree.bio.ed.ac.uk/software/figtree/) were used to visualize the results.

Results

Bayesian and ML trees were constructed from three different alignments: 1) cytB+ctrl reg., 2) the complete mitogenome, and 3) a combination of both datasets. All showed identical topology and trees including the complete mtDNA showed strongly significant support for a basal split separating lions in the northern part of their range (North group: West Africa, Central Africa, and North Africa/ Asia) and lions in the southern part of their range (South group: North East, East/Southern, and South West Africa) (Figure 3A). Within the North group, a clade that included all Asiatic lions and aDNA sequences from North Africa and Iran was significantly supported, as was a clade with Central African lions and the clade with West African lions. Lions from Central Africa and the North Africa/ Asia clade are grouped together on a well-supported branch. In the South group, three major groups can be distinguished: a South West group, an East/Southern group and a North East group. All three clades are significantly supported, as is the branch combining the East/Southern and North East group. The same structure can be seen in the haplotype network based on cytB+ctrl reg. (Figure 4). The observed groups are indicated together with the sample location in Figure 2. Only in two cases did we observe haplotypes from distinct phylogenetic groups in one geographic region: in Ethiopia we found haplotypes from the Central Africa group as well as from the North East group, and in the Republic of South Africa (RSA) we found haplotypes from East/Southern and the South West group (Figure 2, shaded areas).

The most recent common ancestor of all modern lions was estimated to have existed around 291.7 thousand years ago (291.7 ka) (95% Highest Posterior Density (HPD): 178.0-417.7 ka). The split of the South group is older than the North group, estimated to be 231.3 ka (95% HPD: 132.3-338.7 ka) and 174.7 ka (95% HPD: 94.9-276.7 ka), respectively. For all major clades the date of the most recent common ancestor was estimated and compared to results from previous publications (Supplemental Table S4, Figure 3A+B).



B

Clade	Node	Date (95% HPD)	Clade	Node	Date (95% HPD)	Clade	Node	Date (95% HPD)	Clade	Node	Date (95% HPD)
Modern Lion	a	291.7 ka (178.0-417.7)	North	d	174.7 ka (94.9-276.7)	South West Africa	g	113.8 ka (55.2-189.0)	West Africa	j	63.4 ka (15.1-129.1)
South	b	231.3 ka (132.3-338.7)	Central Africa - India	e	141.1 ka (71.7 - 216.6)	East/ Southern Africa	h	78.1 ka (37.2-132.1)	Central Africa	k	49.6 ka (20.7-91.0)
East/Southern - North East Africa	c	183.7 ka (107.0-271.1)	India - North Africa	f	110.8 ka (39.7-200.1)	North East Africa	i	63.9 ka (18.0-118.6)			

Figure 3. Phylogenetic analyses for the complete lion dataset, including sixteen mitochondrial genomes and 175 cytb+ctrl reg. sequences. A: Phylogenetic tree of lion populations throughout their complete geographic range, based on complete mitochondrial genomes and cytB+ctrl reg. sequences. Branch colours correspond to haplotype colours in Figure 4. Support is indicated as posterior probability (Bayesian analysis)/bootstrap support (ML analysis). Branches with a single haplotype have been collapsed to improve readability. Support for these branches is indicated by a black triangle at the tip of the branch (support shown in the label). Nodes which have been included for divergence time estimates are indicated with letters and 95% HPD node bars. Distance to outgroup and nodes without dated split is not in proportion to divergence time. B: divergence estimates and 95% HPD from BEAST analysis, also indicated as error bars in Figure 3A.

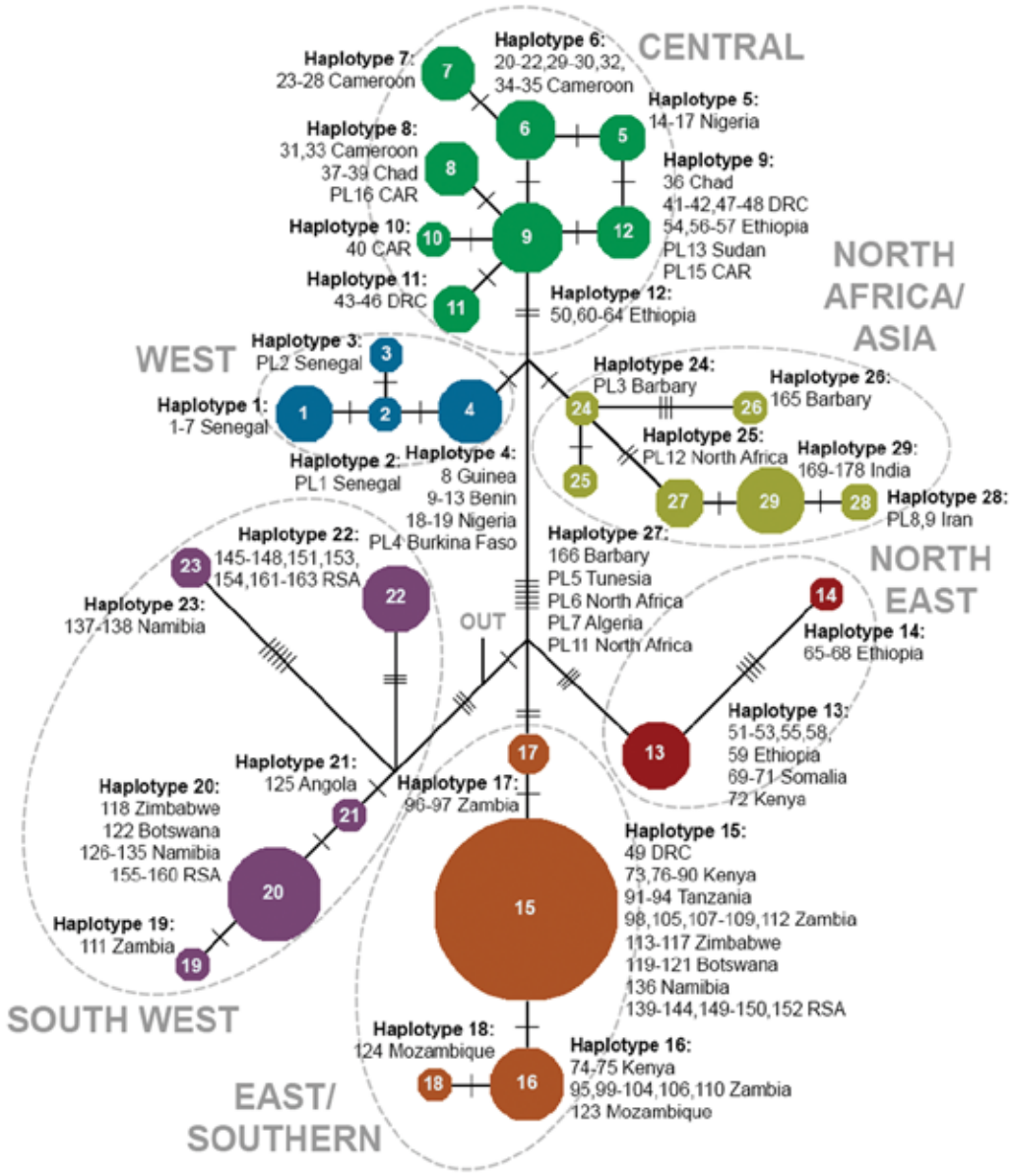


Figure 4. Haplotype network based on cytB+ctrl reg. sequences of lions throughout their entire geographic range. Dashed lines indicate the groups discerned by Bayesian/ML analysis in Figure 3. Haplotype size is proportional to its frequency in the dataset. Hatch marks represent a change in the DNA sequence. The connection to outgroup species is indicated by "OUT".

Discussion

In this paper, we describe the phylogenetic relationships of lion populations throughout their entire geographic range based on 194 sequences of cytB+ctrl reg., including 30 aDNA sequences, and 16 sequences of the complete mitogenome. This has led to strong support for a basal dichotomy between lion populations from the northern part of their range and those from the southern part. Six major phylogenetic groups are identified: West Africa, Central Africa and North Africa/Asia (North group) and North East, East/Southern and South West (South group).

This study included samples from 22 lion range states, including all LCUs with a confirmed lion population in West and Central Africa, and including extinct populations, covering a major part of the historical geographic range of the modern lion. Our results show that lion populations that were previously described as unique, as was the case for the Addis Ababa lions (Bruche *et al.* 2012) and for the Sabi Sands lions (Dubach *et al.* 2005) are most likely the result of incomplete sampling. Angola is represented by one aDNA sample only, which clusters to the South West group. Although it is difficult to draw conclusions for the entire Angolan lion population, this suggests that the captive Angolan lions that were included in previous phylogenetic studies (Antunes *et al.* 2008; Bertola *et al.* 2011) are not pure-bred Angolan. Pedigree information also shows that there is no complete documentation of the female lineage in this captive population (Steinmetz *et al.* 2006). Samples from zoos and museums were only included in our study when decisive information was available on the origin of the individual or its free-ranging ancestors (for additional information see Supplemental Information S2).

Based on the available datapoints, a proposed range of the haplogroups is shown in Figure 2. Two areas of admixture between distinct lineages are indicated by shading. Although the Rift Valley has been proposed as a barrier for gene flow in lions (Burger *et al.* 2004; Dubach *et al.* 2005; Barnett *et al.* 2006a; b, 2009; Bertola *et al.* 2011), our denser sampling of the connecting region between the North and South groups shows that it does not completely prevent a mixture of haplotypes from the two basal branches in the phylogenetic tree (haplotype 9, 12-14). The second admixture zone is located around Kruger National Park (NP) and Limpopo-Venetia National Reserve (NR), RSA, in which we detect haplotypes from the South West group (haplotype 20 and 22) in addition to haplotypes from the East/Southern group (haplotype 15). Since lions from other parts of RSA and the Southern range of Botswana and Namibia also cluster to the East/Southern group, it is likely that the mixture of haplotypes in the Kruger/Limpopo area is the result of human-induced translocations. Lions from Etosha have frequently been used in translocations and it is known that private reserves adjacent to Kruger NP, that were initially fenced off, are now connected to the park (Miller *et al.* 2013).

The pattern we describe for the lion is highly congruent with phylogeographic data from different taxonomic groups on a range of trophic levels, indicating environmentally driven evolution. Several phylogeographic studies on African savannah mammals have described three main clades: West/Central Africa, East Africa and Southern Africa, suggesting that there may have been major refugial areas in these regions during the more recent part of the Pleistocene climatic cycles (Hewitt 2004; Lorenzen *et al.* 2012). These three clades are clearly distinguishable in the lion based on mtDNA (nodes c, d and g in Figure 3) and autosomal data (Bertola *et al.*, submitted). A model-based study on the habitat suitability for mammals and birds during the last glacial maximum (LGM) suggests that there were five possible refugia sub-Saharan Africa: one in Upper Guinea, one or two in the

Cameroon Highlands – Congo Basin, one in the Ethiopian Highlands, one in Angola-Namibia, and one in East/Southern Africa (Levinsky *et al.* 2013). These areas represent the five sub-Saharan lion groups, described in this study, being West, Central, North East, South West, and East/Southern, respectively. In addition, we find corroboration in phylogeographic patterns from other savannah mammals with a distribution similar to that of the lion. Apart from the most basal dichotomy, shown for other species in Table 1 and Figure 1, the South West clade, which harbors lion populations from Angola and Namibia, is also recognized in giraffe (*Giraffa camelopardalis*) (Brown *et al.* 2007), zebra (*Equus zebra*) (Moodley & Harley 2006), impala (*Aepyceros melampus*) (Nersting & Arctander 2001), greater kudu (*Tragelaphus strepsiceros*) (Nersting & Arctander 2001) and sable antelope (*Hippotragus niger*) (Pitra *et al.* 2002). Within East Africa, the North East clade is also found in kob (*Kobus kob*) (Lorenzen *et al.* 2006), oryx (*Oryx beisa*) (Masembe *et al.* 2006), impala (*Aepyceros melampus*) (Nersting & Arctander 2001) and greater kudu (*Tragelaphus strepsiceros*) (Nersting & Arctander 2001). Finally, the distinction we find between the West and the Central African lion is also recognized in the phylogeographic pattern of roan antelope (*Hippotragus equinus*) (Alpers *et al.* 2004), potentially resulting from the lower Niger River as a permanent barrier for gene flow. Climatological events have also heavily influenced migration of early humans (Castañeda *et al.* 2009; Blome *et al.* 2012) and as a result, similar major clades and phylogeographic patterns are found in human datasets (Templeton 2002; Gonder *et al.* 2007; Tishkoff *et al.* 2009).

Phylogenetic variation within the six geographic groups of the modern lion appears to have mainly emerged within the last c. 100,000 year (100 kyr), including the cool last glacial (Marine Oxygen Isotope Stage (MIS) 4, 3 and 2) and two warmer periods (MIS 5 and 1) (Carto *et al.* 2009; Cronin 2010). Phylogenetic structure which had evolved in regional lineages during the previous glacial-interglacial cycles, mostly disappeared by c. 100 ka through various events, including genetic bottlenecks involving expansions and contractions from/to regional refugia (Migliore *et al.* 2013; Levinsky *et al.* 2013; Dauby *et al.* 2014). Since the HPD intervals are relatively large, we add a palaeoclimatic context to be able to propose a possible scenario that has contributed to the current phylogeographic pattern. The two major vegetation zones that likely influenced lion distribution through exclusion on the African continent are dry desert and dense rain forests (Nowell & Jackson 1996; Yamaguchi *et al.* 2004), both reflecting hydrological extremes. In the tropics the hydrological cycle is mainly driven by 21 kyr precession cycle of orbital climate forcing which is somewhat independent of the interglacial-glacial variations (Clement *et al.* 2004; Cronin 2010). The last coalescence between the North and South lineage (node a in Figure 3) in the lion is estimated at ~292 ka, positioned after the first cold interval (MIS 8.6 at 299 ka) of glacial MIS 8 (303-245 ka) (ages after (Imbrie *et al.* 1984)). This period is characterized by a maximum monsoon index, allowing dense wet forest to expand maximally northwards along an east-west axis in lower latitude Africa (Dupont & Hooghiemstra 1989; De Vivo & Carmignotto 2004; Kingdon 2007; Staver *et al.* 2011; Lehmann *et al.* 2011; Hardy *et al.* 2013; Dauby *et al.* 2014). Such vegetation pattern likely reduced or possibly eliminated the connection between northern and southern populations. The second oldest split between South West group and East/Southern & North East groups (node b in Figure 3) occurred at around ~231 ka, a moment positioned close to the first cool interval (MIS 7.4 at 228 ka) of interglacial MIS 7 (245-186 ka). The monsoon index was still high (Dupont & Hooghiemstra 1989) and the belt with rain forest may also in this interval have prevented a connection between lion populations, while simultaneously individuals belonging to the East/Southern group are distributed in a large area across East and Southern Africa (De Vivo & Carmignotto 2004; Barnett *et al.* 2006b). More recent radiation of the South West group (node g in Figure 3), estimated to have occurred ~114 ka, coinciding with the last part of the Eemian interglacial (MIS 5.5.), following a period of droughts, notably in the Kalahari region, in which suitable habitat was reduced in Southern Africa (Dupont 2011). The splits between East/Southern and North East Africa (node c in Figure 3), West Africa and Central & North Africa/Asia (node d in Figure 3), and Central and North Africa/Asia (node e in Figure 3) all appear to have occurred during MIS6 (186-128 ka) when relatively dry and cool conditions prevailed (Dupont & Hooghiemstra 1989; Petit *et al.* 1999). The splits in the North group are likely due to the periodically maximum north-south extension of the Sahara desert (Dupont & Hooghiemstra 1989; Hooghiemstra *et al.* 1992; Andel & Tzedakis 1996; Hoelzmann *et al.* 2004; De Vivo & Carmignotto 2004; Barnett *et al.* 2006b; Migliore *et al.* 2013). A connection between the North Africa/Asia group and the Central Africa group may have persisted during short periods that the monsoon front reached high latitudes, explaining the close genetic relationship to the North Africa/Asia clade (Dupont & Hooghiemstra 1989; Hooghiemstra *et al.* 1992; Hoelzmann *et al.* 2004). The West African population possibly became isolated and reduced in numbers by the significant southwards expansion of the Sahara during MIS 4 (71-59 ka) (Dupont & Hooghiemstra 1989; Hooghiemstra *et al.* 1992; Hoelzmann *et al.* 2004; Castañeda *et al.* 2009; Dupont 2011), and started radiating around 63 ka. There are no indications from our data that the current lion population in India was sourced or reinforced by introductions from sub-Saharan African lions, as was recently hypothesized (Thapar *et al.* 2013).

The deep ancestral split within the African lion and the topology of the phylogenetic tree, along with the nested position for the Asiatic subspecies, clearly illustrate and support the notion that the current taxonomic division does not reflect the evolutionary history of the lion. Consequentially, it hampers proper priority setting for lion conservation, particularly in West and Central Africa. Since the distinct genetic lineages within the African lion are further supported by nuclear data (Antunes *et al.* 2008; Bertola *et al.* submitted) and morphological data (Hemmer 1974; Mazák 2010), we suggest to recognize a northern subspecies, including West Africa, Central Africa and North Africa/Asia, and a southern subspecies, including the lineages North East, East/Southern and South West, in line with the proposed revision by Barnett *et al.* (2014). Within these two subspecies, the distinct phylogeographic clades should be managed as Evolutionary Significant Units (ESUs), sensu Moritz (1994) (Moritz 1994), in the absence of conflicting conclusions based on other genetic markers. Data from more nuclear loci, and from sampling locations at the geographical borders of the proposed haplogroup ranges may provide a better insight, but are not likely to change the main pattern as is described in this paper.

Our study shows a fine-scale phylogeographic pattern for the lion, with strongly significant support for a basal north-south dichotomy, as is also observed in other African savannah mammals. By analysing samples from more localities, the phylogenetic position of the Asiatic subspecies was resolved and it was possible to propose ranges and connectivity zones for six major phylogenetic clades: West Africa, Central Africa and North Africa/Asia (North group) and North East, East/Southern and South West (South group). In context with the timing of the nodes in the phylogenetic tree, our results contribute to understanding the evolutionary forces that shaped the genetic makeup of several African savannah mammals and the extant lion clades in particular. Current nomenclature of the lion, recognizing an African and an Asiatic subspecies, is not in line with the evolutionary history of the species; we therefore propose a revision of the current taxonomy distinguishing a northern and a southern subspecies.

Acknowledgements

Samples were kindly provided by N. Vanherle (CURESS, Chad), S. Adam and selected members of the West/Central African lion network (ROCAL), E. Sogbohossou (Benin), A. Dunn and Bauchi State Government (Nigeria, Yankari GR), P. Tumenta, S.Adam, R. Buij and B. Croes (Cameroon), ICCN, Garamba NP (DRC), NABU (Ethiopia, Kafa BR), T. Jirma (Kenya, Amboseli NP, Maasai Mara NR and Tsavo East NP) and B. Patterson (Kenya, Aberdare NP and Tsavo NP), O. Aschenborn (Namibia), South Africa National Parks (SANParks) (RSA), S. Miller, R. Groom, Save Valley Conservancy and DeBeers Venetia-Limpopo Nature Reserve/The Diamond Route (RSA and Zimbabwe). We are grateful for the permit of the Ethiopian Wildlife Conservation Authority (formerly Wildlife Conservation Department) to obtain and export blood samples of the lions from Addis Ababa Zoo (Ethiopia1) through Zoo Leipzig, Germany. This is contribution no XX from A.P Leventis Onithological Research Institute (Nigeria, Yankari GR). We thank the following people and institutions for access to samples from captive populations and natural history collections: Diergaarde Blijdorp (Rotterdam, The Netherlands), BornFree Ethiopia (S. Brend and R. Tschopp), Leipzig Zoo (F. Oberwemmer) and IZW Leibnitz Institute for Zoo and Wildlife Research (Addis Ababa Zoo samples), Sanaa Zoo (Yemen), Safaripark Beekse Bergen (Hilvarenbeek, The Netherlands), Dierenpark Amersfoort (Amersfoort, The Netherlands), Zoo Basel (Basel, Switserland), Ouwehands Dierenpark (Rhenen, The Netherlands), Planckendael (Muizen, Belgium), Zoological Museum Amsterdam (Amsterdam, The Netherlands), Swedish Museum of Natural History (Stockholm, Sweden), Smithsonian Institution (Washington, USA), Royal Belgian Institute of Natural Sciences (Brussels, Belgium), Bulawayo museum (Bulawayo, Zimbabwe), Naturalis Biodiversity Centre (Leiden, The Netherlands) and Humboldt museum (Berlin, Germany). We further thank H. Visser, L. Müller, M. van Meenen and M. Schoe for contributing to collecting and/or processing of the samples. The investigations were supported by the Division for Earth and Life Sciences (ALW) with financial aid from the Netherlands Organization for Scientific Research (NWO) (project no. 820.01.002).

References

Alpers DL, Van Vuuren BJ, Arctander P, Robinson TJ (2004) Population genetics of the roan antelope (Hippotragus equinus) with suggestions for conservation. *Molecular ecology*, **13**, 1771–84.

Andel TH van, Tzedakis PC (1996) Palaeolithic landscapes of europe and environs, 150,000-25,000 years ago: an overview. *Quaternary Science Reviews*, **15**, 481–500.

Antunes A, Troyer JL, Roelke ME *et al.* (2008) The evolutionary dynamics of the lion Panthera leo revealed by host and viral population genomics. *PLoS genetics*, **4**, e1000251.

Barnett R, Shapiro B, Barnes I *et al.* (2009) Phylogeography of lions (Panthera leo ssp.) reveals three distinct taxa and a late Pleistocene reduction in genetic diversity. *Molecular ecology*, **18**, 1668–1677.

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). *Proceedings of the Royal Society B: Biological Sciences*, **273**, 2119–2125.

Barnett R, Yamaguchi N, Shapiro B *et al.* (2014) Revealing the maternal demographic history of Panthera leo using ancient DNA and a spatially explicit genealogical analysis. *BMC Evolutionary Biology*, **14**, 70.

Bauer H, Nowell K, Packer C (2012) IUCN Red List of Threatened Species: Panthera leo. *Version 2013.2*.

Bertola LD, van Hooft WF, Vrieling K *et al.* (2011) Genetic diversity, evolutionary history and implications for conservation of the lion (Panthera leo) in West and Central Africa. *Journal of Biogeography*, **38**, 1356–1367.

Blome MW, Cohen AS, Tryon C a., Brooks AS, Russell J (2012) The environmental context for the origins of modern human diversity: A synthesis of regional variability in African climate 150,000–30,000 years ago. *Journal of Human Evolution*, **2012**.

Brown DM, Brenneman R a, Koepfli K-P *et al.* (2007) Extensive population genetic structure in the giraffe. *BMC biology*, **5**, 57.

Bruche S, Gusset M, Lippold S *et al.* (2012) A genetically distinct lion (Panthera leo) population from Ethiopia. *European Journal of Wildlife Research*.

Burger J, Rosendahl W, Loreille O *et al.* (2004) Molecular phylogeny of the extinct cave lion Panthera leo spelaea. *Molecular phylogenetics and evolution*, **30**, 841–849.

Carto SL, Weaver AJ, Hetherington R, Lam Y, Wiebe EC (2009) Out of Africa and into an ice age: on the role of global climate change in the late Pleistocene migration of early modern humans out of Africa. *Journal of human evolution*, **56**, 139–51.

Castañeda IS, Mulitza S, Schefuss E *et al.* (2009) Wet phases in the Sahara/Sahel region and human migration patterns in North Africa. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 20159–63.

Clement AC, Hall A, Broccoli AJ (2004) The importance of precessional signals in the tropical climate. *Climate Dynamics*, **22**, 327–341.

Cronin (2010) *Paleoclimates. Understanding climate change past and present.* Columbia University Press.

Dauby G, Duminil J, Heuertz M *et al.* (2014) Congruent phylogeographical patterns of eight tree species in Atlantic Central Africa provide insights into the past dynamics of forest cover. *Molecular ecology*, **23**, 2299–312.

Davis BW, Li G, Murphy WJ (2010) Supermatrix and species tree methods resolve phylogenetic relationships within the big cats, Panthera (Carnivora: Felidae). *Molecular phylogenetics and evolution*, **56**, 64–76.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology*, **7**, 214.

Dubach JM, Briggs MB, White PA, Ament BA, Patterson BD (2013) Genetic perspectives on "Lion Conservation Units" in Eastern and Southern Africa. *Conservation Genetics*, **1942**.
Dubach J, Patterson BD, Briggs MB *et al.* (2005) Molecular genetic variation across the southern and eastern geographic ranges of the African lion, Panthera leo. *Conservation Genetics*, **6**, 15–24.
Dupont L (2011) Orbital scale vegetation change in Africa. *Quaternary Science Reviews*, **30**, 3589–3602.
Dupont LM, Hooghiemstra H (1989) The Saharan-Sahelian boundary during the Brunhes chron. *Acta Botanica Neerlandica*, **38**, 405–415.
Gonder MK, Mortensen HM, Reed F a, de Sousa A, Tishkoff S a (2007) Whole-mtDNA genome sequence analysis of ancient African lineages. *Molecular biology and evolution*, **24**, 757–68.
Hardy OJ, Born C, Budde K *et al.* (2013) Comparative phylogeography of African rain forest trees: A review of genetic signatures of vegetation history in the Guineo-Congolian region. *Comptes Rendus Geoscience*, **345**, 284–296.
Hemmer H (1974) *Untersuchungen zur Stammesgeschichte der Pantherkatzen (Pantherinae) 3 Zur Artgeschichte des Löwen Panthera (Panthera) leo (Linnaeus 1758).* Veröffentlichungen der Zoologischen Staatssammlung München, München, Germany.
Henschel P, Coad L, Burton C *et al.* (2014) The Lion in West Africa Is Critically Endangered (M Hayward, Ed,). *PLoS ONE*, **9**, e83500.
Hewitt GM (2004) The structure of biodiversity - insights from molecular phylogeography. *Frontiers in zoology*, **1**, 4.
Hoelzmann P, Gasse F, Dupont LM *et al.* (2004) Palaeoenvironmental changes in the arid and subarid belt (Sahara-Sahel-Arabian Peninsula) from 150 kyr to present. In: *Past climate variability through Europe and Africa* (eds Battarbee RW, Gasse F, Stickley CE), pp. 219–256. Springer, Dordrecht, The Netherlands.
Hooghiemstra H, Agwu COC, Lydie MD (1992) Vegetational and climatic changes at the northern fringe of the Sahara 250 , 000-5000 years BP : evidence from 4 marine pollen records located between Portugal and the Canary Islands. *Review of Plaeobotany and Paluology*, **74**, 1–53.
Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics (Oxford, England)*, **17**, 754–5.
IUCN (2014) The IUCN Red List of Threatened Species. Version 2014.1.
IUCN SSC Cat Specialist Group (2006a) *Conservation strategy for the lion in West and Central Africa*. IUCN, Gland, Switzerland.
IUCN SSC Cat Specialist Group (2006b) *Conservation strategy for the lion (Panthera leo) in Eastern and Southern Africa*. IUCN, Gland, Switzerland.
Imbrie J, Hays JD, Martinson DG *et al.* (1984) The orbital theory of Pleistocene climate support from a revised chronology of the maribe ð18O record. In: *Milankovitch and Climate, Part 1* (eds Berger A, Imbrie J, Hays H, Kukla G, Saltzman B), pp. 269–305. Reidel Publishing, Dordrecht, The Netherlands.
Janczewski DN, Modi WS, Stephens JC, O'Brien SJ (1995) Mitochondrial 12S Molecular Evolution of Mitochondrial 12S RNA and and Cytochrome *b* Sequences in Pantherine lineage of Felidae. *Molecular biology and evolution*, **12**, 690–707.
Johnson WE, Eizirik E, Pecon-Slattery J *et al.* (2006) The late Miocene radiation of modern Felidae: a genetic assessment. *Science*, **311**, 73–77.
Kingdon J (2007) *The Kingdon Field Guide to African Mammals*. A&C Black Publishers, Ltd., London.
Kurten B, Anderson E (1980) *Pleistocene mammals of North America*. Columbia University Press, New York.
Lehmann CER, Archibald SA, Hoffmann WA, Bond WJ (2011) Deciphering the distribution of the savanna biome. *The New phytologist*, **191**, 197–209.
Levinsky I, Araújo MB, Nogués-Bravo D *et al.* (2013) Climate envelope models suggest spatio-temporal co-occurrence of refugia of African birds and mammals. *Global Ecology and Biogeography*, **22**, 351–363.

Lorenzen ED, Heller R, Siegismund HR (2012) Comparative phylogeography of African savannah ungulates. *Molecular ecology*, **21**, 3656–70.
Lorenzen ED, Simonsen BT, Kat PW, Arctander P, Siegismund HR (2006) Hybridization between subspecies of waterbuck (Kobus ellipsiprymnus) in zones of overlap with limited introgression. *Molecular ecology*, **15**, 3787–99.
Masembe C, Muwanika VB, Nyakaana S, Arctander P, Siegismund HR (2006) Three genetically divergent lineages of the Oryx in eastern Africa: Evidence for an ancient introgressive hybridization. *Conservation Genetics*, **7**, 551–562.
Mazák JH (2010) Geographical variation and phylogenetics of modern lions based on craniometric data. *Journal of Zoology*, **281**, 194–209.
Migliore J, Baumel A, Juin M *et al.* (2013) Surviving in mountain climate refugia: new insights from the genetic diversity and structure of the relict shrub Myrtus nivellei (Myrtaceae) in the Sahara Desert. *PloS one*, **8**, e73795.
Miller SM, Bissett C, Burger A *et al.* (2013) Management of reintroduced lions in small , fenced reserves in South Africa : an assessment and guidelines. *South African Journal of Wildlife Research*, **43**, 138–154.
Moodley Y, Harley EH (2006) Population structuring in mountain zebras (Equus zebra): The molecular consequences of divergent demographic histories. *Conservation Genetics*, **6**, 953–968.
Moritz C (1994) Defining "Evolutionarily Significant Units" for conservation. *Trends in ecology & evolution (Personal edition)*, **9**, 373–375.
Nersting LG, Arctander P (2001) Phylogeography and conservation of impala and greater kudu. *Molecular ecology*, **10**, 711–9.
Nowell K, Jackson P (1996) *Wild Cats*. IUCN, Gland, Switzerland.
Nylander JAA (2004) MrModeltest v2.3.
Petit JR, Raynaud D, Basile I *et al.* (1999) Climate and atmospheric history of the past 420 , 000 years from the Vostok ice core , Antarctica. *Nature*, **399**, 429–436.
Pitra C, Hansen AJ, Lieckefeldt D, Arctander P (2002) An exceptional case of historical outbreeding in African sable antelope populations. *Molecular ecology*, **11**, 1197–1208.
Riggio J, Jacobson A, Dollar L *et al.* (2012) The size of savannah Africa: a lion's (Panthera leo) view. *Biodiversity and Conservation*, **22**, 17–35.
Ronquist F, Teslenko M, van der Mark P *et al.* (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic biology*, **61**, 539–42.
Staver AC, Archibald S, Levin S (2011) Tree cover in sub-Saharan Africa: rainfall and fire constrain forest and savanna as alternative stable states. *Ecology*, **92**, 1063–72.
Steinmetz A, Eulenberger K, Thielebein J *et al.* (2006) Lens-anomalies and other ophthalmic findings in a group of closely-related angola lions (Panthera leo bleyenberghi). *Zoo Biology*, **25**, 433–439.
Templeton A (2002) Out of Africa again and again. *Nature*, **416**, 45–51.
Thapar V, Thapar R, Ansari Y (2013) *Exotic Aliens: The lion & the cheetah in India*. Aleph Book Company.
Tishkoff S a, Reed F a, Friedlaender FR *et al.* (2009) The genetic structure and history of Africans and African Americans. *Science (New York, N.Y.)*, **324**, 1035–44.
De Vivo M, Carmignotto AP (2004) Holocene vegetation change and the mammal faunas of South America and Africa. *Journal of Biogeography*, **31**, 943–957.
Yamaguchi N, Cooper A, Werdelin L, Macdonald DW (2004) Evolution of the mane and group-living in the lion (Panthera leo): a review. *Journal of Zoology*, **263**, 329–342.
Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. The University of Texas at Austin.

