Conservation Genomic Analyses of African and Asiatic Cheetahs (*Acinonyx jubatus) Across Their Current and Historic Species Range.*

Stefan Prost[1,2], Ana Paula Machado[3], Julia Zumbroich[4], Lisa Preier[4], Sarita Mahtani-Williams[4], Katerina Guschanski[5], Jaelle C. Brealey[5], Carlos Fernandes[6], Paul Vercammen[7], Lena Godsall-Bottriell[8], Paul Bottriell[8], Desire Lee Dalton[2], Antoinette Kotze[2], Pamela Anna Burger[4]

[1]LOEWE-Center for Translational Biodiversity Genomics, Senckenberg Museum, 60325 Frankfurt, Germany
[2]South African National Biodiversity Institute, National Zoological Gardens of South Africa, Pretoria 0001, South Africa
[3]Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland
[4]Research Institute of Wildlife Ecology, Vetmeduni Vienna, 1160 Vienna, Austria
[5]Animal Ecology, Department of Ecology and Genetics, Evolutionary Biology Centre, Science for Life Laboratory, Uppsala Universitet, Uppsala, Sweden
[6]CE3C - Centre for Ecology, Evolution and Environmental Changes, Department of Animal Biology, Faculty of Sciences, University of Lisbon, 1749-016 Lisbon, Portugal
[7]Breeding Centre for Endangered Arabian Wildlife, Sharjah, United Arab Emirates
[8]Rex Foundation, UK

Summary (244 words)

Cheetahs (*Acinonyx jubatus)* are majestic carnivores and the fastest land animals; yet, they are quickly heading towards an uncertain future. Threatened by habitat loss, human-interactions and illegal trafficking, there are only approximately 7,100 individuals remaining in the wild. Cheetahs used to roam large parts of Africa, and Western and Southern Asia. Today they are confined to about 9% of their original distribution. To investigate their genetic diversity and conservation status, we generated genome-wide data from historical and modern samples of all four currently recognized subspecies, along with mitochondrial DNA (mtDNA) and major histocompatibility complex (MHC) data. We found clear genetic differentiation between the subspecies, thus refuting earlier assumptions that cheetahs show only little population differentiation. Our analyses further showed that cheetahs from East Africa are more closely related to *A. j. soemmeringii* than they are to *A . j. juabtus*. This raises the need for a reclassification of the cheetah subspecies, as East African cheetahs are currently included into *A . j. juabtus*. We detected stronger inbreeding in the Critically Endangered *A. j. venaticus* (Iran) and *A. j. hecki* (Northwest Africa), and show that overall genome-wide heterozygosity in cheetah is lower than that reported for other threatened and endangered felids, such as tigers and lions. Furthermore, we show that MHC class II diversity in cheetahs is generally higher than previously reported, but still lower than in other felids. Our results provide new and important information for efficient genetic monitoring, subspecies assignments and evidence-based conservation policy decisions.

Introduction (max 500 words)

Cheetahs (*Acinonyx jubatus*, Schreber 1775) are currently divided into four subspecies by the Cat Classification Task Force of the IUCN Cat Specialist Group, namely *A. jubatus hecki*

(Northwest Africa), *A. j. soemmeringii* (Northeast Africa), *A. j. jubatus* (Southern and East Africa) and *A. j. venaticus* (presently only found in Iran) (Charruau et al. 2011; Kitchener et al. 2017). Krausman and Morales (2005) list *A. j. raineyi* (East Africa) as a fifth subspecies, but its status is under debate (Charruau et al. 2011, Kitchener et al., 2017). It is often recognized as a synonym of *A. j. jubatus*, because of its close genetic relationship to this subspecies inferred from mitochondrial DNA (Charruau et al. 2011, Kitchener et al., 2017). Furthermore, there is an ongoing discussion about the possible presence of *A. j. venaticus* in Africa. In 2006 the North African Region Cheetah Action Group (NARCAG) highlighted the need for genetic analyses to resolve the subspecies status of cheetahs from Algeria (Belachir 2007). The most comprehensive phylogeographic analysis to date was carried out by Charruau and colleagues based on 94 cheetah samples from 18 countries (mitochondrial and microsatellite data), including regions in which cheetahs occur today or are extinct (Charruau et al. 2011). In their study, they show that *A. j. jubatus* and *A. j. raineyi* display very little genetic differentiation, a finding that has previously been reported by O'Brien and colleagues (O'Brien et al. 1987). They further showed that Asiatic cheetahs form a separate phylogenetic group, that split from African cheetahs between 4,700 and 67,400 years ago (ya; Charruau et al. 2011). O'Brien et al. 2017 dated this split to about 6,500 ya. The divergence of *A. j. jubatus* and *A. j. soemmeringii* was dated to 1,600 - 72,300 ya in Charruau et al. (2011) and 5,000 ya in O'Brien et al. (2017), respectively. Phylogeographic analyses using short mitochondrial DNA (mtDNA) fragments further showed a clear distinction of *A. j. hecki* and the other subspecies (Charruau et al. 2011). Unfortunately, the study did not include samples from the current range of *A. j. hecki.* Schmidt-Kuentzel et al. 2018 review the phylogeography of modern-day cheetahs and argue that the phylogeography of *A. j. hecki* can not be assessed due to a lack of genetic data from its current range. Furthermore, a recent study argues that cheetahs show very low genetic differentiation between the subspecies (O'Brien et al. 2017). Unfortunately, this study only included three of the four recognized subspecies (they did not include samples of *A. j. hecki*).

There are approximately 7,100 adult and adolescent wild cheetahs distributed across 33 subpopulations in Africa and Asia (Durant et al. 2017; see Fig. 1). More than half (~60%) of the wild cheetahs occur in one large population in Southern Africa (ssp *A. j. jubatus;* Durant et al. 2017), with fewer than 50-70 found in three populations in Iran (*A. j. venaticus*; Iranian Cheetah Society 2013, Farhadinia et al. 2016). The species' range drastically declined over the past decades and its current extent is likely 9% of their historical distribution (2% in Asia and 13% in Africa; Durant et al. 2017). At the end of the nineteenth century, their distribution comprised most non-rainforest parts of Africa and much of Western and Southern Asia, from the Arabian Peninsula all the way to India, and northwards until Kazakhstan (Durant et al. 2017). Their Asian distribution is presently limited to the central deserts of Iran (Iranian Cheetah Society 2013, Farhadinia et al. 2016). The cheetah is one of the most wide-ranging carnivores with movements of up to 1,000km (Weise et al. 2015) and home ranges up to 3,000km^2 (Marker et al. 2008, Weise et al. 2015). Drivers of decline include armed conflicts, decline in prey, habitat change, human occupancy and related human-cheetah conflict, hunting, and illegal wildlife trade (Ray et al. 2005, Lindsey et al. 2011, Durant et al. 2014, Tricorache et al. 2018). Populations of *A. j. venaticus* in Iran and *A. j. hecki* in northwest Africa are listed as Critically Endangered by the IUCN (Belbachir 2008, Jowkar et al. 2008, Durant et al. 2015). All others are listed as Vulnerable (Durant et al. 2015). However, a recent conservation and threat assessment by Durant et al. (2017) pro-

posed changing the current IUCN Red List threat assessment from Vulnerable to Endangered for the cheetah as a species.

In this study, we investigated genome-wide Single Nucleotide Polymorphisms (SNP), mitochondrial DNA (mtDNA) and major histocompatibility complex (MHC) class II DRB immune response gene data of the four currently recognized subspecies to provide genomic evidence to support subspecies assignments, and on-going and future conservation measures. We aim that our data can build the basis for comprehensive range-wide genetic monitoring of cheetahs through the development of reduced SNP sets. Furthermore, it can be used to guide subspecies specific conservation measures as it calls for a reclassification of cheetah subspecies, and help in evidence-based decision making e.g. for planed re-wilding projects.

Results

To provide genomic evidence to support cheetah conservation and to understand their phylogenetic relationship, we (i) shotgun sequenced two historical cheetah samples from Algeria and Western Sahara, (ii) generated double-digest restriction site associated DNA sequencing (ddRAD) reads for 55 modern individuals (including four parent-offspring trios, only used in the relatedness analysis), (iii) sequenced up to 929 base pairs (bp) of mtDNA for 134 modern and historical museum samples, covering wide parts of their present and historical distribution, and (iv) used amplicon sequencing to investigate MHC class II DRB exon2 haplotypes in 46 modern and historical cheetahs. For a complete sample list see Supplementary Table S1. We further downloaded genomic read data from three East African individuals from Dobrynin et al. 2015 (Genbank: SRR2737543 - SRR2737545). Here, we refer to individuals from East Africa as *A. j. raineyi* for simplicity, but acknowledging that this is currently not a recognized subspecies.