Data accessibility

All sequence data generated in this study have been submitted to GenBank. Accession numbers are listed in Supplemental Table S3.

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. Lion samples and outgroup sequences used in this study.

Number	Country	Location	Region	Source	Literature
1	Senegal	Niokolo-Koba NP	West	Wild	Bertola et al., 2014
2	Senegal	Niokolo-Koba NP	West	Wild	Bertola et al., 2014
3	Senegal	Niokolo-Koba NP	West	Wild	Bertola et al., 2014
4	Senegal	Niokolo-Koba NP	West	Wild	Bertola et al., 2014
5	Senegal	Niokolo-Koba NP	West	Wild	Bertola et al., 2014
6	Senegal	Niokolo-Koba NP	West	Wild	Bertola et al., 2014
7	Senegal	Niokolo-Koba NP	West	Wild	Bertola et al., 2014
8	Guinea	-	West	Wild	this study
9	Benin	Pendjari NP	West	Wild	this study
10	Benin	Pendjari NP	West	Wild	Bertola et al., 2011
11	Benin	Pendjari NP	West	Wild	Bertola et al., 2014
12	Benin	Pendjari NP	West	Wild	Bertola et al., 2014
13	Benin	Pendjari NP	West	Wild	Bertola et al., 2014
14	Nigeria	Yankari GR	West	Wild	this study
15	Nigeria	Yankari GR	West	Wild	this study
16	Nigeria	Yankari GR	West	Wild	this study
17	Nigeria	Yankari GR	West	Wild	this study
18	Nigeria	Kainji NP	West	Wild	this study
19	Nigeria	Kainji NP	West	Wild	this study
20	Cameroon	Waza NP	Central	Wild	Bertola et al., 2011
21	Cameroon	Waza NP	Central	Wild	this study
22	Cameroon	Waza NP	Central	Wild	Bertola et al., 2011
23	Cameroon	Waza NP	Central	Wild	Bertola et al., 2011
24	Cameroon	Waza NP	Central	Wild	Bertola et al., 2011
25	Cameroon	Waza NP	Central	Wild	Bertola et al., 2011
26	Cameroon	Waza NP	Central	Wild	Bertola et al., 2011
27	Cameroon	Waza NP	Central	Wild	Bertola et al., 2011
28	Cameroon	Waza NP	Central	Wild	Bertola et al., 2011
29	Cameroon	Faro NP	Central	Wild	this study
30	Cameroon	Faro NP	Central	Wild	this study
31	Cameroon	Bénoué NP	Central	Wild	this study
32	Cameroon	Bénoué NP	Central	Wild	Bertola et al., 2011
33	Cameroon	Bouba Njida NP	Central	Wild	this study
34	Cameroon	Bouba Njida NP	Central	Wild	this study
35	Cameroon	Bouba Njida NP	Central	Wild	this study
36	Chad	Zakouma NP	Central	Wild	Bertola et al., 2011
37	Chad	Zakouma NP	Central	Wild	Bertola et al., 2011
38	Chad	Zakouma NP	Central	Wild	Bertola et al., 2011
39	Chad	Zakouma NP	Central	Wild	Bertola et al., 2011
40	CAR	Birao	Central	Museum (I)/Wild	this study
41	DRC	Garamba NP	Central	Wild	Bertola et al., 2014
42	DRC	Garamba NP	Central	Wild	this study
43	DRC	Garamba NP	Central	Wild	Bertola et al., 2014
44	DRC	Garamba NP	Central	Wild	Bertola et al., 2014
45	DRC	Garamba NP	Central	Wild	Bertola et al., 2014
46	DRC	Garamba NP	Central	Wild	Bertola et al., 2014
47	DRC	-	Central	Captive (i)	this study
48	DRC	-	Central	Captive (i)	Bertola et al., 2011
49	DRC	Ruindi Plains S. of Lake Edward	Central	Museum (II)/Wild	this study
50	Ethiopia	Gambela NP	East	Wild	this study
51	Ethiopia	Kaffa Province	East	Wild	this study
52	Ethiopia	Nechisar NP	East	Wild	this study
53	Ethiopia	Nechisar NP	East	Wild	this study
54	Ethiopia	Bale Mountains NP	East	Wild	this study
55	Ethiopia	Oromia region, Hudet	East	Wild	this study
56	Ethiopia	Somali region, Dolo Ado	East	Wild	this study
57	Ethiopia	Somali region, Kebri Dehar	East	Wild	this study
58	Ethiopia	Somali region	East	Wild/Captive (ii)	this study
59	Ethiopia	Somali region	East	Wild/Captive (ii)	this study
60	Ethiopia	-	East	Captive (iii)	this study
61	Ethiopia	-	East	Captive (iii)	Bertola et al., 2014
62	Ethiopia	-	East	Captive (iii)	Bertola et al., 2014

Number	Country	Location	Region	Source	Literature
63	Ethiopia	-	East	Captive (iii)	Bertola et al., 2014
64	Ethiopia	-	East	Captive (iii)	Bertola et al., 2014
65	Ethiopia	-	East	Captive (iv)	Bertola et al., 2011
66	Ethiopia	-	East	Captive (iv)	Bertola et al., 2011
67	Ethiopia	-	East	Captive (iv)	Bertola et al., 2011
68	Ethiopia	-	East	Captive (iv)	Bertola et al., 2011
69	Somalia	-	East	Captive (v)	Bertola et al., 2011
70	Somalia	-	East	Captive (v)	Bertola et al., 2011
71	Somalia	-	East	Captive (vi)	this study
72	Kenya	Aberdare NP	East	Wild	this study
73	Kenya	Maasai Mara NR	East	Wild	this study
74	Kenya	Maasai Mara NR	East	Wild	this study
75	Kenya	Maasai Mara NR	East	Wild	this study
76	Kenya	Maasai Mara NR	East	Wild	this study
77	Kenya	Maasai Mara NR	East	Wild	this study
78	Kenya	Maasai Mara NR	East	Wild	this study
79	Kenya	Amboseli NP	East	Wild	Bertola et al., 2014
80	Kenya	Amboseli NP	East	Wild	Bertola et al., 2014
81	Kenya	Amboseli NP	East	Wild	Bertola et al., 2014
82	Kenya	Amboseli NP	East	Wild	Bertola et al., 2014
83	Kenya	Amboseli NP	East	Wild	Bertola et al., 2014
84	Kenya	Amboseli NP	East	Wild	Bertola et al., 2014
85	Kenya	Amboseli NP	East	Wild	Bertola et al., 2014
86	Kenya	Kuku group ranch	East	Wild	this study
87	Kenya	Tsavo East NP	East	Wild	this study
88	Kenya	Tsavo East NP	East	Wild	this study
89	Kenya	Tsavo East NP	East	Wild	this study
90	Kenya	Tsavo East NP	East	Wild	this study
91	Tanzania	Serengeti NP	East	Wild	Bertola et al., 2014
92	Tanzania	Serengeti NP	East	Wild	Bertola et al., 2014
93	Tanzania	Serengeti NP	East	Wild	Bertola et al., 2014
94	Tanzania	Ngorongoro Conservation Area	East	Wild	Bertola et al., 2014
95	Zambia	Mpika town	South	Wild	this study
96	Zambia	Mulobezi town	South	Wild	this study
97	Zambia	Mumbwa town	South	Wild	this study
98	Zambia	north of Lusaka	South	Wild	this study
99	Zambia	Luangwa valley	South	Wild	Bertola et al., 2014
100	Zambia	Luangwa valley	South	Wild	Bertola et al., 2014
101	Zambia	Luangwa valley	South	Wild	Bertola et al., 2014
102	Zambia	Luangwa valley	South	Wild	Bertola et al., 2014
103	Zambia	Luangwa valley	South	Wild	Bertola et al., 2014
104	Zambia	Luangwa valley	South	Wild	Bertola et al., 2014
105	Zambia	Luangwa valley	South	Wild	Bertola et al., 2014
106	Zambia	Luangwa valley	South	Wild	Bertola et al., 2014
107	Zambia	Luangwa valley	South	Wild	Bertola et al., 2014
108	Zambia	Luangwa valley	South	Wild	this study
109	Zambia	Luangwa valley	South	Wild	this study
110	Zambia	Luangwa valley	South	Wild	this study
111	Zambia	Victoria Falls	South	Museum (III)/Wild	this study
112	Zambia	Mweru Wantipa	South	Museum (IV)/Wild	this study
113	Zimbabwe	Save Valley Conservancy	South	Wild	this study
114	Zimbabwe	Save Valley Conservancy	South	Wild	this study
115	Zimbabwe	Save Valley Conservancy	South	Wild	this study
116	Zimbabwe	Save Valley Conservancy	South	Wild	this study
117	Zimbabwe	Nuanetsi area	South	Museum (V)/Wild	this study
118	Zimbabwe	Robin's Camp	South	Museum (V)/Wild	this study
119	Botswana	-	South	Captive (vii)	Bertola et al., 2011
120	Botswana	-	South	Captive (vii)	this study
121	Botswana	-	South	Captive (vii)	Bertola et al., 2011
122	Botswana	Aha Hills	South	Museum (III)/Wild	this study
123	Mozambique	Caia	South	Museum (VI)/Wild	this study
124	Mozambique	Changara	South	Museum (III)/Wild	this study
125	Angola	-	South	Museum (VI)/Wild	this study
126	Namibia	Etosha NP	South	Wild	Bertola et al., 2014
127	Namibia	Etosha NP	South	Wild	Bertola et al., 2014
128	Namibia	Easter Caprivi	South	Wild	this study
129	Namibia	East Etosha	South	Wild	this study
130	Namibia	East and West Caprivi/Botswana	South	Wild	this study
131	Namibia	Eastern Etosha	South	Wild	this study
132	Namibia	Eastern Etosha	South	Wild	this study
133	Namibia	West Caprivi and Angola	South	Wild	this study
134	Namibia	Etosha Central	South	Wild	this study
135	Namibia	East Etosha	South	Wild	this study
136	Namibia	Erongo/Walvis Bay	South	Museum (II)/Wild	this study
137	Namibia	-	South	Captive (viii)	this study
138	Namibia	-	South	Captive (viii)	Bertola et al., 2011
139	RSA	Kalahari Gemsbok NP	South	Wild	Bertola et al., 2014
140	RSA	Kalahari Gemsbok NP	South	Wild	Bertola et al., 2014
141	RSA	Kgalagadi to Tswalu	South	Wild	this study
142	RSA	Kgalagadi to Tswalu	South	Wild	this study

Number	Country	Location	Region	Source	Literature
143	RSA	Kgalagadi Transfrontier Park	South	Wild	this study
144	RSA	Kgalagadi Transfrontier Park	South	Wild	this study
145	RSA	Kruger NP: Gogonthaba, Malelane	South	Wild	Bertola et al., 2014
146	RSA	Kruger NP: Gogonthaba, Malelane	South	Wild	Bertola et al., 2014
147	RSA	Kruger NP: Skukuza Phabeni/ Nwaswitshaka watergat pad junction	South	Wild	Bertola et al., 2014
148	RSA	Kruger NP: Skukuza Phabeni/ Nwaswitshaka watergat pad junction	South	Wild	Bertola et al., 2014
149	RSA	Kruger NP: Pretoriuskop, Fayi loop	South	Wild	Bertola et al., 2014
150	RSA	Kruger NP: Lower Sabie, S128	South	Wild	Bertola et al., 2014
151	RSA	Kruger NP: Crocodile Bridge	South	Wild	Bertola et al., 2014
152	RSA	Kruger NP: Crocodile Bridge	South	Wild	Bertola et al., 2014
153	RSA	Kruger NP: Stolznek, North of Biyamiti	South	Wild	Bertola et al., 2014
154	RSA	Kruger NP: Stolznek	South	Wild	Bertola et al., 2014
155	RSA	Venetia-Limpopo NR, Tuli Block	South	Wild	this study
156	RSA	Venetia-Limpopo NR, Tuli Block	South	Wild	this study
157	RSA	Venetia-Limpopo NR, Tuli Block	South	Wild	this study
158	RSA	Venetia-Limpopo NR, Tuli Block	South	Wild	this study
159	RSA	Venetia-Limpopo NR, Tuli Block	South	Wild	this study
160	RSA	Venetia-Limpopo NR, Tuli Block	South	Wild	this study
161	RSA	Kruger NP: Timbavati	South	Captive (ix)	Bertola et al., 2011
162	RSA	Kruger NP: Timbavati	South	Captive (ix)	this study
163	RSA	Kruger NP: Timbavati	South	Captive (ix)	Bertola et al., 2011
164*	RSA	-	South	Museum (I)/Wild	this study
165	Barbary	-	North	Museum (II)/Wild	this study
166	Barbary	-	North	Museum (VI)/Wild	this study
167*	Middle-East	-	North	Museum (VII)/Captive	this study
168*	Middle-East	-	North	Museum (VII)/Captive	this study
169	India	Gir Forest NP	India	Wild	Bertola et al., 2014
170	India	Gir Forest NP	India	Wild	Bertola et al., 2014
171	India	Gir Forest NP	India	Wild	Bertola et al., 2014
172	India	Gir Forest NP	India	Wild (Captive born (x), founders both Wild)	Bertola et al., 2014
173	India	Gir Forest NP	India	Wild (Captive born (x), founders both Wild)	Bertola et al., 2014
174	India	Gir Forest NP	India	Wild (Captive born (x), founders both Wild)	this study
175	India	Gir Forest NP	India	Captive (i)	Bertola et al., 2011
176	India	Gir Forest NP	India	Captive (i)	Bertola et al., 2011
177	India	Gir Forest NP	India	Captive (i)	Bertola et al., 2011
178	India	Gir Forest NP	India	Captive (xi)	Bertola et al., 2011
179	Senegal (PL1)	-	West	GenBank	Barnett et al., 2006+2014
180	Senegal (PL2)	-	West	GenBank	Barnett et al., 2006+2014
181	Barbary (PL3)	-	North	GenBank	Barnett et al., 2006+2014
182	Burkina Faso (PL4)	-	West	GenBank	Barnett et al., 2006+2014
183	Tunisia (PL5)	-	North	GenBank	Barnett et al., 2006+2014
184	North Africa (PL6)	-	North	GenBank	Barnett et al., 2006+2014
185	Algeria (PL7)	-	North	GenBank	Barnett et al., 2006+2014
186	Iran (PL8)	-	North	GenBank	Barnett et al., 2006+2014
187	Iran (PL9)	-	North	GenBank	Barnett et al., 2006+2014
188	North Africa "Tower of London" (PL11)	-	North	GenBank	Barnett et al., 2006+2014
189	North Africa "Tower of London" (PL12)	-	North	GenBank	Barnett et al., 2006+2014
190	Sudan (PL13)	-	Central	GenBank	Barnett et al., 2006+2014
191	CAR (PL15)	-	Central	GenBank	Barnett et al., 2006+2014
192	CAR (PL16)	-	Central	GenBank	Barnett et al., 2006+2014
193	Asiatic lion (*Panthera leo persica*)	India	Support 16 mitochondrial regions	GenBank	Bagatharia et al., 2013
194	Asiatic lion (Panthera leo persica)	India	Support 16 mitochondrial regions	GenBank	Bagatharia et al., 2013
195	Cave lion (*Panthera leo spelaea*)	Germany	Outgroup	GenBank	Burger et al. (2004) + Barnett et al. (2009)
196	Leopard (*Panthera pardus*)	Amur	Outgroup	Captive (xi)	this study
197	Leopard (*Panthera pardus*)	Unknown	Outgroup	GenBank	Wei et al. (2011)
198	Tiger (*Panthera tigris*)	Bengal	Outgroup	GenBank	Kitpipit et al. (2011)
199	Tiger (*Panthera tigris*)	Bengal	Outgroup	GenBank	Kitpipit et al. (2011)
200	Tiger (*Panthera tigris*)	Sumatra	Outgroup	GenBank	Kitpipit et al. (2011)
201	Tiger (*Panthera tigris*)	Sumatra	Outgroup	GenBank	Kitpipit et al. (2011)
202	Tiger (*Panthera tigris*)	Amur	Outgroup	GenBank	Kitpipit et al. (2011)
203	Tiger (*Panthera tigris*)	Amur	Outgroup	GenBank	Kitpipit et al. (2011)
204	Snow leopard (*Panthera uncia*)	Unknown	Outgroup	GenBank	Wei et al. (2011)
205	Clouded leopard (*Neofelis nebulosa*)	Unknown	Outgroup	GenBank	Wu et al. (2007)

* Excluded from analyses presented in the main text. See Supplemental Information 2 for background information and additional analyses.

SOURCE	**MUSEUM**	I: Zoological Museum, University of Amsterdam, Amsterdam, The Netherlands / Naturalis Biodiversity Center, Leiden, The Netherlands
		II: Swedish Museum of Natural History, Stockholm, Sweden
		III: Smithsonian: Smithsonian Institution, Washington D.C., U.S.A.
		IV: Brussels: Royal Belgian Institute of Natural Sciences, Brussels, Belgium
		V: Bulawayo: Natural History Museum of Zimbabwe, Bulawayo, Zimbabwe
		VI: Naturalis Biodiversity Center, Leiden, The Netherlands
		VII: Humboldt: Museum für Naturkunde (MfN)/Humboldt Museum, Berlin, Germany
	CAPTIVE	i: Diergaarde Blijdorp, Rotterdam, The Netherlands
		ii: BornFree Ethiopia, lions confiscated from the Presidential Palace in Addis Ababa, Ethiopia
		iii: Addis Ababa Lion Zoo, Addis Ababa, Ethiopia
		iv: Sanaa Zoo, Sanaa, Yemen
		v: Safaripark Beekse Bergen, Hilvarenbeek, The Netherlands
		vi: Confiscated individual, Breeding Centre, UAE
		vii: Dierenpark Amersfoort, Amersfoort, The Netherlands
		viii: Zoo Basel, Basel, Switserland
		ix: Ouwehands dierenpark, Rhenen, The Netherlands
		x: Sakkarbaug Zoo: Sakkarbaug Zoological Garden, Junagadh, Gujarat, India
		xi: Planckendael: Planckendael, Muizen, Belgium

Supplemental Table S2. (online only) Overview of processing of lion samples included and accompanying Genbank accession numbers.

Supplemental Table S3. Primers used in this study.

Primer Set1	**3 Primers**		
Primer name	**Annealing Temp**	**Sequence (5'-3')**	**Reference**
1F	50-55 ˚C	CGTTGTACTTCAACTATAAGAACTT	Bertola et al., 2011
1R	50-55 ˚C	ATGGGATTGCTGATAGGAGATTAG	Bertola et al., 2011
2F	53-55 ˚C	GTGGGGCCAAATATCCTTTT	Bertola et al., 2011
4R	53-55 ˚C	TTTTTGGTTTACAAGACCAAGGTA	Bertola et al., 2011
5F	53-55 ˚C	AAATCGCCTCCTCAAATGAA	Bertola et al., 2011
5R	53-55 ˚C	AATATTCATGGGAGGGCAGTC	Bertola et al., 2014
Primer Set2	**5 Primers**		
Primer name	**Annealing Temp**	**Sequence (5'-3')**	**Reference**
1F	50-55 ˚C	CGTTGTACTTCAACTATAAGAACTT	Bertola et al., 2011
1R	50-55 ˚C	ATGGGATTGCTGATAGGAGATTAG	Bertola et al., 2011
2F	51-53 ˚C	GTGGGGCCAAATATCCTTTT	Bertola et al., 2011
2R	51-53 ˚C	GAAGGCCTAGGATATCTTTGATTG	Bertola et al., 2014
3F	51-53 ˚C	GACTCAGATAAAATTCCATTCCA	Bertola et al., 2014
3R	51-53 ˚C	CATTATTCCTCGCTGTTTGG	Bertola et al., 2014
4F	51-53 ˚C	CAATTATCCCTGCCCTCCA	Bertola et al., 2014
4R	51-53 ˚C	TTTTTGGTTTACAAGACCAAGGTA	Bertola et al., 2011
5F	53-55 ˚C	AAATCGCCTCCTCAAATGAA	Bertola et al., 2011
5R	53-55 ˚C	AATATTCATGGGAGGGCAGTC	Bertola et al., 2014
Primer Set3	**12 Primers**		
Primer name	**Annealing Temp**	**Sequence (5'-3')**	**Reference**
aDNA1F	50 ˚C	CGTTGTACTTCAACTATAAGAACTT	this study
aDNA1R	50 ˚C	CTAGAAAGAGGCCGGTGAGAA	this study
aDNA2F	50 ˚C	GCTCCTTATTAGGAGTATGCTTAATCC	this study
aDNA2R	50 ˚C	CATGCATGTATAGGCAGATAAAGA	this study
aDNA3F	50 ˚C	TGGCTGAATTATCCGGTACCTA	this study
aDNA3R	50 ˚C	GCACCTCAAAAGGATATTTGG	this study
aDNA4F	50 ˚C	AGCTACAGCCTTCATAGGATATGT	this study
aDNA4R	50 ˚C	TGGAAGGATGAAGTGGAAGG	this study
aDNA5F	50 ˚C	GGAGGCTTCTCAGTAGACAAAG	this study
aDNA5R	50 ˚C	TGATTGTATAGTATGGATGGAATGG	this study
aDNA6F	50 ˚C	CCCCTCAGGAATGGTATCTG	this study
aDNA6R	50 ˚C	ATATGGGGAGGGGTGCTTAG	this study
aDNA7F	50 ˚C	CTCACCAGACCTATTAGGAGATCC	this study
aDNA7R	50 ˚C	GAGGGCAGGGATAATTGCTA	this study
aDNA8F	50 ˚C	GCAATCCTCCGATCTATTCC	this study
aDNA8R	50 ˚C	CCAATTCATGTCAGGGTCAG	this study
aDNA9F	50 ˚C	CTTATTCTGATTCCTAGTAGCGGA	this study
aDNA9R	50 ˚C	CGTTCTCCTTTTTTGGTTTACAAG	this study
aDNA10F	50 ˚C	GCCTCCTCAAATGAAGAGTCT	this study
aDNA10R	50 ˚C	TGCAATATATGAATTGTGAAAGTTACG	this study
aDNA11F	50 ˚C	GCACCCAAAGCTGAAATTCT	this study
aDNA11R	50 ˚C	TCACTTGCTTTTCGTGGGG	this study
aDNA12F	50 ˚C	CTGTGCTTGCCCAGTATGTC	this study
aDNA12R	50 ˚C	CTGTACATGCTTAATATTCATGGG	this study
Primer Set LR	**4 Primers**		
Primer name	**Annealing Temp**	**Sequence (5'-3')**	**Reference**
LR 1F (NADH4)	56 ˚C	CTCACTTTCTGCACCTCTACTAGTCTTA	this study
LR 1R (16S)	56 ˚C	ACGGATCAGAAGTAAGAGACAGTAAAG	this study
LR 2F (16S)	56 ˚C	CATCACCTCTAGCATTTCCAGTATTAG	this study
LR 2R (NADH4)	56 ˚C	ACTAGCCATGAGCATTAGTGGTAGG	this study

Supplemental Table S4. Results of estimates for divergence times for lion clades in years ago (ya), compared to estimates from previous publications. Constraints include the approach and calibration points used. Names of the clades refer to the ingroups.

		This study	**Burger et al., 2004**	**Antunes et al., 2008**	**Barnett et al., 2014**
Software		BEAST	r8s	Lintree	BEAST
Data		complete mtDNA (total: 17 kb) + 1.4 kb fragments	mtDNA (total: 1.1 kbp)	mtDNA (total: 1.9 kbp)	mtDNA (total: 1.1 kbp)
Constraints		0.55 million ya (Cave lion - Lion); 1.6 million ya (Leopard-Lion); 3.8 million ya (Panthera)	clock-based estimate	clock-based estimate	0.55 million ya (Cave lion - Lion)
	Node (Fig. 3)	**Age of nodes (95% HPD)**			
Panthera	not shown	3.443 million ya (2.590-4.373)	1.428–2.295 million ya		
Leopard - Lion	not shown	1.469 million ya (1.135-1.812)	1–1.559 million ya		
Cave lions - Lion	not shown	556,900 ya (510,800-606,100)			
Lion	a	291,700 ya (178,000-417,700)	74,000-203,000 ya	324,000 ya (145,000-502,000)	124,200 ya (81,800-183,500)
South	b	231,300 ya (132,300-338,700)		split not detected	81,900 ya (45,700-122,200)
East/Southern - North East	c	183,700 (107,000-271,100)		split not detected	57,800 ya (26,800-96,600)
North	d	174,700 ya (94,900-276,700)		118,000 ya (28,000-208,000)	61,500 ya (32,700-97,300)
Central - India	e	141,100 ya (71,700 - 216,600)		split not detected	split not detected
India - North Africa	f	110,800 ya (39,700-200,100)		split not detected	21,100 ya (8,300-38,800)
South West	g	113,800 ya (55,200-189,000)		169,000 ya (34,000-304,000)	
East/Southern	h	78,100 ya (37,200-132,100)		101,000 ya (11,000-191,000)	
North East	i	63,900 ya (18,000-118,600)		split not detected	
West - India	-	split not detected		split not detected	51,000 ya (26,600-83,100)
West	j	63,400 ya (15,100-129,100)		split not detected	
Central	k	49,600 ya (20,700-91,000)		split not detected	

Supplemental Information S1. Details on sample storage and processing.

Samples were preserved dried, in 95% ethanol or in buffer (0.15 M NaCl, 0.05 M Tris-HCl, 0.001 M EDTA, pH = 7.5) and stored at -20°C (Supplemental Table 4). For blood and tissue samples DNA was extracted using the DNeasy Blood & Tissue kit (Qiagen) following the manufacturer's protocol. For the scat and the museum samples, a protocol for aDNA extractions from bone and teeth (Rohland & Hofreiter 2007) was followed. In all cases a mock extraction was included to check for contamination. All museum samples were processed in the aDNA facility of DNAmarkerpoint, Leiden University, which is physically isolated from other laboratories and where no previous work on felids had been conducted. In addition, two scat samples which contained strongly degraded DNA, 8.Guinea and 30.Cameroon, were included in the aDNA procedure. Before each extraction, the surfaces in the extraction room were cleaned using 10% bleach and all materials were cleaned and irradiated with UV light for a minimum of one hour.

The complete mitochondrial genome was analyzed for ten individuals by sequencing on an Illumina HiSeq2000 using 99 bp paired-end sequencing with 200-400 bp insert size (Leiden Genome Technology Center, Leiden, The Netherlands). In the first run, two individuals (9.Benin and 89.Kenya) were tagged and pooled with leopard DNA (ratios 1:1:2 for 9.Benin, 89.Kenya and 179.Leopard respectively). In the following two runs, four individuals (21. Cameroon+71.Somalia+162.RSA+ 174.India and 42.DRC+95.Zambia+96.Zambia+131.Namibia) were tagged and equimolarily pooled. Resulting reads were identified based on the unique adapter sequences.

For four individuals the complete mtDNA was analysed by performing two long range PCRs for amplifying all ~18,000 bp. Primers were designed based on known leopard sequences available on Genbank using Primer3v 0.4.0 (Rozen & Skaletsky 2000). Primer sites were chosen such that the forward and corresponding reverse primer were not both located in one of the known numts that have been identified in felids (Lopez *et al.* 1996; Cracraft *et al.* 1998; Kim *et al.* 2006). For amplification either the LA PCR kit (TaKaRa) or the GoTaq Long PCR Master Mix (Promega) was used (Supplemental Table S2). Resulting PCR products were cut out from the gel, cleaned with the Wizard SV gel and PCR Clean-Up kit (Promega) and sonically fragmented. Barcoded Libraries for sequencing were prepared from the fragmented PCR products using the Rapid Library Preparation Kit (Roche). Emulsion PCR and sequencing were performed on the 454/Roche FLX Genome Sequencer Titanium (Forensic Laboratory for DNA Research, Leiden, The Netherlands) according to the protocol.

Cytochrome b, tRNAThr, tRNAPro and the left domain of the control region (hereafter referred to as cytB+ctrl reg.) were amplified using three primer pairs in high quality blood and tissue samples, five primer pairs in the scat samples and twelve primer pairs in the aDNA samples. See Supplemental Table S3 for primer sequences. All primers were designed using the web-based software Primer3v 0.4.0 (Rozen & Skaletsky 2000). The modern samples were amplified using Taq DNA Polymerase (Invitrogen) or Phire Hot Start II DNA Polymerase (Thermo Scientific), depending on the amplification success. Annealing temperature was adjusted according to primer pair and according to previous PCR results (for details see Supplemental Table S2). The museum samples were amplified using AmpliTaq Gold DNA Polymerase (Invitrogen) and following a half-nested approach: in the first round (40 cycles) primer aDNA1F was combined with primer aDNA2R and a 1:50 dilution of the PCR product was used as a template for a second round PCR (40 cycles), in which primer aDNA1F was combined with aDNA1R and primer aDNA2F was combined with aDNA2R etc. In all cases multiple negative PCR controls were included to check for contamination.

Sequencing of the short, non-aDNA PCR products was performed by Macrogen Inc., Amsterdam, The Netherlands. The aDNA samples were sequenced on the Roche/454 platform (Forensic Laboratory for DNA Research, Leiden, The Netherlands). The 12 PCR products for each museum sample were equimolarily pooled, and after a test run containing one sample, the remaining 17 samples were divided in two pools, which were analysed in two separate runs. To check for contamination and to distinguish the samples after sequencing, a unique combination of tags attached to the primers was used for each individual. In addition to the 454 sequencing, 22 PCR products were cloned to confirm sequences with a coverage <10 or inconclusive results (i.e. called base supported by <90% of available reads). Cloning was performed using the Invitrogen TOPO cloning kit following the manufacturer's protocol. From each cloned PCR product, between three and eight colonies were picked. Picked colonies were lysed by heating the cells in 30 µl of water for 10 minutes (min) at 95 °C. Cell lysates were amplified with M13 primers using the following PCR: 2 µl $MgCl_2$ (25 mM), 2 µl 10× PCR buffer, 0.25 µl dNTPs (2.5 mM each), 0.24 µl Taq polymerase, 0.5 µl M13 primers (10 µM each), and 2 µl cell lysate, with water added to a final volume of 20 µl. The PCR program was: 94 °C for 5 min followed by 40 cycles of 94 °C for 30 seconds (s), 55 °C for 45 s, 72 °C for 45 s and a final extension step of 72 °C for 4 min. The PCR products were sequenced by Macrogen Inc., Amsterdam, The Netherlands. Overlap between independent PCR products were used to check for DNA damage and sequencing errors. Unique point mutations (i.e. observed in a single sample) were checked by an independent PCR and sequencing for modern samples, or cloning for aDNA samples.

Read data from Illumina and 454 platforms were analysed using CLC Genomics (CLCBio). A leopard mitochondrial genome available on GenBank (EF551002.1) was used as reference. Mapping was performed by using default settings, except for length fraction and similarity fraction, which were increased to 0.8 and 0.85 respectively. Consensus sequences were extracted and aligned visually with Macrogen sequences. Since we observed one region that seemed to be absent in all Illumina samples, but present in all sequences derived by PCR and Sanger sequencing, and another region where the opposite was true, we constructed a new reference sequence and repeated the mapping of all Illumina and 454 reads, which lead to a more consistent coverage across the reference sequence. Sequences covering cytB+ctrl reg. that had already been analysed in earlier publications (Barnett *et al.* 2006a; Barnett *et al.* 2006b; Barnett *et al.* 2014; Bertola *et al.* 2011; Bertola *et al.* submitted) were added to the dataset for phylogenetic analyses.

Since Roche/454 sequencing does not perform well with mononucleotide repeats, all mononucleotide repeats of >3 bp were manually checked. Gaps resulting from inconclusive base calling were substituted by an ambiguous nucleotide. This was also done for inconclusive results on six positions in three aDNA samples which could not be resolved and a 62bp region with insufficient coverage in sample 165.Barbary. Two repetitive regions in the control region, RS-2 and RS-3, were excluded from the analysis, since aligning was difficult and the region is known to be heteroplasmic (Jae-Heup *et al.* 2001). In addition, a mononucleotide repeat of cytosines of variable length was excluded due to unknown homology (bp 1382-1393 in cytB+ctrl reg.). For phylogenetic analysis 179.Leopard was used as an outgroup and supplemented by six sequences from Genbank: clouded leopard (*Neofelis nebulosa*: DQ257669.1), snow leopard (*Panthera uncia*: EF551004.1), two sequences of tiger (*Panthera tigris*: JF357968.1 (Bengal) and JF357974.1 (Amur)), one sequence of leopard (*Panthera pardus*: EF551002.1) and one sequence of cave lion (*Panthera leo spelaea*: KC701376.1 + DQ899901.1). In addition, two complete mitochondrial genomes from Asiatic lions were included (JQ904290.1 and KC834784.1) (not included in Figures). Since the sequences from Genbank did not align well in the control region, likely due to the assembly method, this region of the Genbank sequences was replaced by ambiquous nucleotides to eliminate the influence of assembly quality.

References

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). *Proceedings of the Royal Society B: Biological Sciences*, **273**, 2119–2125.

Bertola LD, van Hooft WF, Vrieling K *et al.* (2011) Genetic diversity, evolutionary history and implications for conservation of the lion (Panthera leo) in West and Central Africa. *Journal of Biogeography*, **38**, 1356–1367.

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

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

Kim J-H, Antunes A, Luoa S-J *et al.* (2006) Evolutionary analysis of a large mtDNA translocation (numt) into the nuclear genome of the Panthera genus species. *Brain, Behavior and Immunity*, **22**, 629–629.

Lopez JV, Cevario S, O'Brien SJ (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**, 229–46.

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

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds S. K, S. M), pp. 365–386. Humana Press, Totowa, NY.

Supplemental Information S2. Data authenticity.

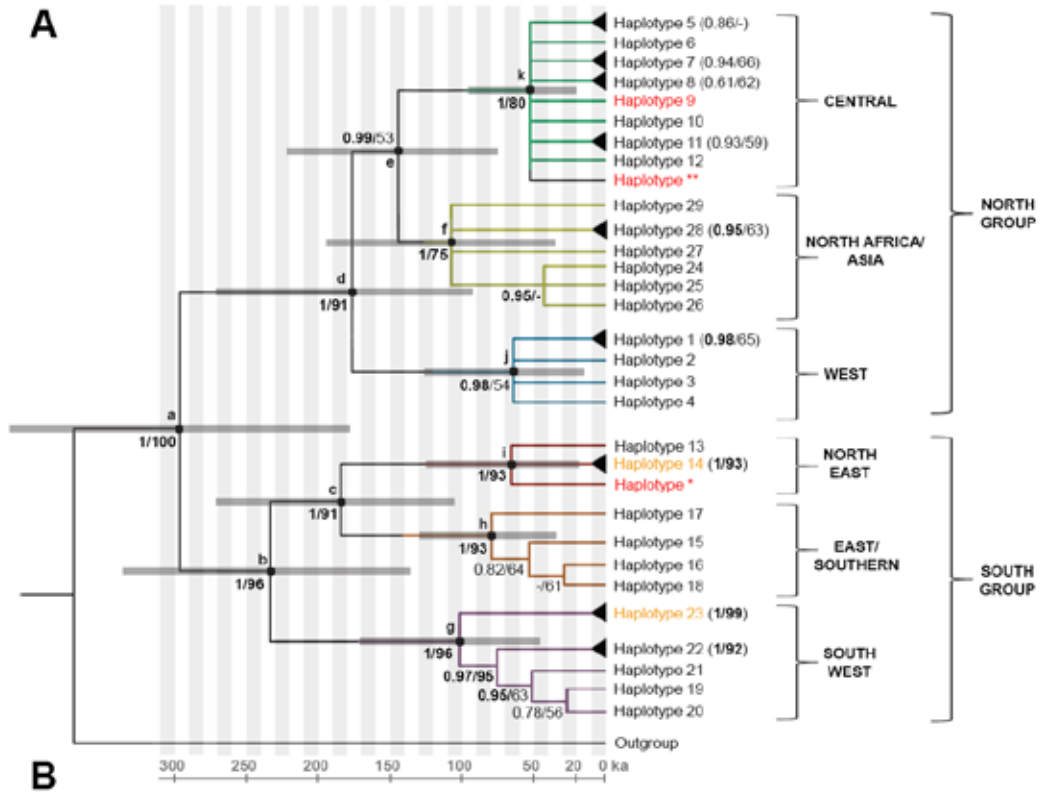
Samples from zoos and museums were only included in our study when sufficient information was available on the origin of the individual or its free-ranging ancestors. In four cases, samples showed unexpected results from the phylogenetic analyses. Since the origin of three samples could not be reconfirmed, they were excluded from analyses presented in the main text. For completeness, results of the analyses including these samples and captive popultions are shown below. In all cases, unique point mutations were double checked by independent PCR and sequencing and laboratory procedures were checked to exclude the possibility of contamination. Addition of these samples does not change the conclusions presented in the main text.

Haplotype 14: Ethiopia captive population (65-68 Ethiopia). This population is located on a long branch, clustering with the North East group. Despite relatively dense sampling of the region, no intermediate haplotypes were identified. Clustering based on mtDNA data and microsatellite data do not contradict the origin of these samples (Bertola *et al.*, submitted). These data were therefore included in all analyses.

Haplotype 23: Namibia captive population (137-138.Namibia). This population is located on a long branch in the South West group, with undetected intermediate haplotypes. Phylogenetic analyses place the population on the expected branch, in the South West group. These data were therefore included in all analyses.

Museum sample 164.RSA (Haplotype *). This sequence was placed in the North East group, with data from Ethiopia, Somalia and Central Kenya. Apart from this specimen, all included samples from the southern part of Kenya and further southward cluster with either East/Southern or the South West group. No samples from the North East group had been processed parallel to this sample and therefore we exclude the possibility of contamination. The specimen was collected by the late L. de Beaufort and comparing this entry to other specimen collected by L. de Beaufort, this entry contained very little information. Because of doubts regarding the authenticity of this entry, and the unexpected position in the phylogenetic tree, this sample was excluded from the phylogenetic analyses presented in the main text. Results for Bayesian, Maximum Likelihood, Network and BEAST analyses including this sample are shown below (Supplemental Figures S2-1 and S2-2).

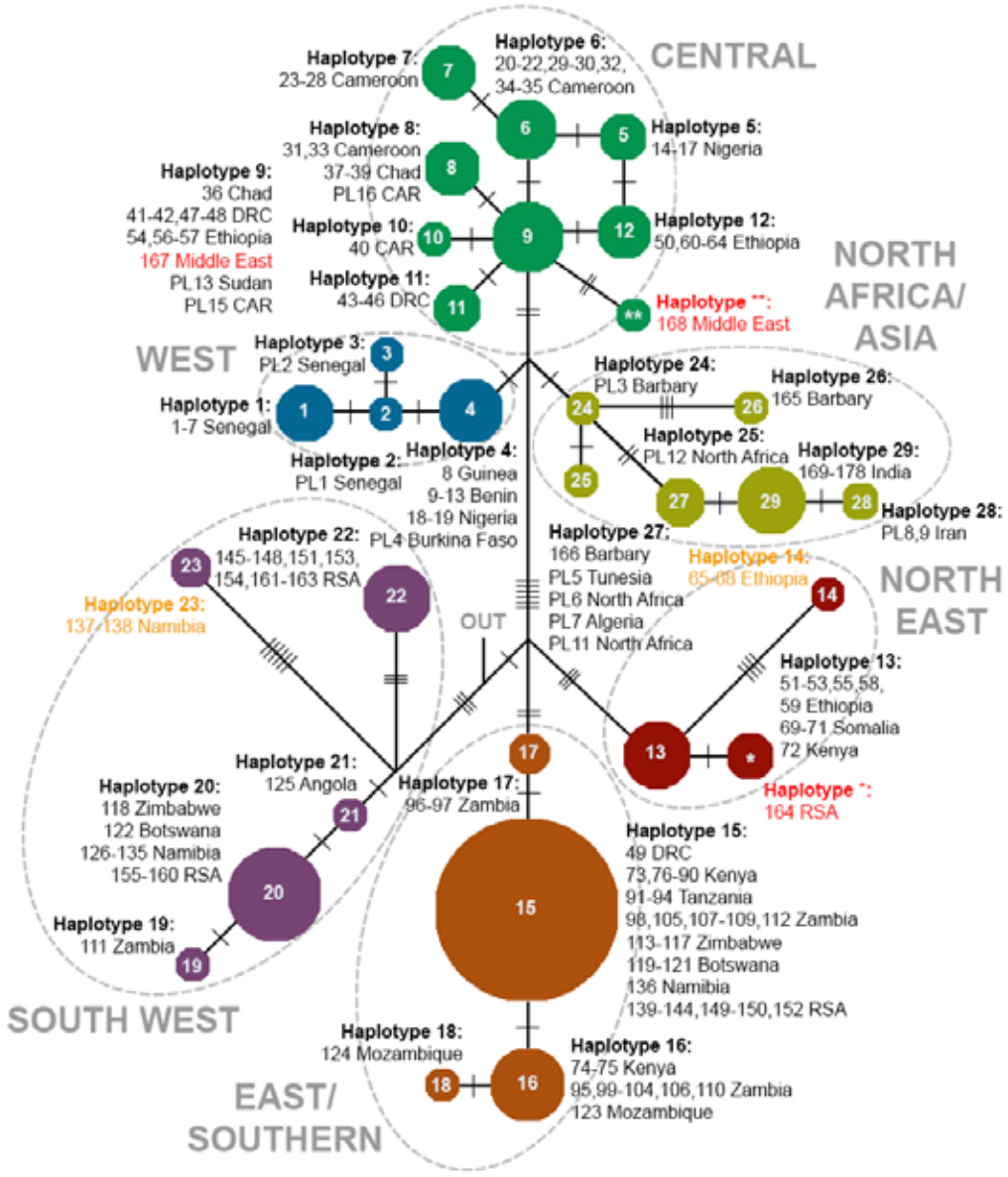
Museum samples 167-168.Middle East (Haplotype 9 and **): these specimen were labeled as hybrids between an Abyssinian male and a female from Mesopotamia (first generation zoo animals). They share a haplotype or cluster close to a haplotype from Central Africa. In contrast, the remaining ten sequences from North Africa and Iran cluster strongly with the Asiatic subspecies. No samples from the Central Africa group had been processed parallel to this sample and therefore we exclude the possibility of contamination. Regarding the sparse information about zoo populations in those times and the unexpected position in the phylogenetic tree, these specimen were excluded from the phylogenetic analyses presented in the main text. Haplotype 9 was retained, since this was found in several other samples from Central Africa. Results for Bayesian, Maximum Likelihood, Network and BEAST analyses including this sample are shown below (Supplemental Figures S2-1 and S2-2).



B

Clade	Node	Date (95% HPD)	Clade	Node	Date (95% HPD)	Clade	Node	Date (95% HPD)	Clade	Node	Date (95% HPD)
Modern Lion	a	293.8 ka (178.6-416.3)	North	d	176.8 ka (92.9-271.8)	South West Africa	g	1.4 ka (43.4-170.2)	West Africa	j	63.4 ka (15.0-127.6)
South	b	231.2 ka (138.7-337.2)	Central Africa - India	e	142.0 ka (75.1 - 222.8)	East/ Southern Africa	h	78.1 ka (34.8-129.7)	Central Africa	k	51.7 ka (19.7-95.0)
East/Southern - North East Africa	c	182.3 ka (104.3-271.4)	India - North Africa	f	106.4 ka (35.8-194.2)	North East Africa	i	65.1 ka (18.1-124.5)			

Supplemental Figure 2-1. Phylogenetic analyses for the complete lion dataset, including sixteen mitochondrial genomes and 178 cytb+ctrl reg. sequences A: Phylogenetic tree of lion populations throughout their complete geographic range, based on complete mitochondrial genomes and cytB+ctrl reg. sequences. Branch colours correspond to haplotype colours in Supplemental Figure 2-2. Populations mentioned above as long branches with missing intermediate haplotypes, are indicated in orange. Populations with limited information regarding their origin, which were excluded from analyses presented in the main text, are shown in red. Support is indicated as posterior probability (Bayesian analysis)/bootstrap support (ML analysis). Branches with a single haplotype have been collapsed to improve readability. Support for these branches is indicated by a black triangle at the tip of the branch (support shown in the label). Nodes which have been included for divergence time estimates are indicated with letters and 95% HPD node bars. Distance to outgroup and nodes without dated split is not in proportion to divergence time. B: divergence time estimates and 95% HPD from BEAST analysis, also indicated as error bars in Supplemental Figure 2-1A.



Supplemental Figure 2-2. Haplotype network based on cytB+ctrl reg. sequences of lions throughout their entire geographic range. Dashed lines indicate the groups discerned Bayesian/ML analysis in Supplemental Figure 2-1A. Populations indicated above as long branches with missing intermediate haplotypes, are shown in orange. Populations with limited information regarding their origin, which were excluded from analyses presented in the main text, are indicated in red. Haplotype size is proportional to its frequency in the dataset. Hatch marks represent a change in the DNA sequence. The connection to outgroup species is indicated by "OUT".

References

Antunes A., Troyer J.L., Roelke M.E., Pecon-Slattery J., Packer C., Winterbach C., Winterbach H., Hemson G., Frank L., 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.

Barnett R., Yamaguchi N., Shapiro B., Ho S.Y., Barnes I., Sabin R., Werdelin L., Cuisin J., Larson G. 2014. Revealing the maternal demographic history of Panthera leo using ancient DNA and a spatially explicit genealogical analysis. BMC Evolutionary Biology. 14:70.

Bertola L.D., van Hooft W.F., Vrieling K., Uit de Weerd D.R., York D.S., Bauer H., Prins H.H.T., Funston P.J., Udo de Haes H. a., Leirs H., van Haeringen W. a., Sogbohossou E., Tumenta P.N., Iongh H.H. de. 2011. Genetic diversity, evolutionary history and implications for conservation of the lion (Panthera leo) in West and Central Africa. Journal of Biogeography. 38:1356–1367.

Bruche S., Gusset M., Lippold S., Barnett R., Eulenberger K., Junhold J., Driscoll C. a., Hofreiter M. 2012. A genetically distinct lion (Panthera leo) population from Ethiopia. European Journal of Wildlife Research.

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, Panthera leo. Conservation Genetics. 6:15–24.

Dubach J.M., Briggs M.B., White P.A., Ament B.A., Patterson B.D. 2013. Genetic perspectives on "Lion Conservation Units" in Eastern and Southern Africa. Conservation Genetics. 1942.

Steinmetz A., Eulenberger K., Thielebein J., Buschatz S., Bernhard A., Wilsdorf A., 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.

Chapter 5

SNP discovery and phylogenetic analyses across ten populations of lions reveals a more complex evolutionary history

(in prep.)

L. D. Bertola, M. Vermaat, P. A. White, H. H. de Iongh, J. Laros and K. Vrieling

Abstract

Next generation sequencing techniques allow for the generation of new magnitudes of unlinked genetic markers, which can be used to infer phylogeographic patterns in non-model organisms, such as the lion. Previous studies in lions, mostly based on mtDNA and microsatellite markers, have shown that the distribution of genetic diversity is not in line with the current taxonomy, only distinguishing an African and an Asiatic subspecies. The addition of genome-wide, unlinked genetic markers provides us with a more complete picture of the underlying genomic complexity. Full genome sequencing and subsequent variant calling has resulted in the discovery of 44,627 SNPs in ten lions, sampled throughout their geographic range, one leopard and one tiger. A total of 18,457 SNPs was variable within the lion. Phylogenetic trees based on autosomal SNPs show a gradual differentiation in the lion, following a north-south axis, and no reciprocally monophyletic groups could be identified. However, the Asiatic subspecies shows a nested position within the African subspecies, indicating that the current nomenclature does not follow the deepest evolutionary split for the distinction of subspecies. Phylogenetic trees based on the mitochondrial genome show a strongly supported split between lions from the northern part of their range, and lions from the southern part of their range. Since autosomal SNP data do not show a conflicting pattern, we suggest that this distinction should be followed in a taxonomic revision of the lion.

Keywords: Single Nucleotide Polymorphism (SNP), SNP discovery, genome sequencing, phylogeography, lion (*Panthera leo*)

Introduction

The rise of next generation sequencing (NGS) techniques has opened up possibilities to apply massive parallel sequencing to non-model organisms, like the lion (*Panthera leo*). Inferring population histories and reconstruction of the evolutionary history of a species can therefore be based on a new magnitude of unlinked data. Species histories are favorably based on data from multiple loci, due to the fact that genetic markers may represent different evolutionary trajectories (mtDNA vs. autosomal DNA) and due to stochasticity in the coalescence of markers (Edwards 2009; Knowles 2009). Mitochondrial DNA (mtDNA) has been proven to be a useful tool for gaining insight in phylogeographic patterns, partially because of its shorter coalescence time compared to nuclear markers. However, it represents one locus only and obtained haplotype trees are therefore not necessarily a true representation of the underlying genomic complexity (Zink & Barrowclough 2008; Edwards & Bensch 2009).

The lion (*Panthera leo*) was subjected to several phylogeographic studies which have contributed to current insights into the distribution of genetic diversity in the African subspecies (*Panther leo leo*) and its connection to the Asiatic subspecies (*Panthera leo persica*). These studies included data from mtDNA (Dubach *et al.* 2005, 2013; Barnett *et al.* 2006a; b, 2014; Antunes *et al.* 2008; Bertola *et al.* 2011; Bertola *et al.* submitted), autosomal DNA (Antunes *et al.* 2008; Dubach *et al.* 2013; Bertola *et al.* submitted) and pathogens (Antunes *et al.* 2008). The general emerging pattern was that of a basal dichotomy, recognizing a Northern group with populations from West and Central Africa including the Asiatic subspecies (*Panthera leo persica*), and a Southern group with populations from East and Southern Africa (Bertola *et al.* 2011a; Dubach *et al.* 2013; Barnett *et al.* 2014). Within these two groups, more phylogenetic lineages can be recognized, with notably long lineages in the Southern group. Admixture between haplogroups was only found in two occasions: 1) haplotypes from both the Central and the North East group are found in the suture zone in Ethiopia, and 2) haplotypes from the South West group and the East/Southern are found in the Kruger/Limpopo area, Republic of South Africa (RSA), likely to be the result of human-mediated translocations (Bertola *et al*. submitted) (Miller *et al.* 2013) (Figure 1). Microsatellite data are roughly congruent, also identifying a distinct position for the West and Central African lion, and a subsequent split between East and Southern Africa in the populations from the southern part of the range. However, there is a geographic discrepancy in the southern part of the range, where admixture is clearly visible, notably between East/Southern and South West Africa based on autosomal data only (Figure 1). The above mentioned studies have not only given a fine scale picture of current genetic diversity in the lion, but also illustrate how this diversity deviates from the current taxonomic nomenclature, only recognizing an African and an Asiatic subspecies (Bertola *et al.* 2011a; Dubach *et al.* 2013; Barnett *et al.* 2014) (Bertola *et al.,* submitted). This has led to requests for a taxonomic revision for this species (Bertola *et al.* 2011a; Dubach *et al.* 2013; Barnett *et al.* 2014) (Bertola *et al.,* submitted)

Although mtDNA and autosomal data have not shown strongly conflicting patterns in lion phylogeography, additional data from genome wide markers would benefit both the understanding of the evolutionary history of the species, and guiding of conservation efforts. According to Moritz (1994) intraspecific genetic diversity can be used as a rational for conservation practices, by following a two-step approach and defining 1) Evolutionary Significant Units (ESUs), and 2) Management Units (MUs). The inclusion of nuclear data for the recognition of ESUs is essential to avoid misclassifying

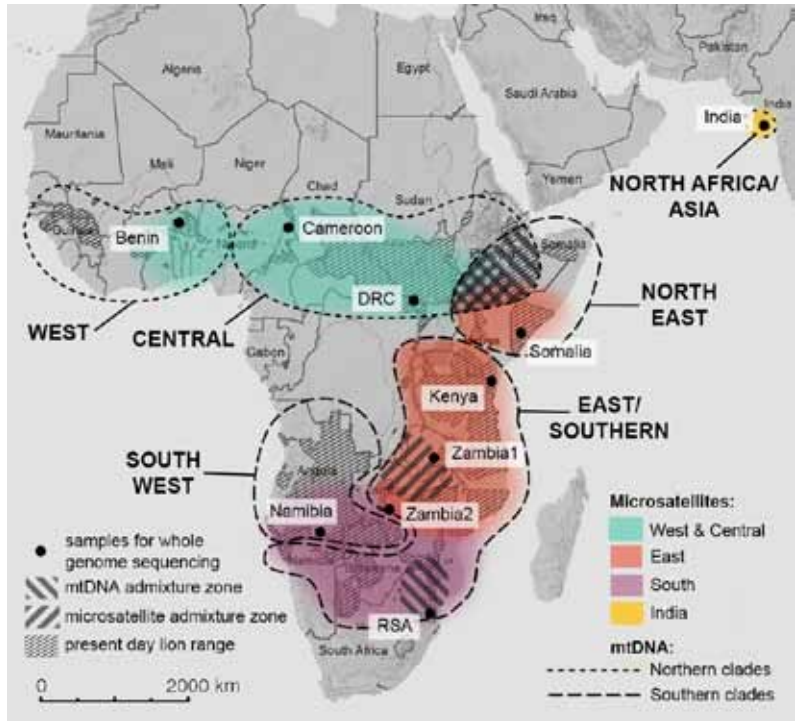


Figure 1. Phylogeographic groups in the lion identified based on mtDNA (Bertola *et al.*, submitted) and microsatellite data (Bertola *et al.*, submitted), and locations of lion samples included in this study. Lion range data from IUCN (2014).

populations which are linked by nuclear, but not by organellar gene flow (Moritz 1994). It is known that female lions exhibit strong philopatry and that male lions are capable dispersers (Pusey *et al.* 1987; Spong & Creel 2001), indicating that this aspect may be relevant in this species. In previous studies on lions which contained autosomal data, few populations were included as representatives. Also autosomal data were mainly represented by microsatellite loci (Antunes *et al.* 2008; Dubach *et al.* 2013; Bertola *et al.* submitted), which are of limited use to infer phylogenetic relationships due to their mutation pattern and high variability. To derive a complete picture of the evolutionary history, a broader range of autosomal markers should be targeted, and compared to the available mtDNA datasets.

In this study, we describe the discovery of Single Nucleotide Polymorphisms (SNPs) by targeting variable positions from whole genome data of ten lions, covering the main phylogeographic groups as were indicated based on previously published mtDNA and microsatellite data (Bertola *et al.* submitted; Bertola *et al.* submitted). The obtained SNPs are analyzed in a phylogeographic framework. Compared to previously published phylogeographic patterns, based on mtDNA and microsatellite data, this provides a more complete overview of the complexity underlying intraspecific genetic diversity in the lion. Finally, a selection of the discovered SNPs can be used for a wider study on more sampling locations, potentially contributing to future studies on lion genetics.

Materials and Methods

Blood or tissue samples of ten lions, representing the main phylogeographic groups (Figure 1), were collected and stored in buffer solution (0.15 M NaCl, 0.05 M Tris-HCl, 0.001 M EDTA, pH = 7.5) at -20 °C. All included individuals were either free-ranging lions or captive lions with proper documentation of their breeding history. A sample from a leopard (*Panthera pardus orientalis*, captive) was included as an outgroup. All samples were collected in full compliance with specific legally required permits (CITES and permits related to national legislation in the countries of origin).

DNA was extracted using the DNeasy Blood & Tissue kit (Qiagen) following the manufacturer's protocol. The DNA was sequenced on 3 lanes of an Illumina HiSeq2000 to 99 bp paired end reads with 200-400 bp insert size (Leiden Genome Technology Center, Leiden, The Netherlands). In the first run, two individuals (Benin and Kenya) were tagged and pooled with leopard DNA as the outgroup (ratios 1:1:2 for Benin, Kenya and Leopard respectively). In the two following runs four individuals (Cameroon+Somalia+RSA+India and DRC+Zambia1+Zambia2+Namibia) were tagged and equimolarily pooled (Supplemental Table S1). Resulting reads were identified based on the unique adapter sequences.

The sequencing run containing Benin, Kenya and Leopard was repeated, since the first run produced read pairs with a severe drop in quality in the second read (Supplemental Figure S1). We hard-clipped these reads after the first 30 bp and added these data to the reads derived from a second run of the same samples. Quality control was performed using the FastQC tool (Andrews 2010) on the raw reads, and after removing adapter sequences with cutadapt (Martin 2011) and quality trimming with Sickle (Joshi & Fass 2011).

Samples Benin and RSA showed bimodal distributions of GC content per read and high average GC content compared to the other samples (55% and 45%, respectively, versus ~40%), indicating contamination with bacterial DNA. A nucleotide Blast search (Altschul *et al.* 1990) was done on a random selection of 10,000 reads per sample. Bacterial genomes of the highest hits were downloaded from GenBank (Supplemental Table S2) and reads for samples Benin and RSA were aligned against these using BWA (Li & Durbin 2009). Only unaligned reads were retained. In a second filtering step, only reads aligning to the reference genome of an Amur tiger (*Panthera tigris altaica*) (Cho *et al.* 2013) were included for downstream analyses. Re-analysis of the GC content distribution for these samples showed that these filtering steps eliminated the second peak (Supplemental Figure S2).

A reference genome was created by concatenating an Amur tiger assembly (Cho *et al.* 2013) and supplementing this with a lion mtDNA genome (30. Cameroon; Bertola *et al.*, submitted). Reads of lions and Leopard were aligned to this reference using BWA (Li & Durbin 2009).

Single Nucleotide Variant (SNV) calling was performed using SAMtools mpileup (Li *et al.* 2009) with default settings on Leopard (outgroup) separately and all lion samples jointly. SNV calling was executed excluding samples Benin and RSA, because of the influence of these samples on the available coverage per sample. We filtered calls based on their quality (phred score ≥20) and per-sample read depth (≥6 for Leopard, ≥3 for all lion samples). Sample alleles at variant sites were derived from Leopard calls and lion calls on non-contaminated samples (i.e. excluding Benin and RSA). This file was enriched with data for Benin and RSA from the joint calling including all lions for positions where enough coverage for all samples was available. All other positions were filled with ambiguous nucleotides (N). This procedure was repeated using only sites that were variant within the lion samples (i.e. excluding outgroups). Calling Y chromosomal SNVs in the eight male samples was done as described above, only on scaffolds supposedly located on the Y chromosome, identified by aligning all known Y chromosomal regions in cat (*Felis catus*) to the genomic data from Cho *et al.* (2013) (Supplemental Table S3). We configured SAMtools to assume a haploid genome and all positions with a heterozygote calling (22 out of 164) were discarded. The resulting sample alleles were serialized to FASTA format and served as input for the phylogenetic analysis. The complete pipeline used in this project and additional information is available at: https://git.lumc.nl/lgtc-bioinformatics/bertola-lion

Phylogenetic analyses were performed using MrBayes v.3.1.2 (Huelsenbeck & Ronquist 2001; Ronquist *et al.* 2012) and Garli (Zwickl 2006), using parameters determined by MrModeltest2 (v.2.3) (Nylander 2004). Branches receiving >0.95 PP in Bayesian analysis (MrBayes) and/or 70 bootstrap support in Maximum Likelihood (ML) analysis (Garli) were considered to be significantly supported. In addition a Principle Component Analysis (PCA) was executed in Genalex (Peakall & Smouse 2012) and R version 3.1.0, using prcomp. Isolation by Distance (IBD) analyses were performed in Genalex using 999 permutations (Peakall & Smouse 2012), excluding the contaminated samples Benin and RSA due to difficulties in estimating genetic distance with a high frequency of ambiguous nucleotides. Levels of differentiation (Fst) were calculated using Arlequin using 1023 permutations (Excoffier *et al.* 2005). The level of heterozygosity was assessed for each lion, taking into account the numbers of scored (non-ambiguous) nucleotides. Further, identified SNPs were attributed to a chromosome, following the genomic architecture in the tiger (Cho *et al.* 2013) and Bayesian analyses and PCA were repeated for individual chromosomes. In addition, mtDNA data were subjected to Bayesian and ML analysis, and PCA by using mitochondrial genomes as identified by Bertola *et al.* (submitted).

Results

The sequencing runs yielded a total of 628,716,470 reads and after quality control a total of 593,632,293 reads (94.4%) were retained for subsequent alignment (Supplemental Table S1). Filtering of variable positions between ten lions, one leopard and one tiger, yielded 44,627 variable positions, of which 18,457 positions were variable within the lion. Assuming identical chromosomal architecture in the lion as in the tiger, we find a strong relationship between discovered SNPs in this study and estimated chromosome sizes in the tiger (Cho *et al.* 2013) (Supplemental Table S4). On the Y chromosome 142 SNPs were identified compared to the outgroup species. Coverage plots for all individuals and all scaffolds illustrate the Y-chromosomal origin, since hardly any coverage is found for the females included (Supplemental Figure S3). Since only 1 Y chromosomal position is variable within the lion, this alignment was not further subjected to phylogenetic analyses. Mitochondrial genomes, consisting of 16,756 bp, excluding repetitive regions RS-2 and RS-3 (Jae-Heup *et al.* 2001), were added to the dataset. On the mtDNA 2,317 SNPs were identified, with 742 variable positions within the lion.

Phylogenetic analyses, based on all lion-specific SNPs, showed a hierarchical pattern in which the populations from the northern part of the lion range represent the most basal branches (Figure 2). Exclusion of the contaminated samples, Benin and RSA, which contained high numbers of missing values, did not influence the topology or support of the tree. Similarly, the exclusion of intermediate populations, i.e. DRC, Somalia and Kenya, did not change the overall topology of the tree. As was previously shown with mtDNA markers, the Asiatic subspecies shows a close genetic relationship to lions from West and Central Africa and does not have an outgroup position. The phylogenetic tree based on the mitochondrial genomes shows a basal dichotomy, although the branch containing the southern populations is not well supported. Phylogenetic trees and PCA from individual chromosomes show largely congruent patterns (Supplemental Figure S4).

IBD analyses showed a strongly significant correlation between genetic and geographic distance, both including and excluding the Asiatic subspecies (Supplemental Table 5). Since tree topology indicates a more gradual differentiation, in contrast to the basal dichotomy observed in the mtDNA, population differentiation was calculated, regarding the geographically intermediate populations Somalia and Kenya as either 1) North, 2) South or 3) Intermediate. Pairwise Fst values were significant ($P<0.05$) in all cases, except when Somalia and Kenya were included as Intermediate, in which case only North and South populations showed significant differentiation from each other.

Individual levels of heterozygosity were assessed and compared to previously published data from Bertola *et al.* (submitted) and (Dubach *et al.* 2013) (Supplemental Table 6). Ranking of these levels between SNP data and microsatellite data finds strong congruence, although contaminated samples Benin and RSA had to be excluded due to the low coverage, which may bias the number of heterozygote positions.

Discussion

This study shows how whole genome sequencing can be used for SNP discovery in a non-model species, in this case the lion. The results are used in a phylogenetic framework to infer evolutionary histories of the lion and compare these results to previously published scenarios. Since this approach mainly served the identification of variable positions in the lion, the number of included samples for a phylogenetic analysis is restricted. However, the identified SNPs can be used as a source for the generation of a SNP panel, based on which a larger number of individuals can be genotyped.

Previously published mtDNA datasets of the lion showed a strongly supported basal dichotomy, clustering all populations from the northern part of the range, including the Asiatic subspecies, and all populations from the southern part of the range. Although the branch with southern populations did not receive significant support when only ten individuals were included (Figure 2), we interpret the tree as having a basal dichotomy, as was previously shown by Barnett *et al.* (2014) and Bertola *et al.* (submitted). This basal dichotomy is less pronounced in the SNP data, notably due to the structure in the northern part of the range. However, the Asiatic subspecies is nested in the African lion tree, close to lions from West and Central Africa, further undermining the validity of its distinct subspecies status. The hierarchical pattern observed in the SNP data may largely be attributed to continent-wide gene flow, explaining the more gradual pattern of population differentiation. The consecution in which the individuals branch off, support this explanation.

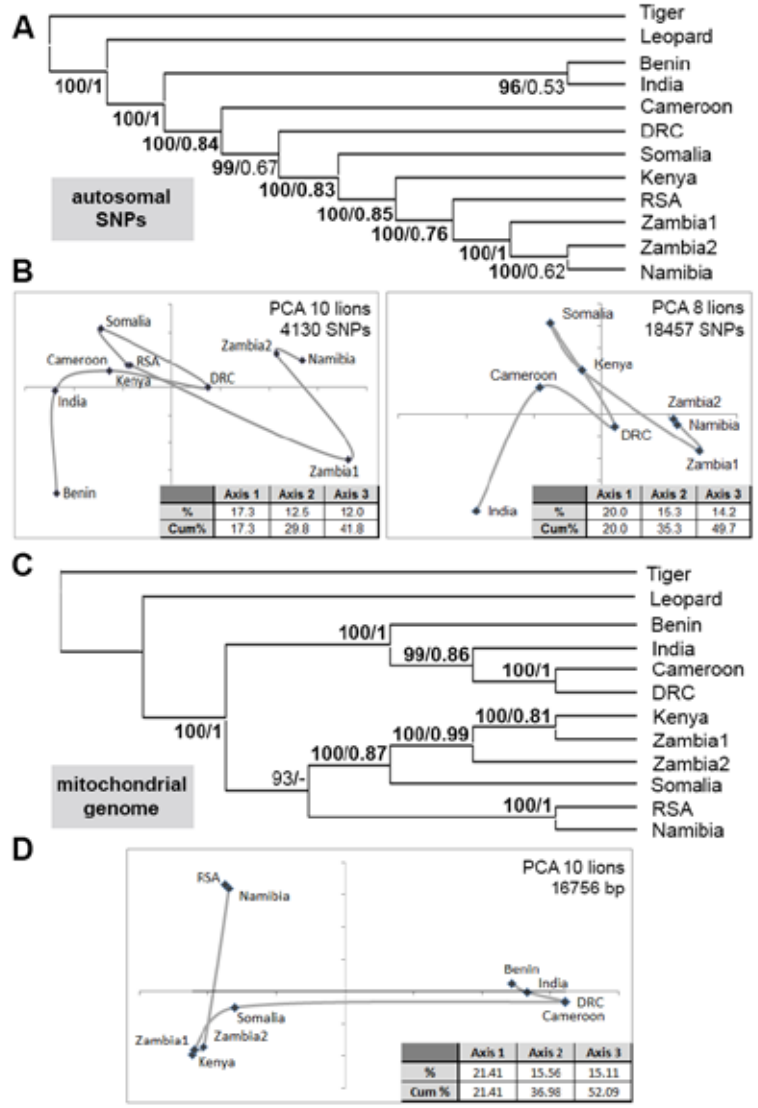


Figure 2. Bayesian analysis and PCA of SNPs in lion, leopard and tiger. A: Bayesian and ML analysis of 18,457 SNPs in ten lions, with posterior probability/bootstrap values indicated at the nodes. B: PCA of all variable positions in the lion, including and excluding the contaminated samples Benin and RSA. The line connects the populations in the same order as the topology of the tree. C: Bayesian and ML analysis of complete mitochondrial genomes of 10 lions, with posterior probability/bootstrap values indicated at the nodes. D: PCA based on the complete mitochondrial genome of ten lions. The line connects the populations in the same order as the topology of the tree.

Since dispersal in lions is biased to the male sex (Pusey *et al.* 1987; Spong & Creel 2001), this may explain why we see a more discrete phylogenetic pattern in the mtDNA. Major barriers for gene flow seem to be restricted to the (recent) population gap in North Africa/Middle East and the Central African rain forest. Although the Rift valley has frequently been mentioned as a barrier for gene flow in the lion (Dubach *et al.* 2005; Barnett *et al.* 2006a; b; Bertola *et al.* 2011a), and gene flow may be reduced, admixture between haplogroups indicates that the Rift valley is not a complete barrier for

lion dispersal (Bertola et al, submitted). In historic times, additional barriers may have existed as a result of expanding rain forest or desert (Bertola *et al.* submitted). The restriction of suitable lion habitat to a small number of refugia may have contributed to the development of discrete genetic lineages. The pattern found in mtDNA data of the lion is congruent with that of other species (Hewitt 2004; Lorenzen *et al.* 2012; Bertola *et al.* submitted) and predicted refugial areas based on climate models (Levinsky *et al.* 2013). Faster coalescence times of mtDNA may have led to reciprocally monophyletic mtDNA clades in the lion, while isolation in refugia may not have lasted long enough for coalescence in autosomal markers, due to the cyclic character of the African climate (Bertola et al, submitted).

Autosomal SNPs and microsatellite data are expected to produce largely congruent patterns because of a similar mode of inheritance and coalescence times. Due to the hierarchical nature of the SNP tree it is difficult to interpret which groups can be considered to be discrete. SNP data may represent a more ancient pattern, in which historic gene flow is strongly represented, while phylogeographic patterns based on microsatellite data may, as a result of their high mutation rate, represent relatively recent evolutionary history, as is the case for fast coalescent markers, like mtDNA. This may explain why distinct clusters are relatively easily retrieved from microsatellite data, but not from SNP data.

Based on microsatellites population Zambia1 was indicated as an admixture zone. IBD analysis from the SNP data seem to confirm this: notably after exclusion of India, Zambia1 forms a relatively distinct cloud, representing low genetic distance, compared to the other pairwise comparisons. We do not find indications for a suture zone between mtDNA haplogroups in this region, indicating the admixture may be the result of male-biased gene flow. An admixture pattern of haplogroups is found in Ethiopia, where the presence of a suture zone is further supported by microsatellite data (Bertola *et al.* submitted). Based on current sampling locations in DRC, Kenya and Somalia, their position in the PCA plots and the formation of a loop connecting these sampling localities also suggests admixture. SNP data from more sampling localities in this region may be able to further support this. Finally, the position of RSA in the PCA plot may be the result of human-mediated admixture in RSA, visible as a mosaic pattern of haplogroups in the Kruger/Limpopo area. This individual contains a South West haplotype, but is likely to be admixed with East/Southern African lions, which explains the close position to Kenya.

Ranking observed heterozygosity values, results in a congruent pattern between SNP and microsatellite data. This supports the notion that, even though a single individual per population has been sampled for the SNP discovery, the number of SNPs identified can give an indication of within-population diversity levels. SNP genotyping for more individuals from a single population could be executed to further strengthen this point.

A genome wide SNP panel, based on ten lions from the main phylogeographic groups, shows a gradual degree of relatedness of lions following a north-south axis, and a nested position of the Asiatic lion within the African subspecies. This suggests that the current nomenclature, recognizing an African and an Asiatic subspecies, conflicts with the distribution of genetic diversity in the species, as was previously shown for mtDNA data only (Dubach *et al.* 2013; Barnett *et al.* 2014; Bertola *et al.* submitted). Although the phylogeographic pattern based on genome-wide autosomal markers is more gradual, without recognizing reciprocally monophyletic clades, suggestions regarding management of lion populations postulated by Barnett *et al.* (2014) and Bertola *et al.* (submitted) still hold. Defining units for conservation management by looking for reciprocal monophyly in autosomal data may be overly restrictive. Following current insights, combining mtDNA, microsatellite and genome-wide SNP data, we confirm six ESUs as previously suggested based on reciprocally monophyletic haplogroups: 1) West Africa, 2) Central Africa, 3) India, 4) North East Africa, 5) East/Southern Africa and 6) South West Africa. Finally, due to the nested position of the Asiatic subspecies, we support a taxonomic revision, distinguishing an northern subspecies, including the Asiatic lion, and a southern subspecies in the lion. Based on the discovered SNPs from this paper a SNP panel has been designed, also including mitochondrial SNPs, which can be used for fast and cost-effective genotyping of large numbers of individuals. This method may also be applied for the establishment of breeding programmes for captive stocks or in a forensics framework to trace source populations of illegal lion products. Analysing more free-ranging lion populations will further improve the understanding of their levels of diverstiy, genetic relationships and evolutionary history.

Acknowledgements

Samples were kindly provided by E. Sogbohossou (Benin), P. Tumenta, S.Adam, R. Buij and B. Croes (Cameroon), ICCN, Garamba NP (DRC), Safaripark Beekse Bergen (Hilvarenbeek, The Netherlands) (Somalia), B. Patterson (Kenya), O. Aschenborn (Namibia), Ouwehands Dierenpark (Rhenen, The Netherlands) (RSA), C.A.Driscoll (India) and Planckendael (Muizen, Belgium) (Leopard). We further thank N. Schidlo, H. Buermans, Y. Ariyurek, and S. Greve-Onderwater for assisting in processing of the samples and Cho et al. for assistance with the reference data. The investigations were supported by the Division for Earth and Life Sciences (ALW) with financial aid from the Netherlands Organization for Scientific Research (NWO) (project no. 820.01.002).

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Search Tool 2Department of Computer Science. *Journal of Molecular Biology*, **215**, 403–410.

Andrews S (2010) FastQC - A Quality Control tool for High Throughput Sequence Data.

Antunes A, Troyer JL, Roelke ME *et al.* (2008) The evolutionary dynamics of the lion Panthera leo revealed by host and viral population genomics. *PLoS genetics*, **4**, e1000251.

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). *Proceedings of the Royal Society B: Biological Sciences*, **273**, 2119–2125.

Barnett R, Yamaguchi N, Shapiro B *et al.* (2014) Revealing the maternal demographic history of Panthera leo using ancient DNA and a spatially explicit genealogical analysis. *BMC Evolutionary Biology*, **14**, 70.

Bertola LD, van Hooft WF, Vrieling K *et al.* (2011) Genetic diversity, evolutionary history and implications for conservation of the lion (Panthera leo) in West and Central Africa. *Journal of Biogeography*, **38**, 1356–1367.

Cho YS, Hu L, Hou H *et al.* (2013) The tiger genome and comparative analysis with lion and snow leopard genomes. *Nature communications*, **4**.

Dubach JM, Briggs MB, White PA, Ament BA, Patterson BD (2013) Genetic perspectives on "Lion Conservation Units" in Eastern and Southern Africa. *Conservation Genetics*, **1942**.

Dubach J, Patterson BD, Briggs MB *et al.* (2005) Molecular genetic variation across the southern and eastern geographic ranges of the African lion, Panthera leo. *Conservation Genetics*, **6**, 15–24.

Edwards SV (2009) Is a new and general theory of molecular systematics emerging? *Evolution; international journal of organic evolution*, **63**, 1–19.

Edwards S, Bensch S (2009) Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008. *Molecular ecology*, **18**, 2930–3; discussion 2934–6.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary bioinformatics online*, **1**, 47–50.

Hewitt GM (2004) The structure of biodiversity - insights from molecular phylogeography. *Frontiers in zoology*, **1**, 4.

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics (Oxford, England)*, **17**, 754–5.

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

Joshi NA, Fass JN (2011) Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files.

Knowles LL (2009) Statistical Phylogeography. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 593–612.

Levinsky I, Araújo MB, Nogués-Bravo D *et al.* (2013) Climate envelope models suggest spatio-temporal co-occurrence of refugia of African birds and mammals. *Global Ecology and Biogeography*, **22**, 351–363.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, **25**, 1754–60.

Li H, Handsaker B, Wysoker A *et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, **25**, 2078–9.

Lorenzen ED, Heller R, Siegismund HR (2012) Comparative phylogeography of African savannah ungulates. *Molecular ecology*, **21**, 3656–70.

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EBMnet.journal*, **17**, 10–12.

Miller SM, Bissett C, Burger A *et al.* (2013) Management of reintroduced lions in small , fenced reserves in South Africa : an assessment and guidelines. *South African Journal of Wildlife Research*, **43**, 138–154.

Moritz C (1994) Defining "Evolutionarily Significant Units" for conservation. *Trends in ecology & evolution (Personal edition)*, **9**, 373–375.

Nylander JAA (2004) MrModeltest v2.3.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics (Oxford, England)*, **28**, 2537–9.

Pusey AE, Packer C, Erhoff-Mulder MB (1987) The Evolution os Sex-biased Dispersal in Lions. *Behaviour*, **101**, 275–310.

Ronquist F, Teslenko M, van der Mark P *et al.* (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic biology*, **61**, 539–42.

Spong GF, Creel S (2001) Deriving dispersal distances from genetic data. *Proceedings. Biological sciences / The Royal Society*, **268**, 2571–2574.

Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular ecology*, **17**, 2107–2121.

Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. The University of Texas at Austin.

Data accessibility

The complete pipeline used in this project and additional information is available at: https://git.lumc.nl/lgtc-bioinformatics/bertola-lion

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. Lion and leopard samples included in this study and results whole genome sequencing.

Run	Sample	Corresponding to Bertola et al. (submitted)	Origin	Sex	Read number	Read number after filtering**	Read number after filtering + QC	% left after QC	GC%
1	**Leopard**	196.Leopard	Amur (captive[1])	Male	172,038,427 (+ 153,603,680 partially failed*)	-	-	94	40
	Benin	9.Benin	Benin - Pendjari NP	Female	18,698,767 (+ 16,318,397 partially failed*)	11,846,142 (+ 9,184,453 partially failed)	11,665,188 (+ 8,834,199 partially failed)	94	55 (bimodal)
	Kenya	89.Kenya	Kenya - Tsavo East NP	Male	32,033,356 (+ 28,473,401 partially failed*)	-	-	94	40
2	**India**	174.India	India - Gir forest[2]	Male	32,453,394	-	-	98	40
	Cameroon	21.Cameroon	Cameroon - Waza NP	Male	37,995,410	-	-	98	40
	Somalia	71.Somalia	Somalia (captive[3])	Male	20,683,546	-	-	94	40
	RSA	162.RSA	RSA (captive[4])	Female	29,190,629	-	-	98	45 (bimodal)
3	**DRC**	42.DRC	DRC - Garamba NP	Male	29,552,205	-	-	97	42
	Zambia1	95.Zambia	Zambia - Luangwa Valley	Male	22,700,633	-	-	98	38
	Zambia2	96.Zambia	Zambia - Mulobezi town	Male	25,796,481	-	-	97	42
	Namibia	131.Namibia	Namibia - Etosha NP	Male	25,496,541	-	-	98	39
					160,870,954 (+ 138,506,569 partially failed*)	-	-	94	40
					17,557,523 (+ 13,741,759 partially failed*)	11,846,142 (+ 9,184,453 partially failed)	11,665,188 (+ 8,834,199 partially failed)	94	55 (bimodal)
					30,217,372 (+ 26,035,900 partially failed*)	-	-	94	40
					31,687,608	-	-	98	40
					37,257,634	-	-	98	40
					19,527,670	-	-	94	40
					28,468,409	28,262,272	28,074,951	98	45 (bimodal)
					28,718,622	-	-	97	42
					22,188,079	-	-	98	38
					25,141,802	-	-	97	42
					24,905,745	-	-	98	39

* "partially failed" indicates the number of reads from a previous run, which were trimmed to 30 bp in the second read only due to a severe drop in quality.

** Filtering refers to filtering of reads from bacterial origin in two contaminated samples, Benin and RSA.

1 Planckendael: Planckendael, Muizen, Belgium

2 Captive born, founders both wild; Sakkarbaug Zoo: Sakkarbaug Zoological Garden, Junagadh, Gujarat, India

3 Safaripark Beekse Bergen, Hilvarenbeek, The Netherlands

4 Ouwehands dierenpark, Rhenen, The Netherlands

Supplemental Table S2. Genbank entries to filter bacterial reads in contaminated samples Benin and RSA.

Genbank Accession	Organism	Details
gi|386716467|ref|NC_017671.1|	Stenotrophomonas maltophilia D457	complete genome
gi|206558403|ref|NC_011000.1|	Burkholderia cenocepacia J2315	chromosome 1, complete sequence
gi|206561868|ref|NC_011001.1|	Burkholderia cenocepacia J2315	chromosome 2, complete sequence
gi|191639869|ref|NC_011002.1|	Burkholderia cenocepacia J2315	chromosome 3, complete sequence
gi|206479926|ref|NC_011003.1|	Burkholderia cenocepacia J2315	plasmid pBCJ2315, complete sequence

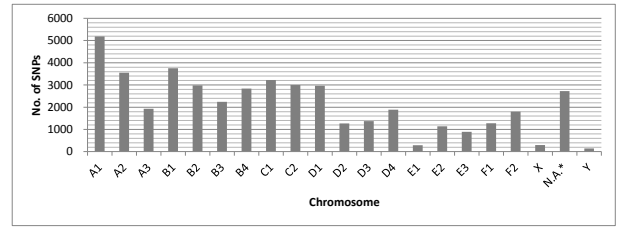
Supplemental Table S3. Scaffolds in the reference sequence (Cho *et al.*, 2013) identified as potentially from Y-chromosomal origin.

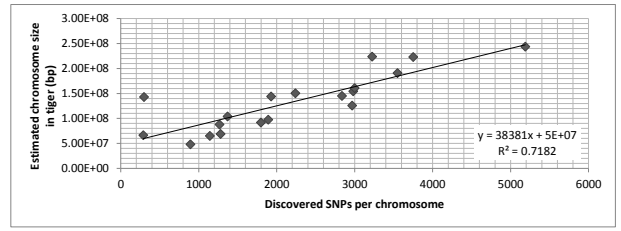
scaffold	Gene	gi	Score	E-value
scaffold725	SRY	77176790	4149	0
scaffold638	UBE1Y	84620608	549	e-153
scaffold363	CYorf15	84620610	172	8.00E-41
scaffold640	CUL4BY	84620611	696	0
scaffold1087	TETY2	84620617	975	0

Supplemental Table S4. Number of discovered SNPs per chromosome and estimated chromosome size in tiger (Cho *et al.*, 2013).

Chromosome	Discovered SNPs	Estimated chromosome size (bp) Amur tiger (Cho et al., 2013)
A1	5,188	243,492,181
A2	3,550	190,495,254
A3	1,929	144,011,757
B1	3,750	222,683,385
B2	2,982	154,295,958
B3	2,240	150,246,213
B4	2,837	144,888,701
C1	3,225	223,586,761
C2	3,000	160,670,131
D1	2,969	125,709,129
D2	1,267	87,703,667
D3	1,371	103,759,264
D4	1,891	97,290,273
E1	290	66,408,731
E2	1,143	64,743,307
E3	893	47,874,673
F1	1,280	68,695,903
F2	1,798	91,576,383
X	298	142,585,357
N.A.*	2,726	-
Y	142	-

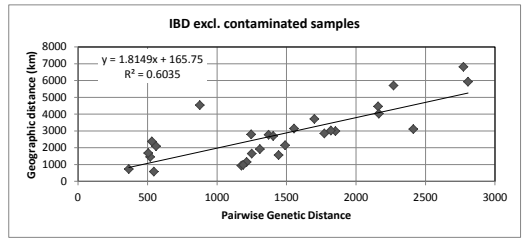
* SNPs which could not be assigned to any of the chromosomes

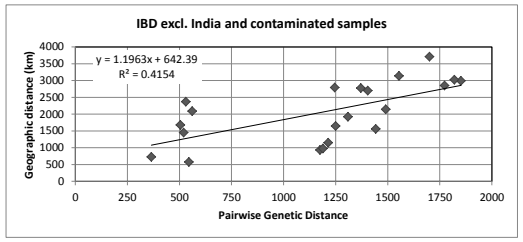




Supplemental Table S5. IBD analysis for 8 lion samples (excluding the contaminated samples Benin and RSA) and for 7 samples (excluding India and the contaminated samples).

Groups	Pair. comparis	Rxy
8 samples (excl. contaminated)	28	0.777 (P≤0.002)
7 samples (excl. India + contaminated)	21	0.645 (P≤0.001)

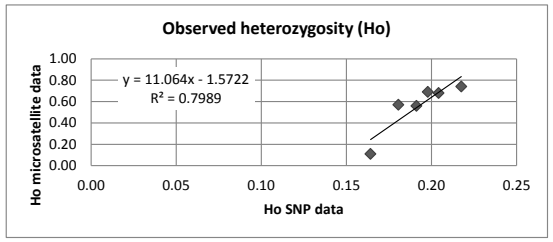


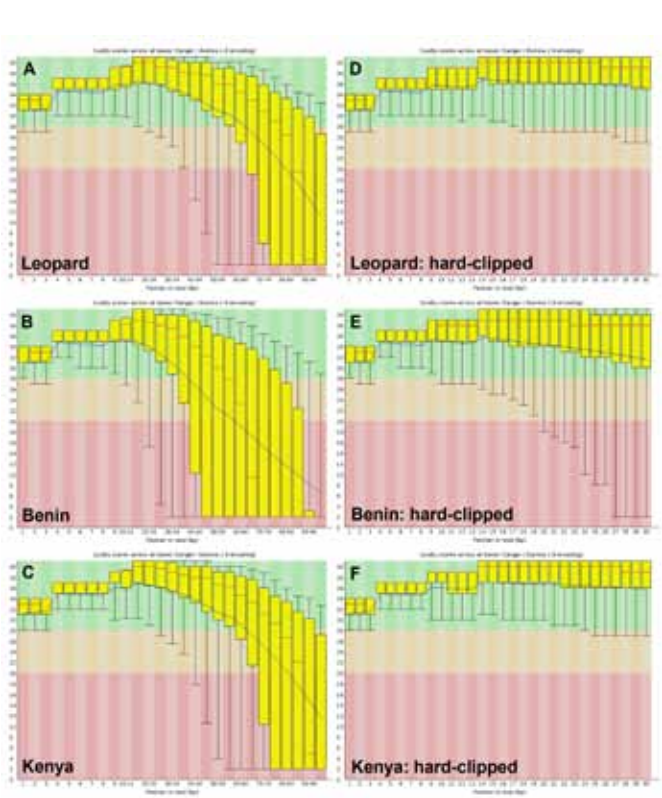


Supplemental Table S6. Observed heterozygosity for all lion samples and comparison with observed heterozygosity based on microsatellite data. Shading indicates the ranking from low heterozygosity (red) to high heterozygosity (green).

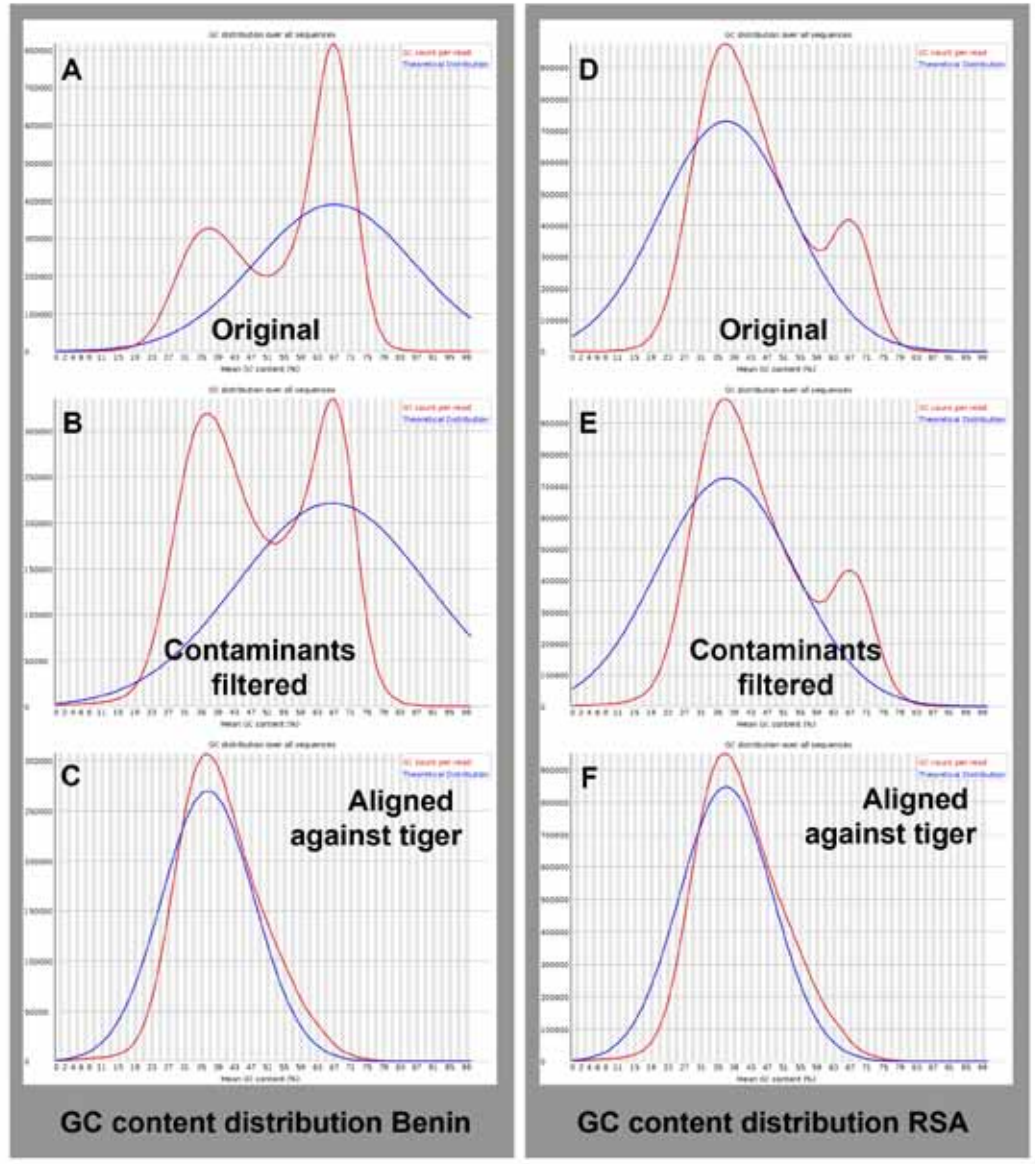
Sample	SNPs scored (non-ambiguous)	heterozygote positions	homozygote positions	Observed heterozygosity (Ho)	Ho based on microsatellite data	Source microsatellite data
Benin	8,106	2,157	5,949	0.27*	0.65	Bertola et al., submitted
India	44,627	7,326	37,301	0.16	0.11	Bertola et al., submitted
Cameroon	44,627	9,111	35,516	0.20	0.68	Bertola et al., submitted
DRC	44,627	9,707	34,920	0.22	0.74	Bertola et al., submitted
Somalia	44,627	8,092	36,535	0.18	-	-
Kenya	44,627	10,004	34,623	0.22	-	-
Zambia1	44,627	8,057	36,570	0.18	0.57	Bertola et al., submitted
Zambia2	44,627	8,828	35,799	0.20	0.69	Dubach et al., 2013
RSA	20,667	5,333	15,334	0.26*	0.69	Bertola et al., submitted
Namibia	44,627	8,532	36,095	0.19	0.56	Bertola et al., submitted

* Contaminated samples were excluded from the ranking due to low coverage

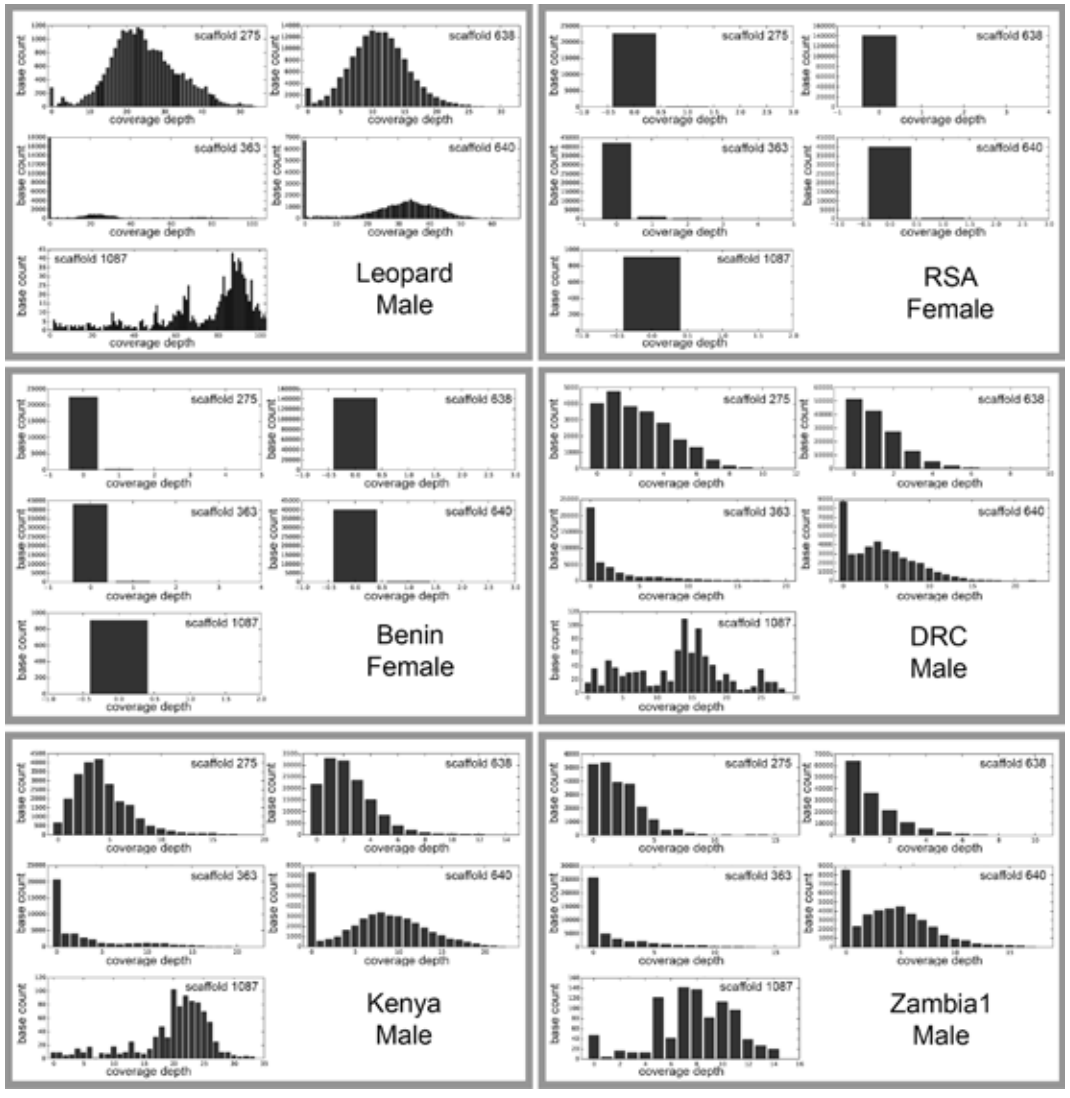


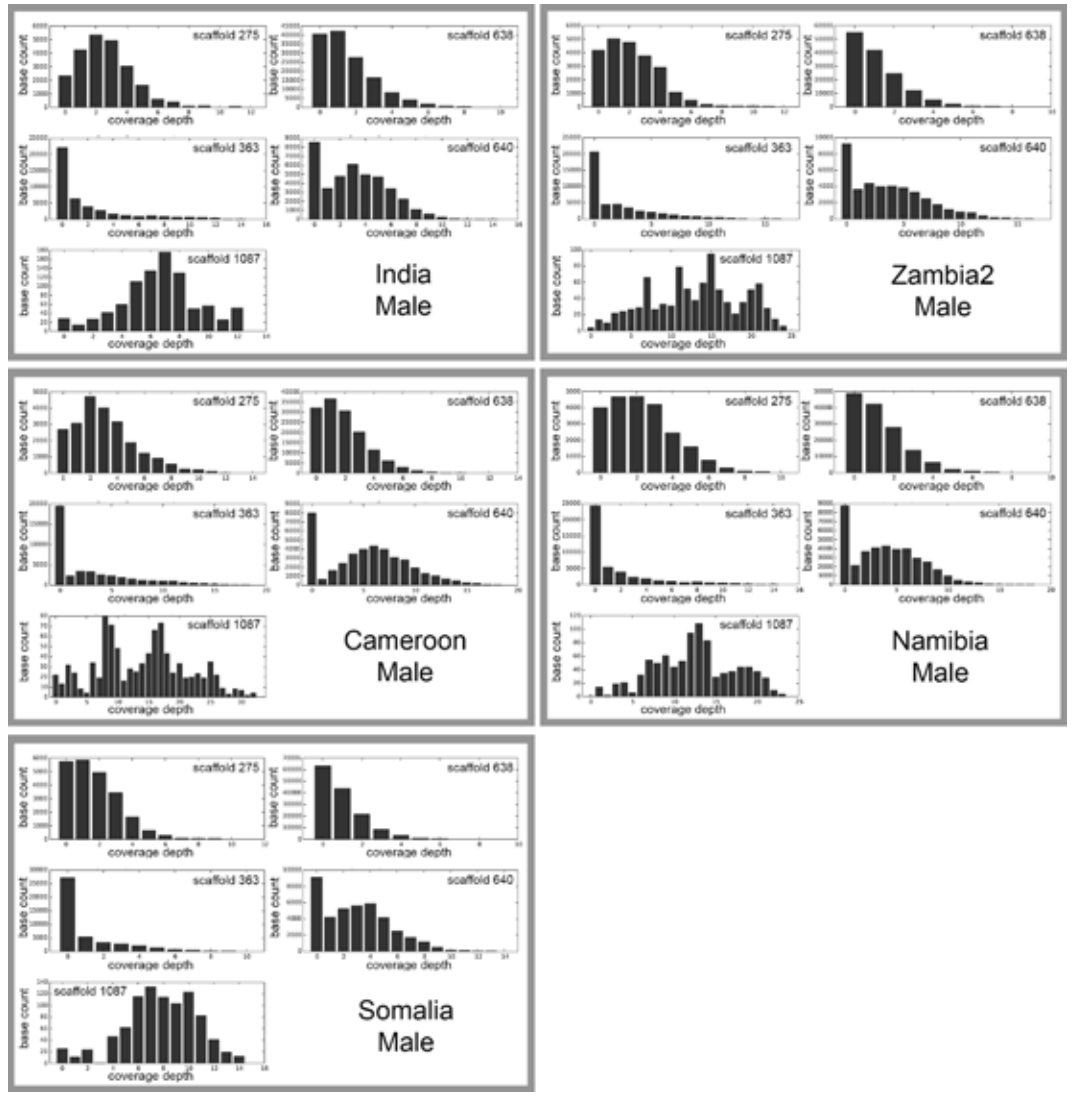


Supplemental Figure S1. Read quality derived from the first run, containing one leopard and two lion samples. Drop in quality scores for (A) Leopard, (B) Benin and (C) Kenya and quality after hard clipping of reads after 30 bp for (D) Leopard, (E) Benin and (F) Kenya.

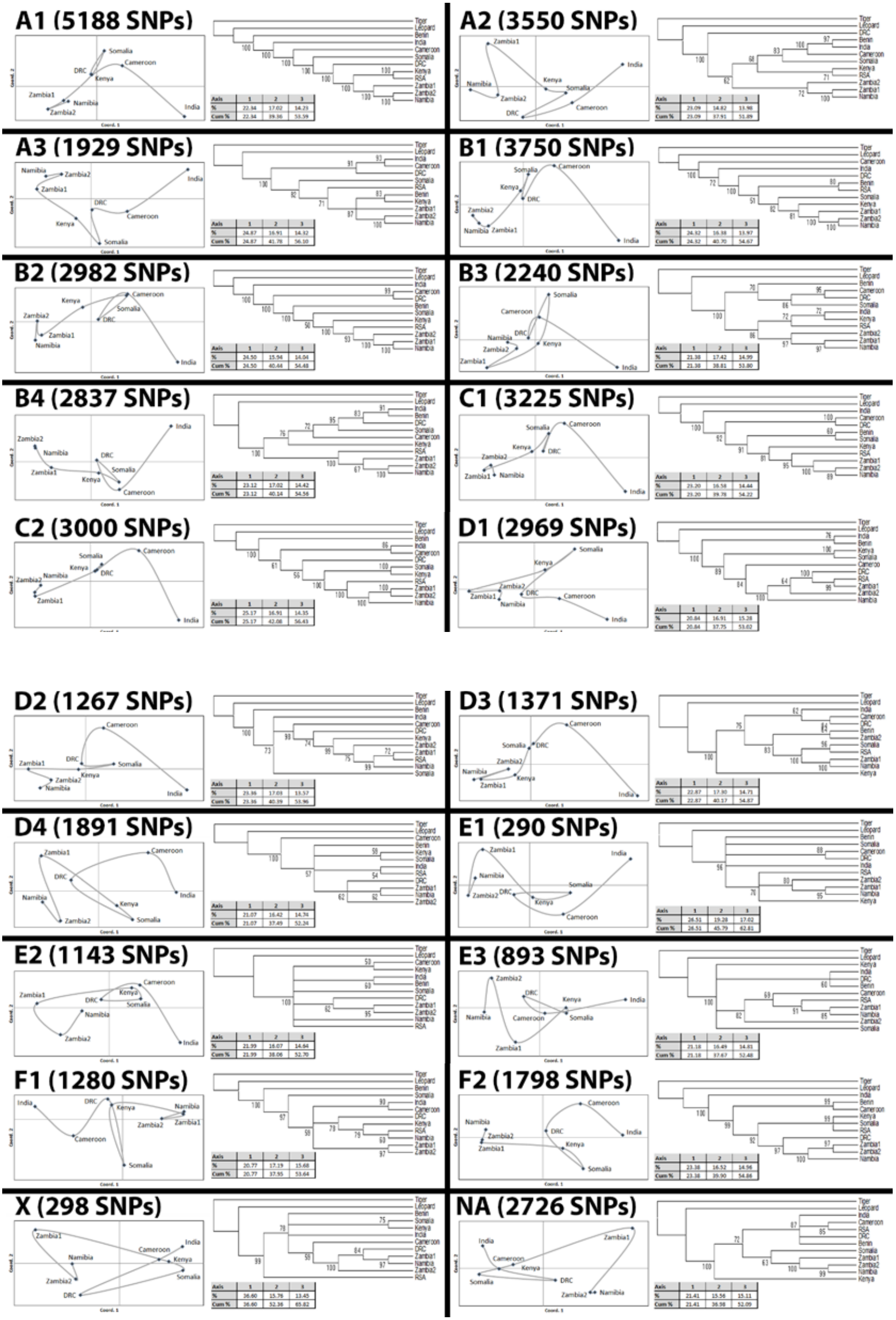


Supplemental Figure S2. GC content distribution for two lion samples showing signs of bacterial contamination. GC content of raw reads of (A) Benin and (D) RSA, (B+E) reads filtered against main contaminants and (C+F) reads aligned again the reference genome of the tiger.





Supplemental Figure S3. Coverage plots for one leopard and ten lions on five scaffolds that had been identified as having an Y-chromosomal origin.



Supplemental Figure S4. Bayesian analyses and PCA for individual chromosomes in the lion.

Chapter 6

General Discussion

General discussion

In this thesis the phylogeography and levels of genetic diversity are investigated for the lion. A number of genetic approaches is used and resulting data are discussed both in a phylogeographic and a conservation genetics framework. The former gives insight into the evolutionary history of the species, by illustrating which genetic lineages can be identified and how they are related to each other. The latter may contribute to species conservation by translation into recommendations for conservation management. This two-fold approach is using the current genetic makeup to infer scenarios from the past, but also to provide recommendations for the future.

Intraspecific genetic diversity of the lion

In order to obtain a complete overview of genetic diversity within a species, all main populations should be subjected to phylogenetic analyses. In the case of the lion, there is a special interest in the populations from West and Central Africa. Firstly, these populations have rarely been included in genetic analyses, and information regarding their position within the phylogenetic tree of the species is sparse. Secondly, taxonomic and phylogenetic data from other mammals, as well as data on the climatic history of West and Central Africa, suggest that evolutionary forces may have differentiated populations in this part of the continent from population in East and Southern Africa. Thirdly, given the vulnerable position of the small and isolated populations in West and (part of) Central Africa, these populations are in strong need of conservation action if we want to safeguard their future. Therefore, there has been a focus on populations from West and Central Africa in this thesis.

In Chapter 2, two previously used mtDNA markers, cytochrome b and part of the control region, were combined, enabling the integration of newly acquired data with sequences available on GenBank (Dubach *et al.* 2005; Barnett *et al.* 2006b). This was an important extension of existing datasets, since Dubach *et al.* (2005) had not included any population from West or Central Africa, thereby omitting an essential part of the lion range, and the genetic marker used by Barnett *et al.* (2006) gave limited insight into the relationships between genetic lineages. Phylogenetic trees and haplotype networks from Chapter 2 illustrate the close genetic relationship between lions from West and Central Africa and the Asiatic subspecies. The distribution of genetic diversity within the lion implies that the current taxonomy, identifying solely an African and an Asiatic subspecies, does not sufficiently reflect the intraspecific diversity of the lion, and the position of the Asiatic subspecies is questioned.

However, the position of the West African population in the northern branch remained unresolved. Also the positions of the three genetic lineages identified in East and Southern Africa were not completely resolved, and therefore their connection to the northern branch was still unclear. Trees based on exclusively the control region are of limited use, due to the lack of structure, and at this point in time, cytochrome *b* data were missing from East Africa, with the exception of Kenya and two captive populations from Ethiopia and Somalia. A denser sampling and possible identification of intermediate haplotypes may illustrate whether the long lineages identified in Chapter 2, notably in East and Southern Africa, are the result of sampling strategy, or actually reflect the evolutionary history of these populations.

To assess if the pattern discovered in the mtDNA phylogeny is also visible in autosomal data, 15 populations, representing the main genetic lineages identified in Chapter 2, were subjected to microsatellite analyses in Chapter 3. Data from the literature (Driscoll *et al.* 2002; Bruche *et al.* 2012) were combined with five populations from West and Central Africa, which were missing in earlier publications. In addition, a population intermediate to East and Southern Africa, located in Zambia, was added to minimize clustering of sampling locations. Results indicated three clusters in Africa, corresponding to 1) West and Central Africa, 2) East Africa, and 3) Southern Africa, in addition to the Asiatic subspecies. This showed that the intraspecific genetic diversity in the African lion is confirmed by autosomal data.

In Chapter 4, the mtDNA dataset was extended further by more sampling locations, including museum specimen from regions where lions are currently extinct (i.e. North Africa and Middle East) or from areas from which it was not possible to obtain samples from free-ranging lions (e.g. Angola). In addition, cytochrome *b* sequences for previously processed samples by Barnett *et al.* (2006) (Barnett *et al.* 2014), permitted denser sampling, notably in North Africa and the Middle East. Based on a preliminary topology, a total of fourteen samples were selected from the main haplogroups and the complete mitochondrial genomes were sequenced. Phylogenetic trees based on the new, extended dataset display a strongly supported dichotomy, differentiating lions from the northern part of their range (West and Central Africa, and North Africa/Asia), and lions from the southern part of their range (East and Southern Africa). Six main phylogeographic haplogroups are distinguished: 1) West Africa, 2) Central Africa, 3) North Africa/Asia, 4) North East Africa, 5) East/Southern Africa, 6) South West Africa. These results provide a better resolved and more strongly supported phylogenetic tree, and do not contradict the clades identified in other phylogeographic studies on lions, with the exception of the location of individual samples of which the genetic integrity could be questioned (i.e. captive individuals supposedly from Angola or Zimbabwe, included in Antunes et al. (2008)) (Dubach *et al.* 2005, 2013; Barnett *et al.* 2006a; b, 2009, 2014; Antunes *et al.* 2008). This further confirms that current lion taxonomy does not only underestimate intraspecific diversity in the lion, but that the distinction between an African and an Asiatic subspecies is not in line with the most basal split found in the phylogenetic tree of the lion.

Chapter 5 contains phylogenetic trees based on 44,627 SNPs, identified by full genome sequencing of 10 lions, representing the main haplogroups, one leopard and comparison to the reference genome of the tiger (Cho *et al.* 2013). The resulting trees show a hierarchical pattern, an fail to recognize reciprocally monophyletic clades. However, the trees do not contradict previously described patterns based on mtDNA or microsatellite data, and the nested position of the Asiatic subspecies seems to confirm that the current taxonomic distinction is not in line with the evolutionary history of the lion.

Evolutionary history of the lion

Based on the absence of long mtDNA lineages in the northern part of the lions range, it was hypothesized in Chapter 2 that lions in West and Central Africa may represent a relatively recent recolonization of the area after local extinction. Climate data indicate that West and Central Africa

was characterized by severe aridity, which may have reduced the number of prey and subsequently the number of lions. The close genetic relationship between West and Central African lions and the Asiatic subspecies suggests that source populations for recolonization may have been in close geographic proximity to North Africa/Asia.

In Chapter 4, the most recent common ancestor of the two major lineages was estimated at ~300 thousand years ago (kya), and major radiation of haplogroups during the last c. 100 thousand years (kyr). In combination with a review of data on the climatic history of the African continent, this enabled the reconstruction of a possible scenario for lion evolution. Two major vegetation zones, dry desert and dense rain forests, representing hydrological extremes, may have contributed to the differentiation between genetic lineages in the lion. Populations may have exhibited contractions to regional refugia when suitable habitat was reduced, and expanded after more favourable changes in the climatic conditions. Last coalescence between the North and South groups coincides with the expansion of dense rain forest along an east-west axis in lower latitude Africa, which may have hampered gene flow between these two major genetic lineages. North-south expansions of the Sahara desert coincided with the major splits in the northern range of the lion, although a connection between North Africa and Central Africa may have persisted during short periods that the monsoon front reached high latitudes, explaining their close genetic relationship.

The haplogroups identified in the lion are further congruent with patterns described in other species, based on taxonomy and/or phylogenetic datasets. For large savannah mammals with a similar range as the lion, numerous species show the distinction between populations in West and Central Africa, and populations in East and Southern Africa (see Table 1 and Figure 1 in Chapter 4). This suggests environmentally driven evolution and possibly common refugia for a range of co-occurring species. Based on phylogeographic patterns in large mammals three major refugial areas have been suggested, being West/Central Africa, East Africa, and Southern Africa (Hewitt 2004; Lorenzen *et al.* 2012). This is in line with main haplogroups described in the lion. A more detailed picture arises from a study using bioclimatic envelope models, describing five possible refugia in sub-Sahara Africa: one in Upper Guinea, one or two in the Cameroon Highlands – Congo Basin, one in the Ethiopian Highlands, one in Angola-Namibia, and one in East/Southern Africa (Levinsky *et al.* 2013). These geographic locations are congruent with the five main groups identified in the lions: West Africa, Central Africa, North East Africa, South West Africa, and East/Southern Africa, respectively (Figure 1). Since delineation of the haplogroups, as shown in Figure 1, is inferred from available sampling locations and current lion range, there may not be a complete overlap with refugial areas as indicated by the model (e.g. in Central Africa). However, as lion range may have shifted through the course of the history, we still conclude that there is a good fit to the proposed refugial areas. The congruence between the lion data presented in this thesis, patterns in other species, and climate data further corroborate the distribution of intraspecific genetic diversity found in the lion.

Genetic diversity and bottlenecks

Due to habitat fragmentation and strong lion population declines in West and (part of) Central Africa, it was hypothesized that these populations would be particularly vulnerable to declines in genetic diversity. Previous studies have shown declines and local extinctions of lion populations in this region

(Riggio *et al.* 2012; Henschel *et al.* 2014), a trend also documented in other species (Craigie *et al.* 2010; Bouché *et al.* 2012). However, analyses of 20 microsatellite loci, presented in Chapter 3, did not show significant heterozygote deficiencies, reduced number of alleles or fixed loci in any of the six sampled populations in this region.

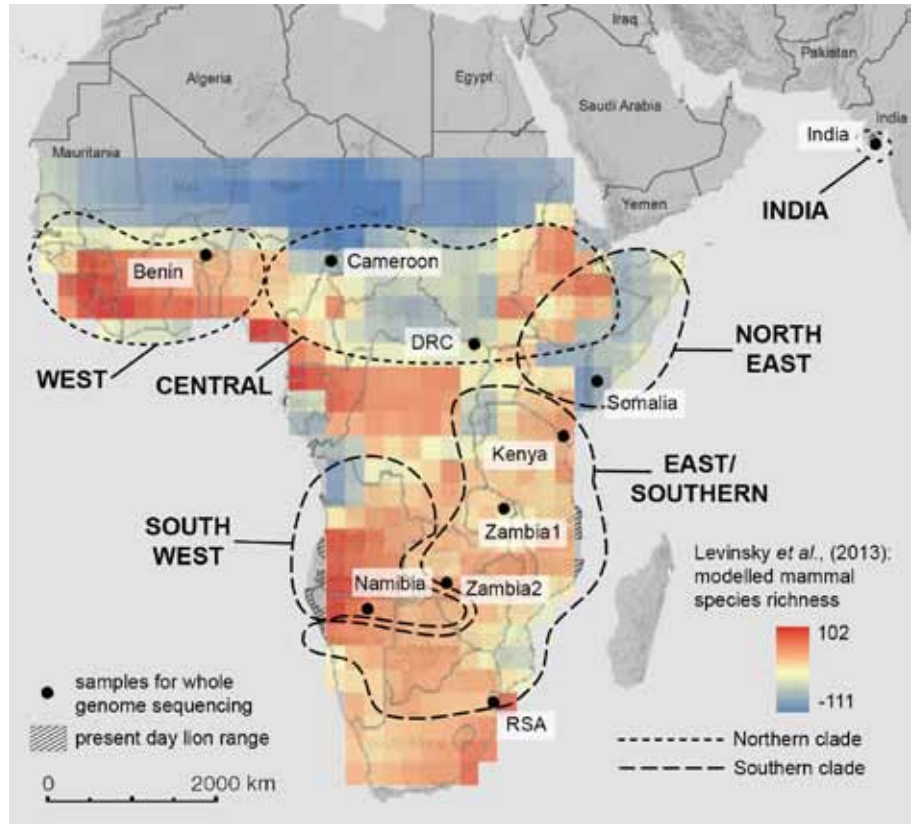


Figure 1. Identified haplogroups in the lion based on mtDNA data, and modelled Last Glacial Maximum species richness for mammals endemic to sub-Saharan Africa, based on Levinsky *et al.* (2013). Lion range data are from IUCN (2014).

Fis values, indicating the inbreeding coefficient of an individual relative to the subpopulation, presented in Chapter 3, were significantly positive in two populations: RSA1 (Kalahari-Gemsbok NP) and Ethiopia1 (Addis Ababa Zoo) (excluding the significantly positive Fis value found in Zambia, possibly due to a Wahlund effect). It was previously shown that RSA1 contains relatively low levels of genetic diversity (Dubach *et al.* 2013), and that the Ethiopia1 is based on a very limited number of founders (Bruche *et al.* 2012). Interestingly, in Chapter 3 significantly negative Fis values were found for the populations Benin (Pendjari NP) and Cameroon 1 (Waza NP), indicating an excess of heterozygotes. A possible explanation would be the mating system, or that these populations have been under pressure, e.g. by hunting in adjacent hunting zones, or due to recent declines (Iongh *et al.* 2009; Riggio *et al.* 2012), and there has been a selection for individuals with high levels of genetic diversity, i.e. heterozygosity. Bottleneck analysis further confirmed the low genetic diversity

in RSA1 and Ethiopia1. In addition, Cameroon1 (Waza NP) and Kenya (Amboseli NP) were identified as having experienced a recent population reduction by the program Bottleneck, consistent with observations obtained from monitoring studies (Iongh *et al.* 2009; Riggio *et al.* 2012). In Chapter 5, observed heterozygosity values, based on 18,457 SNPs, showed that ranking of the populations was congruent with previous ranking based on microsatellite data. Although for SNP analyses a single individual was included for each population, this approach follows the "few individuals, several genes" notion, as was mentioned in Chapter 1 (McMahon *et al.* 2014).

Main reason for the absence of clear signs of a reduction of genetic diversity in West and Central African lion populations, might be that range contraction and population declines are too recent to be visible as a genetic signature. In addition, active inbreeding avoidance, such as selection for the most heterozygote individuals as breeders, as was previously shown in wolves (Bensch *et al.* 2006), may contribute to a relatively high diversity in small populations with limited levels of gene flow. However, the significantly negative Fis values may be a first sign of high pressures on these lion populations. Since genetic diversity is rapidly lost in small and isolated populations, caution is warranted and inbreeding may become a problem in the near future. Management interventions may be needed to safeguard these populations against detrimental genetic effects and subsequent declines in fitness in the future.

The genomic complexity of intraspecific genetic diversity in the lion

As different genetic markers represent different evolutionary trajectories (mtDNA vs. autosomal DNA) and due to stochasticity in the coalescence of markers (Edwards 2009; Knowles 2009), species histories should preferably be based on data from multiple, unlinked loci. Many studies use mtDNA markers to make phylogeographic inferences (Arctander *et al.* 1999; Flagstad *et al.* 2001; Moodley & Bruford 2007; Zinner *et al.* 2009; Gaubert *et al.* 2011; Haus *et al.* 2013). Due to the high copy numbers per cell, mtDNA is relatively easy to amplify for sequencing. Therefore it is a useful marker when using samples in which DNA content is low and the DNA is fragmented, such as scat samples or museum specimens. Drawbacks of mtDNA include the fact that it consists of a single locus and that it represents the maternal lineage only. Despite of this, it may provide useful first insight into phylogeographic patterns (Zink & Barrowclough 2008). Since mtDNA is typically inherited through the mother and does not recombine, its effective population size (N_e) is ¼ of the N_e of autosomal markers. Lineage sorting is therefore more rapid, leading to a structure sometimes not (yet) visible in gene trees based on autosomal data.

Autosomal data can be used to complement phylogeographic patterns derived from mtDNA. Microsatellites are useful markers to infer population structure and within-population diversity. Frequency-based approaches (such as STRUCTURE) cluster individuals and provide insight into admixture between the identified clusters. Reticulate events, such as hybridization between clusters or clades, is usually not allowed for in phylogenetic trees, posing a problem when studying intra-specific phylogenetic relationships. However, the phylogenetic relationship between clusters based on microsatellite allele frequencies is difficult to infer. Their mutation pattern, the unit of analysis (typically a pooled sample of individuals, rather than a single haplotype) and difficulties regarding rooting of a microsatellite-based tree, make them less suitable for larger scale phylogeographic questions (Zink & Barrowclough 2008). Despite of this, including microsatellite data may identify cases in which mtDNA and autosomal data show conflicting patterns, e.g. if populations would be linked by nuclear, but not by organellar gene flow (i.e. female gene flow much less than male gene flow). For the reconstruction of evolutionary histories and inferring a phylogenetic tree, genome-wide nuclear markers would be favourable over frequency-based approaches, such as microsatellites.

In this thesis, mtDNA, microsatellites and genome-wide SNP data are included to gain insight into the intraspecific genetic diversity of the lion. Based on mtDNA data, presented in Chapter 4, six reciprocally monophyletic haplogroups can be described: 1) West Africa, 2) Central Africa, 3) North Africa/Asia, 4) North East Africa, 5) Southern/East Africa and 6) South West Africa. Including complete mitochondrial genomes leads to a well supported basal dichotomy, in which the first three haplogroups are placed in the northern clade, and the remaining three haplogroups are placed in a southern clade. The nested position of the Asiatic subspecies in the northern clade led to the notion that the current nomenclature, only recognizing an African and an Asiatic subspecies, does not only underestimate the complete genetic diversity in the species, but also does not follow the most basal genetic differentiation in the species. The only region in which we find evidence for natural admixture based on regional co-occurrence of haplotypes from different haplogroups is Ethiopia, where haplotypes from Central Africa and from North East Africa co-occur. Although the Rift Valley has often been mentioned as a barrier for lion dispersal (Burger *et al.* 2004; Dubach *et al.* 2005; Barnett *et al.* 2006b, 2014), these data indicate that gene flow across the Rift Valley does exist. A mixture of haplotypes is further identified in the Kruger NP/Limpopo area in RSA, likely to be the result of human-mediated translocations as Etosha NP is known to have been a source for several reintroduction projects (Miller *et al.* 2013).

In Chapter 3, microsatellite data were analyzed for fifteen populations, representing the complete geographic range of the lion. Four clusters are recognized: 1) West and Central Africa, 2) East Africa, 3) Southern Africa, and 4) Asia. This illustrates that the genetic structure of the lion within Africa is confirmed by an autosomal marker. However, the relationship of the West and Central African lion to the Asiatic subspecies is difficult to infer from these data. The high level of fixation of alleles in the Asiatic lion is contributing to its distinct status, which is not necessarily a reflection of an ancient evolutionary split, but rather a result of severe bottlenecks and isolation. Although both Ethiopian populations included in the microsatellite analyses, were captive populations, microsatellite data confirm admixture, as was seen based on mtDNA data. The only other population with evident admixture based on microsatellite data, is Zambia. The fact that all included Zambian individuals contain a haplotype from the same haplogroup (East/Southern), suggests that the admixture pattern is the result of male-mediated gene flow. Lions are known to exhibit sex-biased dispersal, in which males leave their natal pride and tend to move further than females (Pusey *et al.* 1987; Spong & Creel 2001). Haplotypes from the East/Southern Africa haplogroup are found from Kenya southwards, across populations in Zimbabwe, Botswana, RSA and Central Namibia, with the exception of the mosaic pattern in the Kruger NP/Limpopo area mentioned above. However, based on microsatellite data, both included RSA populations (Kalahari-Gemsbok NP and Kruger NP) show a strong assignment to the same clusters as Namibia, representing a South West haplotype. This may be the result of the homogenizing effect of male-biased dispersal, whereas the Zambian population may represent the fringe at which admixture in the autosomal data is still evident.

The phylogenetic tree based on 18,457 SNPs, discovered in Chapter 5, shows a hierarchical topology, and fails to identify reciprocally monophyletic clades. The consecution in which the individuals branch

off suggests that it reflects continent wide gene flow in a north-south direction. Again, India is nested within the African populations, close to West and Central Africa, thereby provoking the validity of the subspecies status of the Asiatic lion. The loop that is formed between the individuals from DRC, Kenya and Somalia in the PCA, may be an indication of admixture, as we see on a smaller geographic scale in mtDNA and microsatellite data. The position of RSA2 in the PCA and SNP tree should be interpreted with caution, since this is an individual from RSA with a Namibian haplotype, and therefore it is likely that hybridization between two haplogroups has occurred.

Discrepancies between phylogenies based on mtDNA and autosomal data can be explained by lineage sorting times, which is four times shorter for mtDNA. This may lead to discrete, monophyletic groups, which are not retrieved from autosomal data. Therefore, Moritz (1994) notes that requiring reciprocal monophyly for both nuclear and mtDNA markers for the recognition of an ESU seems overly restrictive. However, nuclear allele frequencies should be diverged, to avoid misclassification of populations that are linked by nuclear, but not by organellar gene flow (Moritz 1994). This divergence is visible in the microsatellite data, distinguishing three African clusters. The fact that these clusters are not recognizable in the SNP tree, can be explained by the different mode in which these markers establish their diversity. The mutation rate in multi-allelic microsatellites is much higher than in bi-allelic SNP markers. Therefore the tree based on the SNP data may represent a more ancient pattern of continent-wide gene flow, before populations became isolated in refugia as discussed above. The combination of three types of genetic markers, reveal the underlying complexity of the intra-specific genetic diversity in the lion (Figure 2).

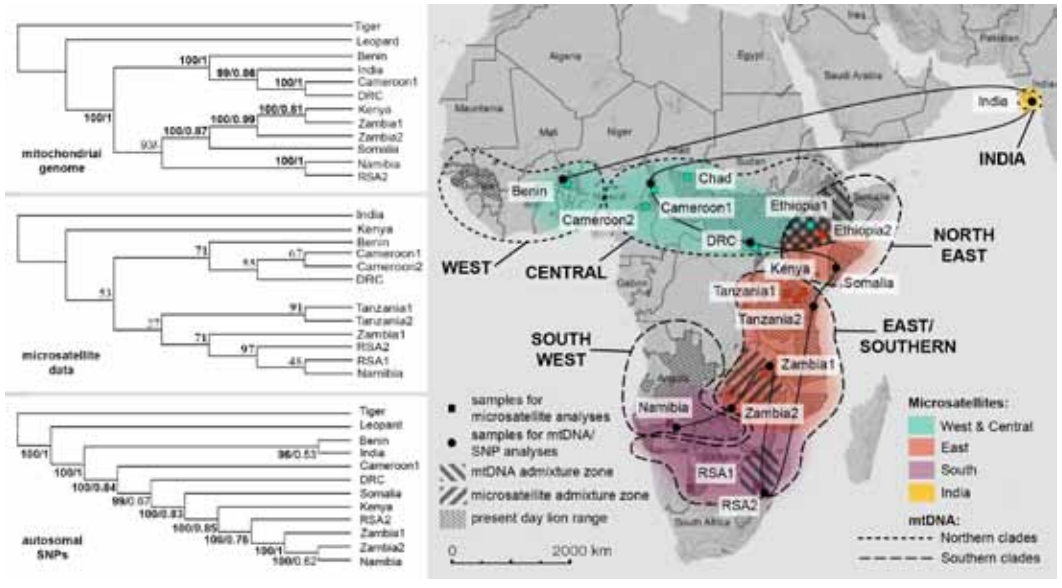


Figure 2. Overview of phylogeographic data from complete mitochondrial genomes, 20 microsatellite loci and 18,457 autosomal SNPs. Support is indicated as posterior probability (Bayesian analysis)/bootstrap support (ML analysis) for mtDNA and SNP trees, the microsatellite tree is based on Da genetic distance and 100 bootstrap replicates. Chad and the Ethiopian populations have not been included in the tree due to missing data, but were added to the map to illustrate admixture patterns. Delineation of haplogroups is based on sampling locations as indicated in Chapter 4. Lion range data are from IUCN (2014).

Implications for taxonomy and management

The classification of populations into subspecies, should ideally be based on the evolutionary history of the species, since this reflects the biological background of the distinguished units. Including phylogeographic data may be a useful tool to assess this. In the case of the lion, the two officially recognized subspecies, the African and the Asiatic lion, do not represent reciprocally monophyletic clades, based on mtDNA or autosomal data. The deepest split in the haplotype tree makes a distinction between lions from the northern part of the range, including the Asiatic subspecies, and lions from the southern part of the range. Admixture between these groups is noticeable in the connecting zone, represented by sampling locations in Ethiopia. Recent hybridization, however, should not be an argument to not recognize these groups as distinct units for conservation, as was previously noted by Moritz (1994). Following the data presented in this thesis, a taxonomic revision would therefore result in two subspecies: a northern subspecies, enclosing West Africa, Central Africa, and India, and a southern subspecies, enclosing North East Africa, East/Southern Africa, and South West Africa.

Propositions for taxonomic revision based on phylogeographic data have also been made by other research groups: Dubach et al. (2013) argues that the limited population size and its unique claim to being Asia's only surviving lion population warrants the distinct subspecies status for the Indian population, in addition to a northern and a southern subspecies. Barnett et al. (2014), however, reasons that subspecies distinctions should follow the recognized phylogeographic groups. This would lead to either 1) elevation of the five African haplogroups to distinct subspecies, or 2) to clustering the Asiatic lion with populations from West and Central Africa as one subspecies, and populations from the southern clade as the second subspecies. Although there may be arguments to retain the subspecies status for the Indian population, based on the ecological and geographic position, this would entail that one of the other subspecies is a paraphyletic group. The alternative, to upgrade all identified haplogroups to subspecies status, is not desirable for two main reasons. Firstly, autosomal data do not show significant divergence of allele frequencies, following the proposed boundaries of the haplogroups, notably in East and Southern Africa. Recognizing the three haplogroups in this area as distinct subspecies may be overly sensitive, as the mtDNA pattern may represent a relic of historic isolation, only sustained by strong female philopatry. Secondly, taxonomic inflation, leading to an increase in (sub)species numbers by splitting existing taxonomic entities, may not benefit conservation (Isaac *et al.* 2004). Therefore it is proposed to follow the deepest split in the haplotype tree for recognizing two lion subspecies, and additional haplogroups for the recognition of ESUs, sensu Moritz (1994). It must be noted that listing activities should be assessed independently from recovery planning, as different scales and units may be appropriate for each of these processes (Mace 2004; Isaac *et al.* 2004).

Comparing the depth of this split with dated splits in other large cats shows that the coalescence age does not provide a clear guideline for taxonomy (Table 1). Notably the coalescence of six subspecies of tigers is relatively recent (Luo *et al.* 2004), whereas in the jaguar no subspecies are recognized but splits are much deeper (Eizirik *et al.* 2001). In leopard, a deep split is found distinguishing African from Asiatic leopards, in which subsequently seven Asiatic subspecies are recognized (Uphyrkina *et al.* 2001). Comparison of these species shows that the depth of the split found in the lion (this thesis) would not be an argument to consider all African lions as a single subspecies.

Table 1. Estimated coalescence times in species from the genus *Panthera.*

Species	Split	Coalesence data	Source
Lion	North (incl. Asiatic subspecies) – South	291,700 ya	This thesis
(*Panthera leo*)		(95% HPD 178,000-417,700)	
Leopard	Africa (incl. South Arabia) – Asia	471,000 ± 102,000 ya	Uphyrkina *et al.* (2001)
(*Panthera pardus*)	7 Asiatic subspecies	169,000 ± 49,000 ya	
Tiger	6 subspecies	72,000 ya	Luo *et al.* (2004)
(*Panthera tigris*)		(95% CI 39,000–104,000)	
Jaguar	1 subspecies, several lineages	280,000–510,000 ya	Eizirik *et al.* (2001)
(*Panthera onca*)		(95% CI 137,000–830,000	

In addition to the distribution of genetic diversity within the species, congruence is found with morphological characters and the historical distinction of subspecies (Hemmer 1974; Mazák 2010) (Figure 3). Up to eight "subspecies" are recognized by some sources (Haas *et al.* 2005), with the Barbary lion (*P. l. leo*) very likely to be extinct and the Cape lion (*P. l. melanochaita*) a possible con(sub) specific with *P. l. krugeri* (Barnett *et al.* 2006a). Hemmer (1974) further suggests to not include *P. l. bleyenberghi* and *P. l. azandica* as distinct subspecies, since differentiation based on morphometric characters is not conclusive. Therefore Hemmer (1974) suggests to follow a basal dichotomy with a distinction between *P. l. senegalensis* and *P. l. persica* in the northern range, and between *P. l. nubica* and *P. l. krugeri* in the southern group. The main discrepancy between genetic data and proposed subspecies delineations in Figure 3, is the border between the West and Central African lion and the East African lion. However, it was not possible to defer information on which this delineation was based, and the illustrations by Hemmer (1974) seem to suggest that the boundary between *P. l. senegalensis* and *P. l. nubica* is located further to the East, more in line with delineation found based on mtDNA haplotypes and miscrosatellite data. The proposed West and Central African identity of individuals in Congo could not be confirmed by an earlier phylogeographic study, in which an included individual from Gabon contained a haplotype widespread in southern Africa (Barnett *et al.* 2006a). In summary, the combination of genetic and morphological data suggest that the current taxonomy, only recognizing the African and the Asiatic lion as distinct subspecies, does not reflect the diversity in the lion. A revision of the nomenclature is suggested in which all populations from the northern part of the range, encompassing West Africa, Central Africa, and India are considered one subspecies, and populations from the southern part of the range, including North East Africa, East/Southern Africa, and South West Africa, are recognized as the second subspecies. The six mentioned lineages should be managed as ESUs, sensu Moritz (1994). We propose the adoption of the names *Panthera leo leo* (Linnaeus, 1758) for the northern subspecies, and *Panthera leo melanochaita* (Smith, 1848) for the southern subspecies, in accordance with the rule of nomenclature (http://iczn.org/iczn/index.jsp). This revision has immediate implications for conservation. Recognizing the northern clade as *P. leo leo*, eliminates separate taxonomic listing for Asiatic lions, but both Asiatic and West African units are considered to be endangered (Breitenmoser *et al.* 2008; Bauer *et al.* 2012; Henschel *et al.* 2014). In West and Central Africa lion numbers are strongly declining (Riggio *et al.* 2012; Henschel *et al.* 2014), along with other species (Craigie *et al.* 2010; Bouché *et al.* 2012). No strongholds (sensu Riggio *et al.* 2012) are identified in this region, whereas several strongholds are recognised within the range of *P. l. melanochaita*. Because West and Central Africa harbour unique genetic lineages, in the lion, as well as in other species, conservation of populations in this region is of utmost importance.

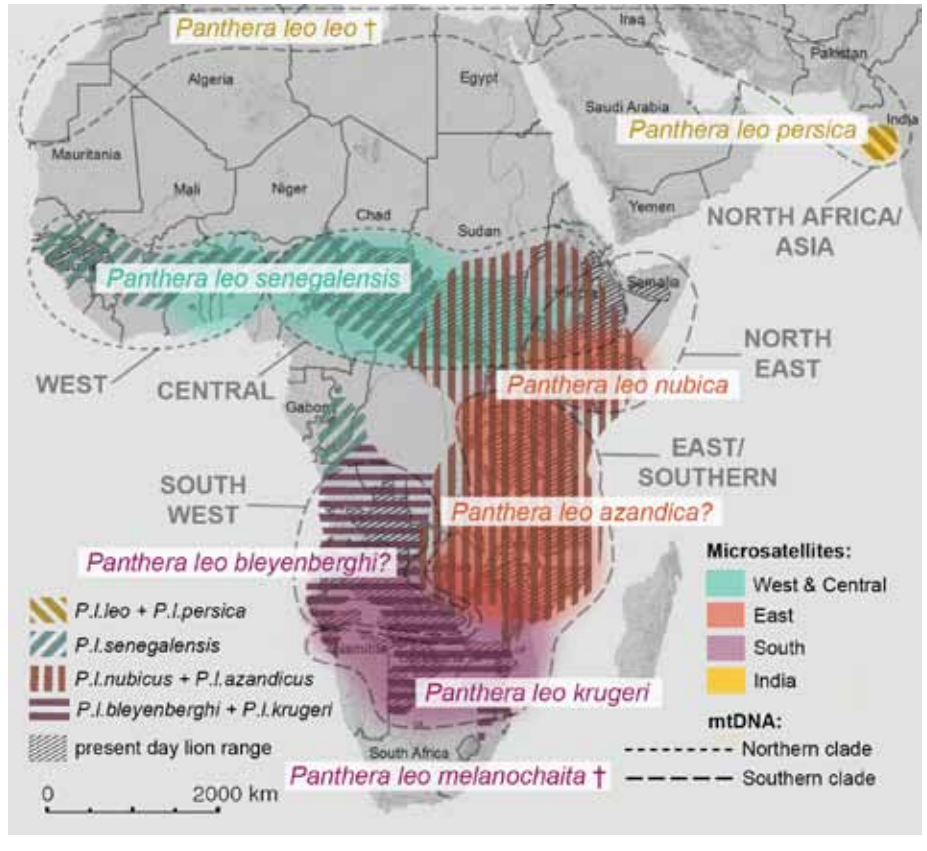


Figure 3. Lion subspecies delineation, based on Haas (2005), with haplogroup delineation and microsatellite clustering indicated by shading. SNP data have been omitted to improve readability. *P. l. azandica* and *P. l. bleyenberghi* are indicated with a '?', following suggestions by Hemmer (1974). Lion range data are from IUCN (2014).

The recognition of units below the subspecies level, such as ESUs and MUs, may be beneficial for management interventions, including translocations, reintroductions and reinforcements of extant populations. Although hybridization of individuals from different phylogenetic lineages may be useful to counteract the effects of inbreeding, it is desirable to maintain the integrity of phylogenetic lineages. Disruption of local adaptation (possibly leading to outbreeding effects) and disturbing natural patterns of genetic diversity within the species by human interference should be avoided, if possible. When considering conservation interventions, it would be advisable to follow the assignment to a haplogroup as a first indication, although admixture, not visible in mtDNA data, may play a role at boundaries of identified haplogroups. For captive-bred lions, often with limited information regarding their free-ranging ancestors and breeding history, assignment to a haplogroup may be less useful, because of the frequent occurrence of admixture in a zoo setting (ISIS 2014) (see 'Captive stocks and implications for breeding'). In general, when assessing source populations for a translocation, it would be advisable to prioritize populations from the same haplogroup, in close proximity to the destination, from which off-take is sustainable and does not disrupt social structures. This is of exceptional importance for countries that harbour lion populations from

different haplogroups, e.g. Nigeria and Kenya. In the case of Nigeria, it would not be advisable to translocated individuals between Yankari GR and Kainji Lake NP, but to target populations from the same haplogroup in neighbouring countries as a possible source.

Based on the distribution of genetic diversity and the conservation status of the lion in some regions, certain populations should be prioritized in order to maintain the genetic diversity in the African lion. Following from the vulnerable position of lions in West and Central Africa and the recognition of potential strongholds, sensu Riggio *et al.* (2012), a pragmatic view should be employed. Prioritization of conservation projects in Niokolo-Giunea (Senegal – Guinea Bissau – Guinea – Mali) and W-Arly-Pendjari (Benin – Burkina Faso – Niger) and the Bénoué complex-Gashaka Gumti (Cameroon – Nigeria – Chad), as well as Chad-CAR, may provide a starting point, following recent assessments (Riggio *et al.* 2012; Iongh *et al.* 2014; Henschel *et al.* 2014). Transboundary conservation agreements may be challenging, but also have great potential to involve multiple partners and to address conservation practices on a meso-scale (Sodhi *et al.* 2011).

Although in a few cases reduced genetic diversity is recognized, there is no clear evidence for inbreeding depression in any of the free-ranging lion population included in this study. Despite of this, on the longer term management interventions may be needed to maintain genetic diversity in the small and isolated populations. Assignment of populations to genetic lineages may be helpful to guide these interventions. However, deciding on when a conservation action such as the reinforcement of an existing populations is justified, should best be confirmed by monitoring of demographic factors, such as fecundity and mortality. Certainly, it should be ensured that the pressures, driving the original population to decline or even extinction, are no longer present. In addition, it must be noted that in the consideration of a reintroduction or translocation project, apart from the genetic requirements other biological, socio-economic and cultural aspects need to be taken into account (IUCN SSC 2013).

Captive stocks and implications for breeding

Due to the challenges linked to collecting samples from different source populations, some studies partially relied on captive individuals (Antunes *et al.* 2008; this thesis). Antunes *et al.* (2008) included two captive individuals supposedly from Angola (individuals from the same stock also included in Chapter 2 of this thesis) and one individual supposedly from Zimbabwe. These samples clustered close to Central Africa in their analyses. This is not in line with previously published phylogeographic patterns in lions, leading to questions regarding the origin and breeding history of these individuals. In the case of the Angola lions, a breeding history is published in a study on lens-anomalies in an inbred zoo population (Steinmetz *et al.* 2006). A pure-blooded Angolan origin cannot be confirmed and hybridization with individuals with a Central African haplotype may have occurred. Another well-known example is the recently extinct lion subspecies, the Barbary lion (*P. l. leo*), which was last sighted in 1942 in the western Magreb (Black *et al.* 2013). Several zoos still claim to have a breeding stock, and researchers have called for further captive breeding to avoid definitive extinction (Burger & Hemmer 2005; Black *et al.* 2009). However, a comparison of a part of the mtDNA of five captive Barbary lions to ancient specimen from North Africa held in natural history collections, revealed that none of the captive lions were maternally of Barbary origin (Barnett *et al.* 2006a). Instead, the identified haplotype was also found in lions from Central African Republic (CAR) and Democratic Republic of the Congo (DRC), proposing a Central African origin, at least for the maternal lineage (Barnett *et al.* 2006a).

Comparison of captive populations to an incomplete reference set, may pose another difficulty. Based on the phenotypical distinctiveness of some Ethiopian lions, including the heavily and dark-maned individuals held in the zoo of Addis Ababa, microsatellite and mtDNA data were used to assign these individuals to a reference dataset (Bruche *et al.* 2012). It was concluded that the Addis lions compose a unique, previously unidentified clade. However, the used reference set included solely populations from East and Southern Africa and therefore only contained a part of the genetic diversity found in the African lion. In Chapter 3 and 4, samples of the Addis lions were re-analysed and compared to a larger dataset, including populations from West and Central Africa (Chapter 3 of this thesis). Both mtDNA and microsatellite markers illustrate that the Addis lions are closely related to the West and Central African lion, a signature also found in part of the free-ranging Ethiopian lions.

Assessing genetic diversity, as well as the assignment to genetic lineages, may be helpful for the design of an effective breeding and management program for captive stocks. Current lion holdings (October 2014) show that a total of 2,095 individuals are registered in the International Species Information System (ISIS 2014). Following the putative designation into eight subspecies, 17% of the lions belong to the East and Southern Africa group, 16% represent the Asiatic subspecies, and a mere 4% are West and Central African lions (including the putative Barbary lions, genetically analysed by Barnett *et al.* 2006a). The remaining 63% of the individuals are known hybrids or have an unknown history. Origin and breeding history are often not well-documented or anecdotal, which may mean that the number and purity of genetic lineages present in the captive stocks are severely over-estimated (Bertola *et al.* 2011). In addition, these figures indicate that the West and Central African lion is severely underrepresented in the captive population.

Despite the subspecies designation followed by ISIS, there is no studbook for African lions in captivity. Asiatic zoo lions on the other hand, breed following a subspecies-specific studbook. A Species Survival Plan (SSP) was established in 1981 and was later assessed for its genetic purity (O'Brien *et al.* 1987). It was concluded that two of the seven founder lions used for this stock were of African origin and consequently all captive Asiatic lions are the result of admixture between the African and Asiatic subspecies (O'Brien *et al.* 1987). Even for the free-ranging Asiatic subspecies, the purity of the lineage has been questioned. According to some authors, African mammals were imported to restock royal menageries in India for at least five centuries (Thapar *et al.* 2013). The discussion whether individuals could have escaped their enclosures and populated the local habitat, leading to the "exotic alien" status of the Asiatic lion (and cheetah), or whether they could have hybridized with existing populations is still ongoing (O'Brien 2013; Packer 2013). However, phylogeographic data do not point into the direction of hybridization or replacement of the Asiatic lion by sub-Sahara African individuals (O'Brien 2013; this thesis). Because of the close genetic relationship between the Asiatic lions and the extinct North African/Middle East populations, there may have been gene flow in this region, although this could be both natural as human-mediated.

Different views on the management of a captive stock, poor bookkeeping, and the inclusion of confiscated individuals of unknown origin have led to admixture of different genetic lineages in the vast majority of zoo lions (ISIS 2014). Therefore, it is advisable to not give too much weight to the

subspecies designation as found in ISIS. However, even in the absence of pure genetic lineages, the captive population may still be valuable in terms of diversity. The strategy of a breeding programme could be to mimic the genetic diversity of the species, taken into account the available resources to house captive animals. Genotyping the captive stock and identifying a number of "zoo lineages" may contribute to the design of such a breeding programme. When lions are not bred with the ultimate goal to be used in reintroduction programmes, pure genetic lineages are less relevant. Reintroduction projects using captive individuals as a source are still highly controversial, and may not be a suitable conservation strategy for the near future (Hunter & Rabinowitz 2009). Notwithstanding, a certain level of vigilance is needed as several organizations state that they aim to reintroduce captive-bred lions into the wild, although release sites are often fenced areas (ALERT 2014; Vier Voeters 2014).

In the course of this research, five captive lions, suggested for a reintroduction project in Ghana, were analysed for their mtDNA haplotype. Since none of the individuals contained a haplotype from the West African haplogroup (Figure 4), it was advised to exclude them from reintroduction in Ghana. Since past hybridization events cannot be detected by mtDNA analyses, for the confirmation that an individual is genetically suitable for reintroduction in a specific area, mtDNA data should be combined with autosomal data.

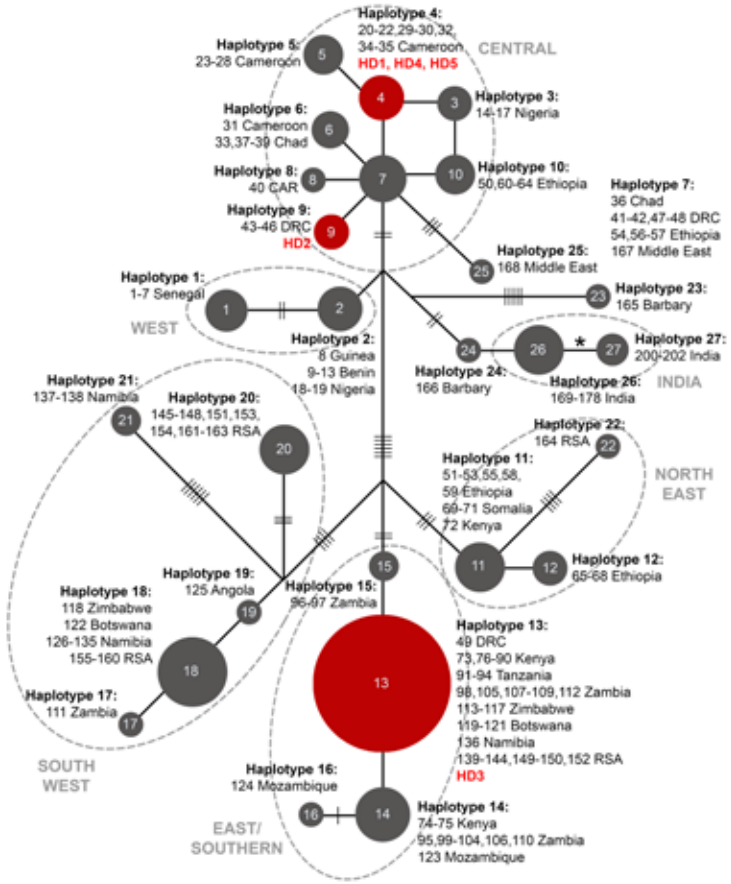


Figure 4. Phylogeographic origin for five captive individuals (HD1-HD5), compared to haplotype data from Chapter 4.

Using a similar forensics approach, as described for the captive individuals above, sources of illegal lion products may be identified, as was previously done for confiscated shipments of elephant tusks (Wasser *et al.* 2008). For a lion skin, confiscated in Libreville, Gabon in 2011, the haplotype was analysed in the course of this project, pointing towards a Central African origin. Although the resulting haplotype is widespread and the country of origin could therefore not be pinpointed, the claim of the arrested trader that the skin was coming from Benin (Henschel, pers. comm.), seems unlikely. As this type of illegal products are likely to be from free-ranging lions, a haplotype may be sufficiently indicative for a source population/region, contrary to analysed individuals in a zoo setting.

Due to frequent hybridization in captive populations, and the absence of a studbook for African lions, it should be considered to genotype existing stocks for assignments to breeding lineages. Upcoming NGS techniques, such as SNP genotyping, provide cost-effective opportunities for large-scale genotyping. Conservation strategies based on maintaining pure genetic lineages should not be overly puristic and it should be noted that management decisions are scale-dependent. Regardless of the ultimate goal of captive breeding, including genetic aspects in the management of captive stocks is of utmost importance, since high degrees of relatedness lead to loss of genetic diversity and, consequently, inbreeding.

Perspectives for the Future

Looking into the future, modelling studies may give insight into the possible scenarios for conservation, although socio-economic and cultural aspects are often not taken into account. Data on population diversity in tigers have been used to model the increase in tiger numbers needed to maintain current levels of genetic diversity in the next 150 years (Bay *et al.* 2014). Also for relatively large populations, considered as tiger strongholds, gene flow between populations is required. More strongly, the authors state that in the absence of gene flow current genetic diversity cannot be maintained, knowing the species' demographic parameters (Bay *et al.* 2014). In addition, the authors test for the effect of a delay in iniation of gene flow between isolated populations, and show that a subsequent increase in the rate of population expansion is needed when implementation of connectivity is postponed (Bay *et al.* 2014). Another study using tiger data, modelled the risk of inbreeding and extinction in populations of different sizes and different levels of gene flow (Kenney *et al.* 2014). Even relatively high levels of gene flow do not safeguard small populations, but significantly lowers the extinction risk in medium to large populations. The relevance of connectivity is stressed to avoid increased extinction risk, however, the future for already small populations looks grim according to the applied model (Kenney *et al.* 2014).

Due to a different population structure, such as social organization in prides (Pusey *et al.* 1987; Gilbert *et al.* 1990; Packer *et al.* 1991a; Spong *et al.* 2002), these models may not be applicable for lions. Although many of the small populations in RSA are fenced, also in unfenced reserves across Africa lions are secluded from gene flow in peripherally isolated populations (Newmark 2008). The presence of private alleles in most studied LCUs in East and Southern Africa and the strong isolation-by-distance (IBD) patterns confirm a general lack of gene flow (Dubach *et al.* 2013). This effect may be even stronger in West and (part of) Central Africa, as these populations are located in a severely fragmented habitat (Riggio *et al.* 2012) and intermediate populations have recently been extirpated

(Henschel *et al.* 2014). Björklund (2003) used deterministic population genetics models and an individual-based stochastic model to assess the risk of inbreeding in the lion in relation to habitat loss. Most influential parameters tested were the number of prides and male dispersal rates. To sustain a large outbred population of lions, the area should sustain a continuous population of at least 50 prides, but preferably 100 prides, and allow for unrestricted male dispersal (Björklund 2003). Assuming an average territory size for a pride of lions of 50 km^2, as it holds true for lions in Selous GR, Tanzania, this would entail that an area of 5000 km^2 is needed to satisfy the requirements of the model (Björklund 2003). Surely other parameters may be more applicable for lions living in a different type of habitat, with different prey densities, and different social structure. However, since gene flow is a prerequisite and only few protected areas contain the area needed according to the model, this illustrates the relevance of including genetic aspects in management of wild lion populations. Further loss of habitat or decrease of suitability is one of the major challenges in lion conservation. A recent study modelled the distribution of the African lion in response to a changing global climate (Peterson *et al.* 2014). Previously, it had already been shown that climatic parameters had a great effect on the variance of demographic parameters (Celesia *et al.* 2009) and home ranges (Tuqa *et al.* 2014). In the face of global climatic changes, it was predicted that few new areas will become suitable for the African lion, however, large areas in West and southern Africa are projected to decrease in suitability. Notably for West Africa, where remaining lion habitat is small and fragmented, this may pose a serious problem.

On a more positive note, management interventions in highly inbred populations have shown that these effects can be reversed by reinforcing the inbred population with new genetic lineages. High percentages of abnormal and immotile sperm were found in the Florida panther (*Puma concolor coryi*), a subspecies which is confined to a single population. As a result of habitat destruction the population was reduced to 30-50 individuals (Roelke *et al.* 1993; Barone *et al.* 1994). This led to severe inbreeding effects, such as an unusually high frequency of kinked tails, cowlicks, low sperm quality and heart defects (Roelke *et al.* 1993). As the population decline continued, these defects increased in frequency. This led to the notion that the only way to preserve the Florida panther was by hybridization with individuals from another subspecies, thereby affecting the "genetic integrity" of the Florida panther. Eight female pumas from Texas (*Puma concolor stanleyana*) were introduced in 1995 to genetically enrich the existing population. The population was monitored intensively and over a course of 12 years the number of individuals had increased threefold, heterozygosity doubled, survival and fitness measures improved and inbreeding correlates declined significantly (Pimm *et al.* 2006; Johnson *et al.* 2010). However, the authors also warned that, although the introduction of new genetic material seemed to benefit the population, this approach does not guard the population for persisting pressures like continued habitat loss.

Data from well known lion bottlenecks, such as in the Ngorongoro Crater and in the Gir forest, illustrate that the lion shows high resilience and is capable of bouncing back after strong bottlenecks (Packer *et al.* 1991b; Driscoll *et al.* 2002). However, it is important to note that a population with low genetic diversity as a result of a bottleneck may be more vulnerable to environmental stochasticity and catastrophes, e.g. disease outbreaks. This suggests that continuous monitoring and managing protected areas as a network with metapopulations is the best way forward.

Final thoughts

The aim of this thesis was to obtain insight into the spatial distribution of genetic diversity in the lion, contributing to the understanding of the evolutionary history of the species, and possible application for guiding conservation decisions. A number of genetic approaches was used to elucidate the underlying complexity of intra-specific genetic diversity. Although one is always limited by the number of sampling locations, and not every single lion population could be included, I believe that the patterns described in this thesis show a relatively complete picture of the intraspecific genetic diversity of the lion.

The development of SNPs for the lion opens up possibilities to directly target these positions by generating a SNP panel, and to analyse a greater number of individuals. This would provide data for a phylogeographic context, but also information on levels of genetic diversity within populations. Scoring SNP data is less error-prone than scoring microsatellites, which especially plays a role when dealing with low quality samples, such as scat or historic samples. Reference samples are not needed, as is the case for the sizing of microsatellite data. In addition, genotyping of large numbers of individuals for SNPs is cost-effective and less labour-intensive, compared to amplifying and sequencing large numbers of PCR products. The use of a SNP panel to study more lion populations, guide breeding and to employ different sample types should be further investigated. Finally, it may be worthwhile to also assess the applicability of the panel for other, related species.

It may be inevitable to follow a pragmatic approach in which certain populations are prioritized. Phylogeographic data can guide conservation decisions and rationalize this prioritization. Insights into the intraspecific genetic diversity, presented in this thesis, will hopefully contribute to the establishment of effective conservation practices, and to safeguarding the lion's future in its full diversity.

References

ALERT (2014) African Lion Rehabilitation and Release into the Wild Program.

Antunes A, Troyer JL, Roelke ME *et al.* (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 MA (1999) Phylogeography of three closely related African bovids (tribe Alcelaphini). *Molecular biology and evolution*, **16**, 1724–39.

Barnett R, Shapiro B, Barnes I *et al.* (2009) Phylogeography of lions (Panthera leo ssp.) reveals three distinct taxa and a late Pleistocene reduction in genetic diversity. *Molecular ecology*, **18**, 1668–1677.

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). *Proceedings of the Royal Society B: Biological Sciences*, **273**, 2119–2125.

Barnett R, Yamaguchi N, Shapiro B *et al.* (2014) Revealing the maternal demographic history of Panthera leo using ancient DNA and a spatially explicit genealogical analysis. *BMC Evolutionary Biology*, **14**, 70.

Barone MA, Roelke ME, Howard J *et al.* (1994) Reproductive Characteristics of Male Florida Panthers: Comparative Studies from Florida, Texas, Colorado, Latin America, and North American Zoos. *Journal of Mammalogy*, **75**, 150–162.

Bauer H, Nowell K, Packer C (2012) Panthera leo. In: *IUCN 2013. IUCN Red List of Threatened Species. Version 2013.2.*

Bay RA, Ramakrishnan U, Hadly EA (2014) A call for tiger management using "reserves" of genetic diversity. *The Journal of heredity*, **105**, 295–302.

Bensch S, Andrén H, Hansson B *et al.* (2006) Selection for heterozygosity gives hope to a wild population of inbred wolves. *PloS one*, **1**, e72.

Bertola LD, Vrieling K, De Iongh HH (2011) Conservation Genetics of the African Lion. *Waza news*, **3**, 2–4.

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

Black SA, Fellous A, Yamaguchi N, Roberts DL (2013) Examining the extinction of the Barbary lion and its implications for felid conservation. *PloS one*, **8**, e60174.

Black S, Yamaguchi N, Harland A, Groombridge J (2009) Maintaining the genetic health of putative Barbary lions in captivity: an analysis of Moroccan Royal Lions. *European Journal of Wildlife Research*, **56**, 21–31.

Bouché P, Nzapa Mbeti Mange R, Tankalet F *et al.* (2012) Game over! Wildlife collapse in northern Central African Republic. *Environmental monitoring and assessment*, **184**, 7001–11.

Breitenmoser U, Mallon DP, Ahmad Khan J, Driscoll C (2008) Panthera leo persica. In: *IUCN 2013. IUCN Red List of Threatened Species. Version 2013.2.*

Bruche S, Gusset M, Lippold S *et al.* (2012) A genetically distinct lion (Panthera leo) population from Ethiopia. *European Journal of Wildlife Research*.

Burger J, Hemmer H (2005) Urgent call for further breeding of the relic zoo population of the critically endangered Barbary lion (Panthera leo leo Linnaeus 1758). *European Journal of Wildlife Research*, **52**, 54–58.

Burger J, Rosendahl W, Loreille O *et al.* (2004) Molecular phylogeny of the extinct cave lion Panthera leo spelaea. *Molecular phylogenetics and evolution*, **30**, 841–849.

Celesia GG, Peterson AT, Peterhans JCK, Gnoske TP (2009) Climate and landscape correlates of African lion (Panthera leo) demography. *African Journal of Ecology*, **48**, 58–71.

Cho YS, Hu L, Hou H *et al.* (2013) The tiger genome and comparative analysis with lion and snow leopard genomes. *Nature communications*, **4**.

Craigie ID, Baillie JEM, Balmford A *et al.* (2010) Large mammal population declines in Africa's protected areas. *Biological Conservation*, **143**, 2221–2228.

Driscoll CA, Menotti-Raymond M, Nelson G, Goldstein D, O'Brien SJ (2002) Genomic microsatellites as evolutionary chronometers: a test in wild cats. *Genome research*, **12**, 414–423.

Dubach JM, Briggs MB, White PA, Ament BA, Patterson BD (2013) Genetic perspectives on "Lion Conservation Units" in Eastern and Southern Africa. *Conservation Genetics*, **1942**.

Dubach J, Patterson BD, Briggs MB *et al.* (2005) Molecular genetic variation across the southern and eastern geographic ranges of the African lion, Panthera leo. *Conservation Genetics*, **6**, 15–24.

Edwards SV (2009) Is a new and general theory of molecular systematics emerging? *Evolution; international journal of organic evolution*, **63**, 1–19.

Eizirik E, Kim JH, Menotti-Raymond M *et al.* (2001) Phylogeography, population history and conservation genetics of jaguars (Panthera onca, Mammalia, Felidae). *Molecular ecology*, **10**, 65–79.

Flagstad O, Syvertsen PO, Stenseth NC, Jakobsen KS (2001) Environmental change and rates of evolution: the phylogeographic pattern within the hartebeest complex as related to climatic variation. *Proceedings. Biological sciences / The Royal Society*, **268**, 667–77.

Gaubert P, Machordom A, Morales A *et al.* (2011) Comparative phylogeography of two African carnivorans presumably introduced into Europe: disentangling natural versus human-mediated dispersal across the Strait of Gibraltar. *Journal of Biogeography*, **38**, 341–358.

Gilbert DA, Packer C, Pusey AE, Stephens JC, O'Brien SJ (1990) Analytical DNA fingerprinting in lions: parentage, genetic diversity, and kinship. *The Journal of heredity*, **82**, 378–86.

Haas SK, Hayssen V, Krausman PR (2005) Mammalian species - Panthera leo. *Mammalian Species*, 1–11.

Haus T, Akom E, Agwanda B *et al.* (2013) Mitochondrial diversity and distribution of african green monkeys (chlorocebus gray, 1870). *American journal of primatology*, **75**, 350–60.

Hemmer H (1974) *Untersuchungen zur Stammesgeschichte der Pantherkatzen (Pantherinae) 3 Zur Artgeschichte des Löwen Panthera (Panthera) leo (Linnaeus 1758).* Veröffentlichungen der Zoologischen Staatssammlung München, München, Germany.

Henschel P, Coad L, Burton C *et al.* (2014) The Lion in West Africa Is Critically Endangered (M Hayward, Ed,). *PLoS ONE*, **9**, e83500.

Hewitt GM (2004) The structure of biodiversity - insights from molecular phylogeography. *Frontiers in zoology*, **1**, 4.

Hunter LTB, Rabinowitz A (2009) Felid reintroduction using captive founders: poor science and worst practices. *Cat news*, 1–3.

ISIS (2014) Species Holdings - Panthera leo.

IUCN SSC (2013) *Guidelines for Reintroductions and Other Conservation Translocations*. Gland, Switzerland.

Iongh HH de, Bauer H, Tumenta PN *et al.* (2014) National lion action plans and strategies in Benin,

Cameroon and Senegal. *Cat news*, **60**, 8–11.
Iongh H de, Tumenta P, Croes B *et al.* (2009) Threat of a lion population extinction in Waza National Park, North Cameroon. *CATnews*, **50**, 26–27.
Isaac NJB, Mallet J, Mace GM (2004) Taxonomic inflation: its influence on macroecology and conservation. *Trends in ecology & evolution*, **19**, 464–9.
Johnson WE, Onorato DP, Roelke ME *et al.* (2010) Genetic Restoration of the Florida Panther. *Science*, **329**, 1641–1645.
Kenney J, Allendorf FW, Mcdougal C, Smith JLD, B PRS (2014) How much gene flow is needed to avoid inbreeding depression in wild tiger populations ? *Proceedings of the Royal Society B: Biological Sciences*, **281**.
Knowles LL (2009) Statistical Phylogeography. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 593–612.
Levinsky I, Araújo MB, Nogués-Bravo D *et al.* (2013) Climate envelope models suggest spatio-temporal co-occurrence of refugia of African birds and mammals. *Global Ecology and Biogeography*, **22**, 351–363.
Lorenzen ED, Heller R, Siegismund HR (2012) Comparative phylogeography of African savannah ungulates. *Molecular ecology*, **21**, 3656–70.
Luo S-J, Kim J-H, Johnson WE *et al.* (2004) Phylogeography and genetic ancestry of tigers (Panthera tigris). *PLoS biology*, **2**, e442.
Mace GM (2004) The role of taxonomy in species conservation. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **359**, 711–9.
Mazák JH (2010) Geographical variation and phylogenetics of modern lions based on craniometric data. *Journal of Zoology*, **281**, 194–209.
McMahon BJ, Teeling EC, Höglund J (2014) How and why should we implement genomics into conservation? *Evolutionary Applications*, n/a–n/a.
Miller SM, Bissett C, Burger A *et al.* (2013) Management of reintroduced lions in small , fenced reserves in South Africa : an assessment and guidelines. *South African Journal of Wildlife Research*, **43**, 138–154.
Moodley Y, Bruford MW (2007) Molecular biogeography: towards an integrated framework for conserving pan-African biodiversity. *PloS one*, **2**, e454.
Moritz C (1994) Defining "Evolutionarily Significant Units" for conservation. *Trends in ecology & evolution (Personal edition)*, **9**, 373–375.
Newmark WD (2008) Isolation of African protected areas. *Frontiers in Ecology and the Environment*, **6**, 321–328.
O'Brien SJ (2013) The "exotic aliens" controversy: a view from afar. *Journal of the Bombay Natural History Society*, **110**, 108–113.
O'Brien SJ, Joslin P, Smith III GL *et al.* (1987) Evidence for African origins of founders of the asiatic lion species survival plan. *Zoo Biology*, **6**, 99–116.
Packer C (2013) Is India's Pride Actually African? *Science*, **340**, 1289–1290.
Packer C, Gilbert DA, Pusea AE, O'Brien SJ (1991a) A molecular genetic analysis of kinship and cooperation in African lions. *Nature*, **351**, 562–565.
Packer C, Pusey AE, Rowley H *et al.* (1991b) Case Study of Population Bottleneck: Lions of Ngorongoro. *Biological Conservation*, **5**, 219–230.
Peterson AT, Radocy T, Hall E, Kerbis Peterhans JC, Celesia GG (2014) The potential distribution of the Vulnerable African lion Panthera leo in the face of changing global climate. *Oryx*, 1–10.

Pimm SL, Dollar L, Bass OL (2006) The genetic rescue of the Florida panther. *Animal Conservation*, **9**, 115–122.
Pusey AE, Packer C, Erhoff-Mulder MB (1987) The Evolution os Sex-biased Dispersal in Lions. *Behaviour*, **101**, 275–310.
Riggio J, Jacobson A, Dollar L *et al.* (2012) The size of savannah Africa: a lion's (Panthera leo) view. *Biodiversity and Conservation*, **22**, 17–35.
Roelke ME, Martenson JS, O'Brien SJ (1993) The consequences of demographic genetic depletion in the endangered reduction and Florida panther. *Current Biology*, **3**, 340–350.
Sodhi NS, Butler R, Laurance WF, Gibson L (2011) Conservation successes at micro-, meso- and macroscales. *Trends in Ecology & Evolution*, 1–10.
Spong GF, Creel S (2001) Deriving dispersal distances from genetic data. *Proceedings. Biological sciences / The Royal Society*, **268**, 2571–2574.
Spong G, Stone J, Creel S, Björklund M (2002) Genetic structure of lions (Panthera leo L.) in the Selous Game Reserve: implications for the evolution of sociality. *Journal of Evolutionary Biology*, **15**, 945–953.
Steinmetz A, Eulenberger K, Thielebein J *et al.* (2006) Lens-anomalies and other ophthalmic findings in a group of closely-related angola lions (Panthera leo bleyenberghi). *Zoo Biology*, **25**, 433–439.
Thapar V, Thapar R, Ansari Y (2013) *Exotic Aliens: The lion & the cheetah in India*. Aleph Book Company.
Tuqa JH, Funston P, Musyoki C *et al.* (2014) Impact of severe climate variability on lion home range and movement patterns in the Amboseli ecosystem, Kenya. *Global Ecology and Conservation*, **2**, 1–10.
Uphyrkina O, Johnson WE, Quigley H *et al.* (2001) Phylogenetics, genome diversity and origin of modern leopard, Panthera pardus. *Molecular ecology*, **10**, 2617–33.
Vier Voeters (2014) Zes leeuwen op weg naar Lionsrock.
Wasser SK, Joseph Clark W, Drori O *et al.* (2008) Combating the illegal trade in African elephant ivory with DNA forensics. *Conservation biology : the journal of the Society for Conservation Biology*, **22**, 1065–1071.
Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular ecology*, **17**, 2107–2121.
Zinner D, Groeneveld LF, Keller C, Roos C (2009) Mitochondrial phylogeography of baboons (Papio spp.): indication for introgressive hybridization? *BMC evolutionary biology*, **9**, 83.

List of Abbreviations

aDNA – ancient DNA
ALWG – African Lion Working Group
CAR – Central African Republic
DRC – Democratic Republic of the Congo
EAZA – European Association of Zoos and Aquaria
ESU – Evolutionary Significant Unit
HPD – Highest Posterior Density
IBD – Isolation by Distance
ISIS – International Species Information System
IUCN – International Union for Conservation of Nature
Kya – Thousand years ago
Kyr – Thousand years
LCU – Lion Conservation Unit
MHC – Major histocompatibility complex
ML – Maximum Likelihood
MP – Maximum Parsimony
mtDNA – mitochondrial DNA
MU – Management Unit
My – Million years
Ne – effective population size
NGS – Next Generation Sequencing
PCA – Principal Component Analysis
RFLP – Restriction fragment length polymorphism
RSA – Republic of South Africa
SNP – Single Nucleotide Polymorphism
SSC – Species Survival Commission
SSP – Species Survival Plan
WAZA – World Association of Zoos and Aquariums

Summary

Insights into the spatial distribution of genetic diversity contribute to the understanding of evolutionary histories which have shaped these patterns. But they also provide us with guidelines on how to effectively conserve this diversity. This may serve to rationalize prioritization of populations for conservation, aiming to minimize the loss of genetic diversity and to preserve genetically distinct lineages. In this thesis, the intraspecific genetic diversity of the lion (*Panthera leo*) is assessed. Top predators are known to fulfill a crucial role in the ecosystem, contributing to both species richness and resilience. However, many large carnivores, including the lion, have shown recent declines in population numbers, as a result of anthropogenic pressures. The importance as a keystone species and the vulnerability of large carnivores are arguments to target them as a model to develop conservation plans.

Currently, only two subspecies of lions are officially recognized by the IUCN: the African lion (*Panthera leo leo*), distributed throughout sub-Sahara Africa with the exception of dense rain forest, and the Asiatic lion (*Panthera leo persica*), confined to a single population in India. However, other species with a similar distribution throughout the African continent, show a basal dichotomy, distinguishing populations from West/Central African and populations from East/Southern Africa. This dichotomy is often reflected in their taxonomy. Morphological data of lions, originally leading to the distinction of up to eight “subspecies”, seem to confirm the existence of different African lineages, although results of these studies should be interpreted with caution since they have not always adequately treated age- and sex-related variation. Later studies, including genetic data, have also confirmed that the genetic variation in the lion is greater than the taxonomy implies, and that several genetic lineages can be recognized within the African subspecies. Since taxonomy is often reflected in the conservation strategy, it is important to ensure that the existing genetic variation is thoroughly documented.

Previous studies on lion phylogeny have mainly included lions from East and Southern Africa, and focused on the position of the Asiatic subspecies in relation to these African populations. At the same time, populations in West and (parts of) Central Africa have shown exceptionally strong declines in lion numbers and other wildlife species. This has resulted in the West African lion to be classified as ‘Regionally Endangered’ on the IUCN Red List, with the recent suggestion to uplist it to ‘Critically Endangered’. The urgent conservation need in this region, together with the notion that West and Central Africa may harbour unique genetic lineages, are the main reasons why there is a focus on the West and Central African lion in this thesis. Samples were included from every West and Central African LCU with a recently confirmed resident lion population.

In Chapter 2 the knowledge gap for the West and Central African lion is filled by analysing populations which had not been previously included in phylogeographic studies. By selecting cytochrome *b* and part of the control region to assess the phylogeography of the lion, it was possible to include data from two previous studies. Resulting phylogenetic trees show differentiated lineages in East/

Southern Africa and a strong genetic relationship between the West/Central African lion and the Asiatic subspecies. This may be the result of severe droughts in the western part of the lion range, leading to a strong bottleneck or local extinction of these populations. Subsequent recolonization, possibly from North Africa or the Middle East would explain the strong genetic relationship with the Asiatic subspecies. Main conclusion is that the current taxonomy does not adequately reflect the genetic diversity of the African lion and that options for a taxonomic revision, notably with regard to the position of the West/Central African lion, should be explored.

Phylogenies based on mitochondrial markers may not reflect the complete genomic complexity due to differences in the mode of inheritance and coalescence time compared to autosomal markers. Therefore, a revision of the taxonomy, with potentially far-reaching ramifications for management, should be based on a combination of unlinked genetic markers. In Chapter 3 the congruence between phylogenetic patterns based on mitochondrial and autosomal markers is assessed by including data of 20 microsatellite loci and mtDNA for 15 lion populations. Results show that four genetic clusters can be distinguished: 1) West/Central Africa, 2) East Africa, 3) Southern Africa, and 4) the Asiatic subspecies. Although microsatellites are a suitable marker to infer population structure, they do not give much insight into phylogenetic relationships. The reduced genetic diversity of the Asiatic population, as a result of severe bottleneks, led to a clear distinction of the Asiatic subspecies which does not necessarily reflect a long evolutionary distance. Therefore, it is not possible to use microsatellite data to confirm the close evolutionary relationship between populations in West/Central Africa and the Asiatic subspecies, as was observed in mtDNA data. There is no indication of reduced genetic diversity in West/Central African populations, as was hypothesized based on recent population histories in this region. Possibly the reduction in lion numbers is too recent to be visible as a genetic signature. However, examples of intensively managed lion populations show how fast inbreeding may arise in small and isolated populations. Management interventions may therefore be needed in the future, to safeguard these populations against further declines of genetic diversity and subsequent fitness effects.

Since autosomal data do not contradict phylogenetic patterns based on mtDNA, this dataset was expanded with data from more sampling localities in Chapter 4. Specimens from natural history collections were included for areas where lions are extinct (i.e. North Africa and Middle East) and for areas from which it was not possible to include modern samples of wild lions. An ancient DNA approach was used for genetic analyses. A total of 194 samples from 22 countries were included and complete mitogenomes were analysed for 14 individuals, covering the main phylogeographic groups. Resulting phylogenetic trees reveal a strongly supported basal dichotomy, distinguishing lions from the northern part of the range, including the Asiatic subspecies (North group), and populations from the southern part of the range (South group). Six main haplogroups are identified: three in the North Group (1) West Africa, 2) Central Africa, 3) North Africa/Asia) and three in the South Group (4) North East Africa, 5) East/Southern Africa, and 6) South West Africa). The basal split into two main phylogenetic groups and the distinction of the other haplogroups are reiterated in several other savannah mammals. This points towards environmentally driven evolution in which similar forces have shaped the phylogeographic patterns of co-occurring species. For the lion, the most recent common ancestor of these major lineages is estimated at ~300 thousand years ago, and radiation of the haplogroups probably occurred during the last ~100 thousand years. It is hypothesized that this is the result of cyclical expansion of the rain forest and desert, which may have hampered gene flow between populations. The temporal contraction to local refugia may have led to clearly distinguishable clades due to fast coalescence of mtDNA markers. The proposed scenario is further confirmed by results of bioclimatic envelop models published in a previous study, which predicts refugia in line with the haplogroups we distinguish in the lion. The degree of divergence between the North and the South group, and especially the nested position of the Asiatic subspecies within the West and Central African lion, supports the notion that the current taxonomy is not in line with the evolutionary history of the lion, and that therefore a taxonomic revision is warranted.

In Chapter 5 we develop a new lion-specific genetic marker by sequencing the complete genome of 10 lions, covering the main phylogenetic groups. The genome data were mined for variable positions and a total of ~18,000 lion-specific SNPs was identified. Phylogenetic analyses based on these SNPs result in a tree with a hierarchical structure in which no reciprocally monophyletic clades can be identified. However, the Asiatic subspecies again shows a nested position with the populations in West and Central Africa, which confirms earlier findings based on mtDNA. The SNPs identified in this chapter provide a source for the generation of a SNP panel which can be used for cost-effectively genotyping a larger number of lions. This may serve high-throughput genetic analysis of free-ranging lions as well as the assessment of genetic lineages present in captive stocks, for the design of breeding programmes.

The datasets presented in this thesis consistently illustrate that the Asiatic subspecies has a nested position within the African lion. Both mtDNA and microsatellite datasets show that the West and Central African lion can be distinguished as a separate clade. In East and Southern Africa there is a geographical discrepancy between the clades identified by mtDNA and microsatellites. This may be the result of male-mediated gene flow, as lions are known to exhibit sex-biased dispersal, with males showing stronger dispersal than females. The fast coalescence time of mtDNA further contributes to clearly recognizable monophyletic clades based on these data. The topology of the tree based on the SNPs indicates a continent-wide pattern of gene flow. Differences in mutation rates between bi-allelic SNP markers and multi-allelic microsatellite markers suggest that the SNP data reflect a more ancient pattern of gene flow, before populations were isolated in local refugia.

To translate these patterns into conservation recommendations, it is advisable to recognize units below the subspecies level, such as ESUs or MUs. Requiring reciprocally monophyletic groups for both mtDNA and autosomal markers may be overly restrictive. However, a degree of nuclear allele divergence should be present to avoid misclassification of units which are linked by nuclear gene flow only, and not by organellar gene flow. In addition, divergence based on mtDNA data may represent a relic of historic isolation, only sustained by strong female philopatry. Therefore, it is suggested to treat the monophyletic mtDNA clades as ESUs and not elevate them to a subspecies status. To ensure that the taxonomy reflects the evolutionary history of the lion, it is suggested to revise the current taxonomy and recognise the following units:

Panthera leo leo (Linnaeus, 1758)
Unit 1: Asia (+ North Africa & Middle East, extinct)
Unit 2: Central Africa
Unit 3: West Africa

Panthera leo melanochaita (Hamilton Smith, 1848)
Unit 4: North East Africa
Unit 5: East/Southern Africa
Unit 6: South West Africa

It is advisable to follow a pragmatic approach, keeping in mind the scale-dependency of the problems and possible solutions related to the management of these units, both for free-ranging lions as well as for captive stocks. The data presented in this thesis give a more complete overview of the distribution of genetic diversity in the lion, by including data from more populations, notably from West and Central Africa, and by analysing a variety of genetic markers. The more detailed phylogeography of the lion gives insight into evolutionary drivers that have shaped this genetic makeup, but will also contribute to the development of effective management plans to conserve the full genetic diversity of the lion.

Samenvatting

Inzicht in de ruimtelijke verdeling van genetische diversiteit draagt bij aan het begrijpen van de evolutionaire geschiedenis die deze patronen heeft gevormd. Maar het voorziet ons ook van richtlijnen hoe deze diversiteit effectief te behouden valt. Dit kan dienen als richtlijn voor het prioriteren van populaties voor soortbehoud. Het doel is hierbij om het verlies aan genetische diversiteit te minimaliseren en genetisch verschillende lijnen te behouden. In dit proefschrift wordt de intraspecifieke genetische diversiteit van de leeuw (*Panthera leo*) onderzocht. Het is bekend dat top predatoren een cruciale rol vervullen in het ecosysteem die bijdraagt aan zowel de soortenrijkdom als de veerkracht van het systeem. Niettemin worden er, als gevolg van antropogene factoren, in veel grote carnivoren sterke afnames in aantallen waargenomen. Hieronder valt ook de leeuw. Het belang als sleutelsoort en de kwetsbaarheid van grote carnivoren zijn argumenten om deze soorten als model te gebruiken voor het ontwikkelen van strategieën voor soortbehoud.

Op dit moment worden binnen de leeuw officieel twee ondersoorten erkent door de IUCN: de Afrikaanse leeuw (*Panthera leo leo*), die verspreid voorkomt in sub-Sahara Afrika behalve in dicht regenwoud, en de Aziatische leeuw (*Panthera leo persica*), beperkt tot een enkele populatie in India. Andere soorten met een soortgelijk verspreidingsgebied over het Afrikaanse continent laten in veel gevallen een basale dichotomie zien waarbij populaties uit West/Centraal Afrika onderscheiden worden van populaties uit Oost/Zuidelijk Afrika. Dit patroon wordt vaak weerspiegeld in de taxonomie van deze soorten. Morfologische data van leeuwen, die oorspronkelijk hebben geleid tot het onderscheid van acht “ondersoorten”, lijken de verschillende genetische lijnen in Afrika te bevestigen. Aangezien niet in alle morfologische studies leeftijds- en geslachtsgebonden variatie adequaat is meegewogen, moeten deze resultaten met voorzichtigheid worden geïnterpreteerd. Meer recente studies waarbij ook genetische data gebruikt zijn, hebben eveneens bevestigd dat de genetische variatie in de leeuw groter is dan de taxonomie impliceert. In de Afrikaanse ondersoort kunnen diverse genetische lijnen herkend worden. Aangezien taxonomie vaak wordt weerspiegeld in strategieën voor soortbescherming, is het van belang dat de bestaande genetische variatie degelijk is gedocumenteerd.

In eerdere studies over de fylogenie van de leeuw zijn met name leeuwen uit Oost en Zuidelijk Afrika meegenomen. De focus lag op de positie van de Aziatische ondersoort ten opzichte van deze Afrikaanse populaties. Daarnaast werden in populaties in West en Centraal Afrika bijzonder sterke dalingen in aantallen waargenomen, zowel voor de leeuw als voor andere soorten. Daarom werd de leeuw in West Afrika op de Rode Lijst van het IUCN gecategoriseerd als ‘Regionaal Bedreigd’, en recent werd voorgesteld om deze zelfs als ‘Ernstig Bedreigd (Kritiek)’ aan te zien. De combinatie van de urgente noodzaak voor soortbehoud in deze regio en de notie dat West en Centraal Afrika unieke genetisch lijnen zou kunnen herbergen, zijn de belangrijkste redenen voor de focus op de West en Centraal Afrikaanse leeuw in dit proefschrift. Hiervoor werden monsters verzameld uit elke LCU in West en Centraal Afrika met een recentelijk bevestigde leeuwenpopulatie.

In Hoofdstuk 2 wordt de kenniskloof voor de West en Centraal Afrikaanse leeuw gevuld door het analyseren van populaties die niet eerder deel uit hadden gemaakt van fylogenetische studies. De fylogeografie van de leeuw is onderzocht door middel van het analyseren van cytochroom *b* en een gedeelte van de controle regio. Hierdoor was het mogelijk om data van twee eerdere studies op te nemen in de analyses. De fylogenetische bomen laten gedifferentieerde genetische lijnen zien in Oost/Zuidelijk Afrika, met daarnaast een sterke genetische verwantschap tussen de West/Centraal Afrikaanse leeuw met de Aziatische ondersoort. Dit zou het resultaat kunnen zijn van zeer droge periode in het westelijke gedeelte van het verspreidingsgebied van de leeuw, wat kan hebben geleid tot een sterke genetische bottleneck of zelfs lokale extinctie van deze populaties. Latere herkolonisatie, mogelijkerwijs vanuit Noord Afrika of het Midden Oosten, zou de sterke genetische verwantschap met de Aziatische ondersoort verklaren. Belangrijkste conclusie is dat de huidige taxonomie de genetische diversiteit van de leeuw niet adequaat weerspiegelt en dat mogelijkheden voor een taxonomische revisie verkend moeten worden, met name met oog op de positie van de West/Centraal Afrikaanse leeuw.

Fylogenieën gebaseerd op mitochondriale markers geven niet de volledige genomische complexiteit weer vanwege verschillende wijze van overerving en verschillen in coalescentietijd ten opzichte van autosomale markers. Daarom moet een herziening van de taxonomie, met mogelijk verstrekkende gevolgen voor management, gebaseerd zijn op de combinatie van niet gekoppelde genetische markers. In Hoofdstuk 3 wordt de congruentie onderzocht van fylogenetische patronen gebaseerd op mitochondriale en autosomale markers door het toevoegen van data van 20 microsatellieten en mtDNA voor 15 leeuwen populaties. Resultaten laten zien dat vier genetische clusters onderscheiden kunnen worden: 1) West/Centraal Afrika, 2) Oost Afrika, 3) Zuidelijk Afrika, en 4) de Aziatische ondersoort. Hoewel microsatellieten geschikte markers zijn voor het bepalen van populatie structuur, geven zij weinig inzicht in de fylogenetische verwantschappen. De verminderde genetische diversiteit van de Aziatische populatie, als gevolg van sterke genetische bottlenecks, zorgt voor een duidelijk onderscheid van deze ondersoort, terwijl dit niet noodzakelijkerwijs een lange evolutionaire afstand weergeeft. Daarom is het niet mogelijk om met microsatelliet data de sterke evolutionaire verwantschap tussen de populaties in West/Centraal Afrika en de Aziatische ondersoort zichtbaar te maken, zoals het geval is met de mtDNA data. Er is nog geen indicatie van een verminderde genetische diversiteit in West/Centraal Afrikaanse populaties, hoewel dit voorspeld was op basis van recente negatieve populatie trends in deze regio. Waarschijnlijk is de afname in aantallen leeuwen te recent om op dit moment als genetische signatuur zichtbaar te zijn. Niettemin laten voorbeelden van intensief beheerde leeuwen populatie zien hoe snel inteelt kan optreden in kleine en geïsoleerde populaties. In de toekomst zouden daarom management ingrepen nodig kunnen zijn om deze populaties te beschermen tegen verdere afname van de genetische diversiteit en de daarop volgende fitness effecten.

Aangezien autosomale data de fylogenetische patronen gebaseerd op mtDNA niet tegenspreken, werd de mtDNA dataset verder uitgebreid met monsters van meerdere locaties, gedocumenteerd in Hoofdstuk 4. Specimen uit natuurhistorische collecties werden opgenomen voor gebieden waar leeuwen uitgestorven zijn (i.e. Noord Afrika en het Midden Oosten) en voor gebieden van waar het niet mogelijk was nieuwe monsters van wilde leeuwen te bemachtigen. Een methode voor het verwerken van historische monsters (ancient DNA) is gebruikt voor de genetische analyse. In totaal zijn 194 monsters uit 22 landen meegenomen en zijn complete mitogenomen geanalyseerd voor 14 individuen, verdeeld over de belangrijkste fylogeografische groepen. Fylogenetische bomen van deze data laten een sterk ondersteunde basale dichotomie zien met leeuwen van het noordelijke gedeelte van hun verspreidingsgebied, inclusief de Aziatische ondersoort (Noord groep) op de ene tak, en populaties van het zuidelijke gedeelte van het verspreidingsgebied (Zuid groep) op de andere. Zes voornaamste haplogroepen worden onderscheiden: drie in de Noord groep (1) West Afrika, 2) Centraal Afrika, 3) Noord Afrika/Azië) en drie in de Zuid groep (4) Noord Oost Afrika, 5) Oost/Zuidelijk Afrika, en 6) Zuid West Afrika). Het basale onderscheid in deze twee fylogenetische groepen en het onderscheid in de overige haplogroepen wordt in verscheidene zoogdiersoorten van de savanne teruggevonden. Dit wijst in de richting van omgevingsfactoren, zoals veranderingen in het klimaat, als de drijvende kracht waarbij soortgelijke factoren de fylogeografische patronen van soorten met overlappende verspreidingsgebieden hebben beïnvloed. Voor de leeuw wordt de meest recente gemeenschappelijke voorouder van deze twee hoofdlijnen geschat op ~300 duizend jaar geleden. Radiatie van de haplogroepen vond waarschijnlijk plaats gedurende de laatste ~100 duizend jaar. Het wordt verondersteld dat dit veroorzaakt wordt door een cyclische expansie van het regenwoud en de woestijn, met als gevolg een reductie in gene flow tussen populaties. Het tijdelijk terugdringen van populaties in lokale refugia zou geleid kunnen hebben tot duidelijk onderscheidbare clades vanwege de snelle coalescentie van mtDNA markers. Het voorgestelde scenario wordt verder bevestigd door eerder gepubliceerde resultaten van modellen gebaseerd op bioklimatologische envelop methoden. Hierin worden refugia voorspeld die overeenkomen met de haplogroepen beschreven in dit proefschrift. De mate van divergentie tussen de Noord Groep en de Zuid Groep, en met name de genestelde positie van de Aziatische ondersoort binnen West en Central Afrika, ondersteunt het idee dat de huidige taxonomie niet overeenkomt met de evolutionaire geschiedenis van de leeuw. Een herziening van de taxonomie is daarom gerechtvaardigd.

In Hoofdstuk 5 ontwikkelen we een nieuwe leeuw-specifieke marker door het sequencen van het complete genoom van 10 leeuwen, verspreid over de belangrijkste fylogenetische groepen. De genoom data zijn gemijnd voor het identificeren van variabele posities en in totaal zijn ~18.000 leeuw-specifieke SNPs ontdekt. Fylogenetische analyses gebaseerd op deze SNPs resulteren in een boom met een hiërarchische structuur waarbij geen reciproke monofyletische clades herkend kunnen worden. Ook in dit geval is de Aziatische ondersoort genesteld in populaties uit West en Centraal Afrika, hetgeen de resultaten op basis van mtDNA bevestigt. De SNPs die in dit hoofdstuk geïdentificeerd zijn, zijn een bron voor het ontwikkelen van een SNP panel waarmee een groter aantal leeuwen kosteneffectief onderzocht kan worden. Dit kan bijdragen aan high-throughput genetische analyses van wilde leeuwen, maar ook aan het in kaart brengen van genetische lijnen die aanwezig zijn in populaties in gevangenschap ter ondersteuning van fokprogramma´s.

De datasets in dit proefschrift tonen consistent aan dat de Aziatische ondersoort een genestte positie heeft binnen de Afrikaanse leeuw. Zowel de mtDNA data als de microsatelliet dataset laten zien dat de West en Centraal Afrikaanse leeuw als aparte groep erkent kan worden. In Oost en Zuidelijk Afrika is een geografische discrepantie zichtbaar tussen de clades gebaseerd op mtDNA en microsatelliet data. Dit kan het gevolg zijn van gene flow die gedomineerd wordt door mannelijke individuen, aangezien bij leeuwen dispersie tussen de seksen niet gelijk is en mannelijke leeuwen sterkere dispersie vertonen dan leeuwinnen. De snelle coalescentietijd van mtDNA draagt er verder toe bij dat clades gebaseerd op deze data duidelijk herkenbaar zijn. De topologie van de boom gebaseerd op SNPs geeft een continentwijd patroon van gene flow aan. Verschillen in de mutatiesnelheden tussen SNPs en microsatellieten suggereren dat SNP data een ouder patroon van gene flow weerspiegelen, voordat populaties geïsoleerd raakten in locale refugia.

Om deze patronen te vertalen naar aanbevelingen voor soortbehoud, verdient het de aanbeveling om eenheden onder het ondersoortsniveau te onderscheiden, zoals ESUs of MUs. Mogelijkerwijs is het te restrictief om reciproke monofyletische groepen te vereisen voor zowel mtDNA als autosomale markers. Desalniettemin, zou een zekere graad van divergentie van nucleaire allelen aanwezig moeten zijn, om misclassificatie te voorkomen van groepen die alleen door nucleaire en niet door mitochondriale gene flow verbonden zijn. Bovendien zou divergentie gebaseerd op uitsluitend mtDNA een afspiegeling kunnen zijn van historische isolatie, die alleen in stand wordt gehouden door vrouwelijke filopatrie. Daarom wordt voorgesteld om de monofyletische mtDNA groepen als ESUs te behandelen en om deze niet verheffen tot aparte ondersoorten. Om te verzekeren dat de taxonomie de evolutionaire geschiedenis van de leeuw weerspiegelt, wordt voorgesteld om de huidige taxonomie te herzien en de volgende eenheden te onderscheiden:

Panthera leo leo (Linnaeus, 1758)
Unit 1: Azië (+ Noord Afrika & Midden Oosten, uitgestorven)
Unit 2: Centraal Afrika
Unit 3: West Afrika

Panthera leo melanochaita (Hamilton Smith, 1848)
Unit 4: Noord Oost Afrika
Unit 5: Oost/Zuidelijk Afrika
Unit 6: Zuid West Afrika

Het is aan te bevelen om een pragmatische aanpak te volgen, waarbij de schaal-afhankelijkheid van de problemen en mogelijke oplossingen voor het management van deze eenheden erkend wordt. Dit geldt zowel voor leeuwen in het wild, als voor het beheer van populaties in gevangenschap. De data die in dit proefschrift gepresenteerd zijn, geven een completer overzicht van de verdeling van de genetische diversiteit in de leeuw. Dit is bewerkstelligd door het toevoegen van data van meer populaties, met name uit West en Centraal Afrika, en door het analyseren van een aantal verschillende genetische markers. De gedetailleerde fylogeografie van de leeuw geeft inzicht in evolutionaire drivers die deze genetische opmaak hebben gevormd, maar zal ook bijdragen aan het ontwikkelen van effectieve management plannen voor het behoud van de volledige genetische diversiteit van de leeuw.

Résumé

L'étude de la distribution spatiale de la diversité génétique contribue à une meilleure compréhension des pressions évolutives qui ont façonné les patterns d'organisation de la variation actuelle. De plus, elle nous fournit des lignes directives sur la façon de préserver efficacement cette diversité et peut donc servir à justifier la conservation prioritaire de certaines populations, en visant à minimiser les pertes de la diversité génétique et de préserver lignées génétiquement distinctes. Dans cette thèse, la diversité génétique intraspécifique du lion (*Panthera leo*) est évaluée. Ces grands prédateurs ont un rôle crucial dans l'écosystème, en contribuant autant à sa richesse spécifique qu'à sa résilience. Cependant, de nombreux carnivores, y compris le lion, voient le nombre de leurs populations décroître à la suite de pressions anthropiques. En raison de leur importance et leur vulnérabilité, les carnivores représentent un modèle fondamental pour l'élaboration de plans de conservation.

Actuellement, deux sous-espèces de lions sont officiellement reconnues par l'UICN: le lion d'Afrique (*Panthera leo leo*), répartis sur l'ensemble de l'Afrique subsaharienne à l'exception de la forêt tropicale dense, et le lion d'Asie (*Panthera leo persica*), confiné à une seule population en Inde. Cependant, d'autres espèces ayant une distribution similaire à travers le continent africain montrent une dichotomie de base, distinguant les populations d'Afrique de l'Ouest/Centrale et les populations d'Afrique Orientale/Australe. Cette dichotomie se reflète souvent dans leur taxonomie. Les études morphologiques des lions qui ont initialement conduit à la distinction de près de huit "sous-espèces", semblent confirmer les deux lignées de lions. Toutefois les résultats de ces études doivent être interprétés avec précaution car les variations d'âge et de sexe ne sont pas toujours traitées de façon adéquate. Les premières études génétiques ont confirmé que la variation génétique du lion est supérieure au nombre de taxon établi et que la sous-espèce africaine est constituée de plusieurs lignées génétiques. Puisque les stratégies de conservation et la taxonomie sont interdépendants, il est important de veiller à ce que la variation génétique existante soit bien documentée.

Jusqu'à présent, les études phylogéniques du lion ont principalement inclus les lions d'Afrique Orientale et Australe et se sont concentrées sur le positionnement des sous-espèces Asiatiques par rapport aux populations africaines. Dans un même temps, les populations d'Afrique de l'Ouest et (en parties) de l'Afrique Centrale ont montré des déclins exceptionnellement forts du nombre de lions et d'autres espèces sauvages. En conséquence, le lion d'Afrique de l'Ouest a été classé comme «régionalement menacé», et il a même récemment été suggéré de le classer «en danger critique». Cette thèse se concentre principalement sur l'étude génétique du lion d'Afrique de l'Ouest et d'Afrique Centrale, en raison du besoin urgent de conservation dans ces régions et du fait de la présence potentielle de lignées génétiques uniques.

Dans le Chapitre 2, la phylogéographie des populations du lion d'Afrique de l'Ouest et d'Afrique Centrale a été analysée. Cette analyse a été faite à l'aide du marqueur génétique cytochrome *b* et d'une partie de la région de contrôle afin d'y inclure les données de deux études antérieures. Dans notre étude, les arbres phylogénétiques montrent lignées bien différenciées en Afrique Orientale/

Australe et une forte relation génétique entre des lions d'Afrique de l'Ouest/Centrale et la sous-espèce Asiatique. Ce résultat peut être expliqué par de graves sécheresses dans la partie ouest de l'aire de distribution du lion, conduisant à un fort goulot d'étranglement ou à l'extinction locale de ces populations. Une recolonisation ultérieure venant de l'Afrique du Nord ou du Moyen-Orient pourrait alors expliquer la forte relation génétique avec la sous-espèce Asiatique. La conclusion principale est que la taxonomie actuelle ne reflète pas adéquatement la diversité génétique du lion africain. Les différentes options pour une révision taxonomique concordante, notamment en ce qui concerne le positionnement du lion d'Afrique de l'Ouest/Centrale, devraient être explorées.

Les phylogénies basées sur les marqueurs mitochondriaux pourraient masquer la complexité génomique en raison des différents modes de transmission génétique et du temps de coalescence par rapport aux marqueurs autosomiques. Puisque les révisions taxonomiques ont des conséquences potentiellement importantes sur la gestion et la conservation de la biodiversité, une telle révision du lion devrait être basée sur une combinaison de marqueurs génétiques non liés. Dans le Chapitre 3, la congruence entre les patterns phylogénétiques basé sur des marqueurs mitochondriaux et autosomiques est évaluée au moyens de 20 loci microsatellites et d'ADN mitochondrial dans 15 populations de lions. Les résultats permettent de différencier quatre groupes génétiques distincts: les populations de 1) l'Afrique de l´Ouest/Centrale, 2) l´Afrique de l'Est, 3) l'Afrique Australe, et 4) Asiatiques. Bien que les microsatellites soient des marqueurs appropriés pour inférer la structure des populations, ils ne permettent pas d'explorer les relations phylogénétiques. La diversité génétique réduite de la population Asiatique permet de distinguer clairement de la sous-espèce Asiatique. Toutefois, ceci ne reflète pas nécessairement une longue distance évolutive. Par conséquent, il n´est pas possible de tester les relations phylogénétiques entre les populations d'Afrique de l´Ouest/Centrale et la sous-espèce Asiatique. De plus, ces analyses n'indiquent aucune réduction de la diversité génétique dans les populations de l'Afrique de l´Ouest/Centrale, comme le suggérait l'hypothèse basée sur l'histoire des populations dans cette région. Ce résultat peut être expliqué par une diminution récente de l'effectif de lions pouvant masquer la signature génétique. Cependant, des exemples de gestion intensive des populations de lion ont montré que la consanguinité peut survenir rapidement au sein de petites populations isolées. Par conséquent, une meilleure gestion des populations peut être nécessaires pour les prémunir contre une perte de la diversité génétique et de ces effets sur la fitness de ces populations.

Comme les données autosomiques ne contredisent pas les patterns phylogénétiques basés sur l'ADN mitochondrial, l'échantillonnage de localité a été étendu et inclus à l'ensemble de données précédent puis analysé dans le Chapitre 4. Des spécimens provenant de collections d'histoire naturelle ont été inclus pour les zones où les lions ont aujourd'hui disparu (ce est à dire, l'Afrique du Nord et Moyen-Orient) et pour les zones dans lesquelles il n´était pas possible d'échantillonner de lions sauvages. Une méthode appelée «ancient DNA» a été utilisée pour l'analyse génétique de ces données. Un total de 194 échantillons provenant de 22 pays ont été inclus et les mitogenomes complets de 14 individus ont été analysés, couvrant ainsi les principaux groupes phylogéographiques. Les reconstructions phylogénétiques révèlent une dichotomie de base fortement soutenue, en distinguant les lions de la partie nord de la distribution, y compris la sous-espèces Asiatique (Groupe du Nord), et les populations de la partie sud de la distribution (Groupe Sud). Six principaux haplogroupes peuvent être identifiés: 1) l´Afrique d'Ouest, 2) l'Afrique Centrale, 3) l'Afrique du Nord/Asie (Groupe Nord), 4) l'Afrique de Nord-Est, 5) l'Afrique Oriental/Australe, et 6) l'Afrique de Sud-Ouest (Groupe Sud). La division de base en deux groupes phylogénétiques principaux et la distinction d'autres haplogroupes se retrouvent dans plusieurs autres mammifères de la savane. Ceci indique que des facteurs environnementaux auraient façonné la répartition phylogéographique d'espèces coexistantes. Pour le lion, le plus récent ancêtre commun de ces lignées est estimé à ~300 mille ans et la diversification des haplogroupes s'est probablement produite durant les derniers ~100 mille ans. L'expansion cyclique de la forêt tropicale et du désert pouvant entraver le flux de gènes entre les populations, expliquerait ce résultat. La contraction temporelle dans des refuges localisés pourrait alors avoir conduit à différents clades clairement distincts dû à la coalescence rapide des marqueurs d'ADN mitochondrial. Le scénario proposé est confirmé par les résultats de modèles d'enveloppes bioclimatiques publiés dans une étude précédente et qui prédit des refuges correspondant aux différents haplogroupes du lion. Le degré de divergence entre le Groupe Nord et le Groupe Sud et en particulier la position de la sous-espèce Asiatique dans le clade d'Afrique d'Ouest/Centrale soutient l'idée que la taxonomie actuelle ne concorde pas avec l'histoire évolutive du lion, et justifierait par conséquent une révision taxonomique.

Dans le chapitre 5, nous développons un nouveau marqueur génétique spécifique pour le lion obtenu par séquençage du génome entier de 10 lions, couvrant les principaux groupes phylogénétiques. Les données génomiques ont été extraites pour des positions variables et un total de ~18 000 SNPs ont été identifiés. Des analyses phylogénétiques basées sur ces SNPs résultent en un arbre avec une structure hiérarchique dans lequel il n'a pas de clades réciproquement monophylétiques. Cependant, la sous-espèce Asiatique se trouve à nouveau imbriquée dans le clade d'Afrique d'Ouest/Centrale. Les SNPs identifiés dans ce chapitre permettent de générer un panel de SNP pouvant être utilisé pour génotyper rentablement un plus grand effectif de lions, comme lors d'analyse génétique à très haut débit de lions sauvages ainsi que l'évaluation de lignées génétiques présentes dans les populations captives.

Les ensembles de données présentés dans cette thèse illustrent d'une manière cohérente que la sous-espèce Asiatique a une position imbriquée dans le lion d'Afrique. Les données d'ADN mitochondrial ainsi que les données microsatellites montrent que le lion de l'Afrique d´Ouest/Centrale est reconnu comme un clade distinct. Cependant, en Afrique Orientale et Australe, les clades géographiques identifiés par l'ADN mitochondrial et les microsatellites divergent. Cela pourrait être le résultat de flux de gènes d'origine masculine, étant donné la dispersion biaisée des sexes chez les lions. De plus, le temps de coalescence rapide de l'ADN mitochondrial contribue à des clades monophylétiques clairement reconnaissables selon nos données. La topologie de l'arbre basé sur les SNPs indique un modèle de flux de gènes à l'échelle continentale. Les différences dans les temps de mutation entre les marqueurs SNP bi-allélique et marqueurs microsatellites multi-alléliques suggèrent que les données des SNPs reflètent un modèle de flux de gènes plus ancien qui se serait produit avant que les populations aient été isolées dans des refuges localisés.

Afin de traduire les données en recommandations pour la conservation, il est conseillé de reconnaître les unités en dessous du niveau de l'espèce, comme par exemples les ESUs ou le MUs. Exiger d'utiliser des groupes réciproquement monophylétiques pour des marqueurs d'ADN mitochondrial et autosomiques peut être trop restrictif. Cependant, un niveau de divergence pour les allèles nucléaires doit être présent afin éviter les erreurs de classification des unités liés par les flux de gènes nucléaires, et non par les flux de gènes des organites. De surcroît, la divergence sur la base de données d'ADN mitochondrial pourrait représenter une relique d'isolations passées et maintenues seulement par la forte philopatrie des femelles. Par conséquent, il est suggéré de traiter les clades monophylétiques obtenus par l'ADN mitochondrial comme les ESUs et ne pas les élever au statut de sous-espèce.

Enfin, pour s'assurer que la taxonomie reflète l'histoire évolutive du lion, il est suggéré de réviser la taxonomie actuelle et reconnaître les unités suivantes:

Panthera leo leo (Linnaeus, 1758)
Unit 1: l'Asie (+ Afrique du Nord et Moyen-Orient, disparu)
Unit 2: l'Afrique Centrale
Unit 3: l'Afrique de l'Ouest

Panthera leo melanochaita (Hamilton Smith, 1848)
Unit 4: l'Afrique de Nord-Est
Unit 5: l'Afrique Oriental/Australe
Unit 6: l'Afrique de Sud-Ouest

Il est conseillé de suivre une approche pragmatique, en tenant compte des problèmes liés à l'échelle et des solutions possibles associées à la gestion de ces unités. Ceci compte autant pour les populations sauvages que captive. Les données présentées dans cette thèse donnent un aperçu plus complet de la distribution de la diversité génétique dans le lion par l'addition de données de plusieurs populations, notamment en provenance d'Afrique de l'Ouest et Centrale, et par l'analyse d'une variété de marqueurs génétiques. Une phylogéographie plus détaillée du lion donne, d'une part, un aperçu sur l'histoire évolutive qui a façonné la structure génétique, et d'autre part, contribue à l'élaboration de plans de gestion efficace pour conserver la diversité génétique complète du lion.

Acknowledgements

Many people have played vital roles during my time as a PhD researcher, but even more have contributed to the path that has led to this point. I would like to start with thanking my two copromotores and supervisors, Hans de Iongh and Klaas Vrieling, for providing me the opportunity to work on this topic, first as a student, later as a PhD researcher. Thanks to your input and faith in the proposal it was possible to obtain funding for a PhD project. Due to your different backgrounds and the somewhat twofold character of the project, I was able to get “the best of both worlds”. Hans, thanks to your extensive network, we were able to set up numerous collaborations and obtain samples from a great part of the lion’s range. I am grateful for your guidance and the opportunity to become a part of this network. Klaas, I value our discussions and your critical questions, which have improved this work greatly. During the course of this project I have gained many new insights, thanks to your expertise. I’m thankful for the input from my promotor, Geert de Snoo. Despite busy schedules there was always time for challenging questions and discussions on how to proceed. I would like to thank my supervisors for their guidance and inspiration during the past years.

Many teachers who I have met when I started studying in Leiden (or even before) have played a role in preparing me for a scientific career. I am happy that many of these connections are still intact. Ruben, you have played an essential role in this when asking me to write an essay about “something that matters”. This proved to be the first step towards working on lion genetics. Although I hadn’t mentioned the lion anywhere in the paper (it just dealt with the use of genetic data to make recommendations for conservation strategies), I’m glad I ended up in Hans’ office and got the chance to work on a true flagship species. In this way, some of the people mentioned above became my colleagues and I’ve continued learning during discussions at the coffee table, both at the CML and the IBL. Collaborations with other institutes, such as NCB Naturalis and the LUMC, have created even more opportunities to benefit from different expertise.

I am very grateful for the opportunities that were given to me to visit conferences, workshops and courses. Visiting several African countries has contributed a great deal to my understanding of the challenges and opportunities in the field of conservation. Also trips outside the lion´s geographical range were sources of inspiration. An African proverb says “Traveling is learning” and I like to keep this in mind on my trips. A great charm of my working environment is the international character which provides the opportunity to gain a more in depth knowledge from different cultural perspectives. I learned much from my African fellow lion researchers, Etotepe, Pricelia and Tuqa, who have surely added new points of view for me to the topic of lion conservation. I have also learned a lot from the members of the African Lion Working Group, and I enjoyed meeting and traveling with some of them. Colleagues who we have directly collaborated with, some of them included as a coauthor on the chapters in this thesis, have been very helpful with their contributions and insights. Some preliminary work and part of the data presented in this thesis are the result of internships of three BSc. students and three MSc. students: Hermen, Lana, Mike, Hester, Laura and Ljamis. It was very instructive for me to supervise you during these projects and I hope that you look back on it as an

interesting and useful experience, even though it was not always easy to obtain the results we were expecting.

Thanks to my fellow PhD researchers who became my friends, working days were never dull. I enjoyed the concerts and events we visited together, the writing weekends, the after work beers, the movie nights and the rather random email messages that would occasionally clutter the mailbox. Although we regularly discussed scientific topics, I especially valued the discussions with Krijn, often leading to more questions than answers. I'm happy that we can continue our discussions in the view of the new project which will hopefully have a fruitful future.

In this place I would also like to express my gratitude to all people who have supported me in the past years. Tim, thanks for your friendship; I doubt that our emails will make much sense to others, who don´t share our fascination for "weird" creatures. Jasper, if we don't see each other in Delft or Madrid, we should simply pick a new destination to meet up. I'm happy that we succeeded in traveling a small part of Africa together (as had been the plan for many years), but there is more exploring to do and endemic species to score. Amar, it is time to catch up again and, after all, the other end of the globe is not that far away. In addition, I would like to thank Jasper and Coen for being my paranymphs.

I would like to finish by thanking my family and especially my parents. Both coming from a scientific background, they understood the path I was taking and were of great support. Dinner table conversation often turned into scientific discussions, and more than once additional research questions were proposed. I'm very grateful for your support.


Cameroon - Bénoué NP


Kenya - Amboseli NP

Curriculum Vitae

Laura Bertola was born on the 30th of August 1984 in Delft. In 2002 she obtained her VWO diploma from the St. Stanislascollege in Delft. In the same year she started studying biology at Leiden University. After internships at the Ministery of Transport, Public Works and Water Management and different institutes in Leiden, she decided that she would like to combine theoretic research with direct applicability. Her Msc. project, supervised by Dr. Klaas Vrieling and Prof. dr. Hans de Iongh, was dealing with the genetics of the West and Central African lion, and served as a pilot for the study presented in this thesis. After earning a MSc. degree in 2008, it was decided to write a proposal for The Netherlands Organisation for Scientific Research (NWO) to obtain funding for a follow-up project on the genetic diversity of the lion. Funding was received in 2010 and the PhD project was started shortly after.

The PhD project was a collaboration between the Institute of Environmental Sciences (CML) and the Institute of Biology Leiden (IBL) of Leiden University, represented by co-promotores Prof. Dr. Hans de Iongh and Dr. Klaas Vrieling, and promotor Prof. Dr. Geert de Snoo. During the course of her PhD, Laura was enabled to visit multiple conferences and workshops abroad and became an active member of the African Lion Working Group (ALWG). She also supervised several students, assisted during courses and gave lectures for Bsc. and MSc. students. In addition, she followed several courses and obtained a diploma of the Netherlands Research School for the Socio-Economic and Natural Sciences of the Environment (SENSE).

Since 2014 Laura has a position as a researcher at the CML, working on environmental DNA. In this project the applicability of DNA isolated directly from aquatic samples is assessed for water quality monitoring.


Namibia - Etosha NP


Senegal - Niokolo Koba NP

Publications

In prep.

Bertola L.D, Vermaat M., White P.A., De Iongh H.H., Laros J. and Vrieling K. SNP discovery and phylogenetic analyses across ten populations of lions reveals a more complex evolutionary history.

Bertola L.D., Vrieling K. & De Iongh H.H. Phylogenetic structure of lions across their geographic range based on a SNP panel.

Submitted

Bertola L.D., Jongbloed H., Van der Gaag K.J., De Knijff P., Yamaguchi N., Hooghiemstra H., Bauer H., Henschel P., White P.A., Driscoll C.A., Tende T., Ottosson U., Saidu Y., Vrieling K. & De Iongh H.H. Phylogeographic patterns in Africa and high resolution delineation of genetic clades in the African lion.

Bertola L.D., Tensen L., Van Hooft P., White P.A., Driscoll C.A,, Henschel P., Caragiulo A., Dias-Freedman I., Sogbohossou E.A., Tumenta P.N., Jirmo T.H., De Snoo G.R, De Iongh H.H. & Vrieling K. Autosomal and mtDNA markers reveal concordant phylogenetic patterns of lion populations over the entire geographic range.

Published

Cho Y.S., Hu L., Hou H., Lee H., Xu J., Kwon S., Oh S., Kim H., Jho S., Kim S., Shin Y., Kim B.C., Kim H., Kim C., Turner J.A., Luo S., Johnson W.E., Koepfli K., Schmidt-ku A., Marker L., Harper C., Miller S.M., Jacobs W., Bertola L.D., Kim T.H., Lee S., Zhou Q., Jung H., Xu X., Gadhvi P., Xu P., Xiong Y., Luo Y., Pan S., Gou C., Chu X., Zhang J., Liu S., He J., Chen Y., Yang L., Yang Y., He J., Liu S., Wang J., Kim C.H., Kwak H., & Kim J. (2013) The tiger genome and comparative analysis with lion and snow leopard genomes. *Nature communications*, **4**, 2433.

Bertola L.D., Vrieling K. & De Iongh H.H. (2012) Conservation Genetics of the Lion: New Approaches to Species Conservation. In: Genetic Diversity: New Research. Cervantes Amaya J.A., Franco Jimenez M.M. (Eds.). Hauppage NY: Nova Science Publishers, Inc.. 61-82.

Bertola L.D., Vrieling K. & De Iongh H.H. (2011) Conservation Genetics of the African Lion. *Waza news*, **3**, 2–4.

Bertola L.D., van Hooft W.F., Vrieling K., Uit de Weerd D.R., York D.S., Bauer H., Prins H.H.T., Funston P.J., Udo de Haes H. a., Leirs H., van Haeringen W. a., Sogbohossou E., Tumenta P.N., & De Iongh H.H. (2011) Genetic diversity, evolutionary history and implications for conservation of the lion (Panthera leo) in West and Central Africa. *Journal of Biogeography*, **38**, 1356–1367.

Bertola L.D., Schoonheere E., Wilms I., Kabli S., Alia A., De Groot H.J.M. & Bagowski C.P. (2011) Effects of hypoxia on buoyancy control and the development of lordosis in physostomous and physoclistous fish species. *Research & Reviews in Biosciences.***5**(1).

Bertola L.D., Griepsma S., Vonk F.J. & Bagowski C.P. (2008) Developmental expression of the alpha-skeletal actin gene. *BMC Evolutionary Biology*, **8**(1): 166-166.

Abstracts for workshops & conferences

Bertola L.D., Vrieling K. & De Iongh H.H. (2015) Insights into the distribution of genetic diversity in the lion and implications for conservation. 1st Annual Meeting in Conservation Genetics – Science and Practice (Birmensdorf – Switserland).

Bertola L.D., Vrieling K. & De Iongh H.H. (2015) Phylogeographic patterns in Africa and high resolution delineation of genetic clades in the African lion. International Biogeography Society – 7th Biennial Conference (Bayreuth – Germany).

Bertola L.D., Vrieling K. & De Iongh H.H. (2013) Importance de la génétique pour la conservation du lion. L'atelier d'élaboration du Plan d'Action National sur la conservation du lion et du Iycaon au Sénégal (Dalaba – Senegal).

Bertola L.D., Vrieling K. & De Iongh H.H. (2012) Genetic methods for Lion Conservation. Felid Taxon Advisory Group (EAZA) (Amersfoort – The Netherlands).

Bertola L.D., Vrieling K. & De Iongh H.H. (2012) Genetics for lion conservation. Lion Range State meeting (Johannesburg – Republic of South Africa).

Bertola L.D., Vrieling K. & De Iongh H.H. (2012) Lion genetics: applications for conservation. African Lion Working Group meeting (Okaukuejo – Namibia).

Bertola L.D., Vrieling K. & De Iongh H.H. (2011) The Importance of Biogeographic Data for Lion Conservation. IUCN Conservation Breeding Specialist Group Annual meeting (Prague – Czech Republic).

Bertola L.D., Vrieling K. & De Iongh H.H. (2011) Use of genetic markers to unravel the phylogenetic position of the lion in West and Central Africa. 8th International Conference on Behaviour, physiology and genetics of wildlife (Berlin – Germany).

Bertola L.D., Vrieling K. & De Iongh H.H. (2011) The use of genetics for lion conservation. Nederland Annual Ecology Meeting (NAEM) (Lunteren – The Netherlands).

Bertola L.D., Vrieling K. & De Iongh H.H. (2010) Genetic diversity and phylogenetics of the lion in West and Central Africa. Carnivore Workshop (Maroua – Cameroon).

Bertola L.D., Vrieling K. & De Iongh H.H. (2010) Genetic diversity, evolutionary history and implications for conservation of the lion (Panthera leo) in West and Central Africa. Carnivore Researchers Meeting (Nairobi – Kenya).

Courses

Systems Biology: Statistical analysis of ~omics data (2010) (Wageningen University + LUMC + NBIC)

Next Generation Sequencing data analysis (2011) (LUMC + NBIC)

LU PhD course Time Management (2012) (Leiden University)

LU PhD course Communication in Science (2012) (Leiden University)

Using R in data analysis (2012) (Leiden University + LUMC)

Advanced NGS: de novo assembly (2013) (Wageningen University)

LU PhD course Effective Communication (2013) (Leiden University)

LU PhD course Scientific Integrity (2013) (Leiden University)

Netherlands Research School for the
Socio-Economic and Natural Sciences of the Environment

D I P L O M A

For specialised PhD training

The Netherlands Research School for the
Socio-Economic and Natural Sciences of the Environment
(SENSE) declares that

Laura Diana Bertola

born on 30 August 1984 in Delft, The Netherlands

has successfully fulfilled all requirements of the
Educational Programme of SENSE.

Leiden, 18 March 2015

the Chairman of the SENSE board

Prof. dr. Huub Rijnaarts

the SENSE Director of Education

Dr. Ad van Dommelen

The SENSE Research School has been accredited by the Royal Netherlands Academy of Arts and Sciences (KNAW)

KONINKLIJKE NEDERLANDSE
AKADEMIE VAN WETENSCHAPPEN

The SENSE Research School declares that Ms Laura Bertola has successfully fulfilled all requirements of the Educational PhD Programme of SENSE with a work load of 47.8 EC, including the following activities:

SENSE PhD Courses

- Environmental Research in Context (2012)
- Research in Context Activity: Organisation of event 'Tiger conservation in Nepal', including discussion meetings and lectures, Kathmandu University, Nepal (2012-2014)

Other PhD and Advanced MSc Courses

- Systems Biology: Statistical analysis of ~omics data, Wageningen University (2010)
- Next Generation Sequencing data analysis, Netherlands Bioinformatics Centre (NBIC) and Leiden University Medical Center (LUMC) (2011)
- Time Management, Leiden University (2011)
- Communication in Science, Leiden University (2012)
- Using R in data analysis, Leiden University Medical Center (LUMC) and Leiden University (2012)
- Advanced Next Generation Sequencing: de novo assembly, Netherlands Bioinformatics Centre (NBIC), Wageningen University and Leiden University (2013)
- Effective Communication, Leiden University (2013)
- Scientific Integrity, Leiden University (2013)

Management and Didactic Skills Training

- Supervising three MSc and three BSc students (2010-2014)
- Teaching in BSc courses 'Moleculair biological techniques', 'Environmental biology theory' (2010-2012)
- Teaching in MSc courses 'Trends in conservation biology', 'Environmental processes and biodiversity' and 'Trends in behaviour and ecology' (2010-2014)
- Member of Institute of Environmental Sciences (CML) council (2011-2014)

Selection of Oral Presentations

- *The use of genetics for lion conservation*. Nederland Annual Ecology Meeting (NEAM), 8-9 February 2011, Lunteren, The Netherlands
- *Use of genetic markers to unravel the phylogenetic position of the lion in West and Central Africa*. 8th International Conference on Behaviour, physiology and genetics of wildlife, 14-19 September 2011, Berlin, Germany
- *The Importance of Biogeographic Data for Lion Conservation*. IUCN Conservation Breeding Specialist Group Annual meeting, 29 September-1 October 2011, Prague, Czech Republic
- *Genetics for lion conservation*. Lion Range State meeting, 29-30 March 2012, Johannesburg, South Africa

SENSE Coordinator PhD Education

Dr. ing. Monique Gulickx

Propositions

Thesis:

GENETIC DIVERSITY IN THE LION
(*Panthera leo* (Linnaeus 1758)):
Unravelling the Past and Prospects for the Future

by Laura Bertola

1. For long-term conservation of a species, it is desirable to maintain the maximum of genetic lineages, which makes phylogeography an important component of conservation biology. (this thesis)

2. Uneven sampling in the geographic range of species may lead to false interpretations regarding the distinctiveness of a population. (Chapter 3)

3. Frequency-based markers may not give insight into evolutionary relationships if the population has gone through one or more severe bottlenecks. (Chapter 3)

4. Despite indications for hybridization between (sub-Saharan) African and Asiatic lions in India from historic publications, there is no indication from genetic data that this has actually occurred. (this thesis)

5. The pre-Darwinian nomenclature that we use to classify the lumpy continuum of living organisms in space and time should not trick us into believing that the species we name are necessarily sensible evolutionary, ecological or conservation units. (Isaac *et al.,* 2004)

6. Beyond the problem of extinction of species per se is the even more complex problem of the extinction of interactions.

7. Phylogenetic diversity is a more suitable measure for biodiversity than species richness, because evolutionary distance is taken into account.

8. Although hybridization of subspecies will affect the genetic integrity of these individuals, human-mediated translocations and resulting hybridization may be a useful tool to ensure survival of a population in some extreme cases.

9. When working towards a solutions for a problem, preconceived ideas may be a larger complication than ignorance. (derived from lectures by Hans Rosling)

10. The capitalist view on maximizing returns is very useful, as long as financial capital is not the only capital that is being considered.

11. The strong emphasis on predictable results and valorization limits opportunities for serendipity.

12. Possibly due to physical nature of intervals between pitches, bullfinches tend to “improve” songs that are taught to them by human teachers who whistle off-key. (derived from lectures by Tim Birkhead)