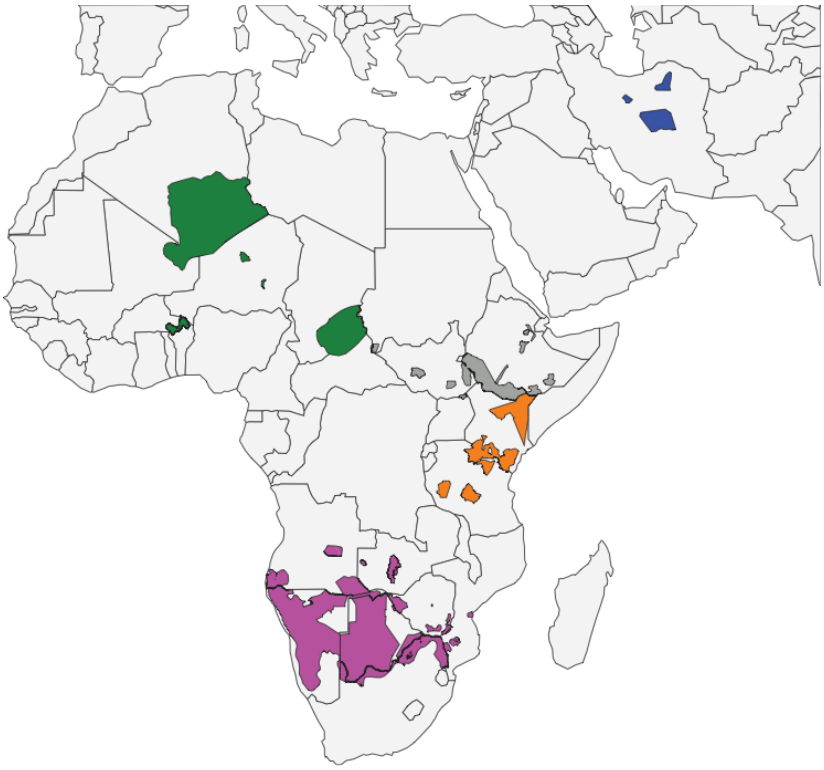

Figure 1. Current distribution of the four recognized subspecies. Blue: *A. j. venaticus*, green: *A. j. hecki*, grey: *A. j. soemmeringii*, orange: *A. j. rainyei* and purple: *A. j. jubatus*. The distribution range was adopted from the IUCN Red List (Durant et al. 2015) and Durant et al. 2017. Subspecies were assigned to the distributions using the results of Charruau et al 2011 and this study.

Distinct genomic differentiation and conservation status of the four cheetah subspecies

Subspecies or conservation unit assignments are crucial to carry out targeted conservation efforts. Therefore, we generated genome-wide SNP data (3,743 SNPs after filtering) for 46 individuals of the four currently recognized cheetah subspecies (see Supplementary Table S1; including the three individuals from Dobrynin et al. 2015). On the contrary to the current classification, the principal component analyses (PCA) supported a genomic differentiation into five clusters (Fig. 2A and Supplementary Fig. S1). The five clusters correspond to the four currently recognized subspecies (*A. j. jubatus*, *A. j. soemmeringii*, *A. j. hecki and A. j. venaticus*) and *A. j. raineyi*. This is further confirmed by admixture analyses (Fig. 2B and Supplementary Fig. S2). The analyses support a separation of the five subspecies with little to no signatures of admixture (Fig. 2B and Supplementary Fig. S2). The limited admixture we detected was between *A. j. raineyi* and either *A. j. soemmeringii* or *A. j. jubatus*. We further ran a separate admixture analysis including all individuals of *A. j. soemmeringii* and *A. j. jubatus.* We did not detect any signatures of admixture between the two subspecies (Supplementary Fig. S3). Genetic distances, measured using F_{ST}, were the highest (0.49696*)* between the two endangered subspecies *A. j. hecki* and *A. j. venaticus* and lowest (0.15660) between *A. j. soemmeringii* and *A. j. raineyi* (Ta-

ble 1). Currently, there is little known about the evolutionary history of the genus *Acinonyx* (summarized in Van Valkenburgh et al. 2018). We thus reconstructed a maximum likelihood tree based on genetic distances from the SNP data (Fig. 2C) using the puma (*Puma concolor*) as an outgroup. The first divergence event separates Asiatic and African cheetahs (with a bootstrap support of 100). We found each subspecies to form monophyletic clades with high support (bootstrap support: 90-100; Fig. 2C). The only subspecies with slightly lower support was *A. j. hecki* (bootstrap support: 70). The splits separating the African subspecies show lower bootstrap support values. Individuals of *A. j. raineyi* form the sister clade to *A. j. soemmeringii* (bootstrap support: 90).

High inbreeding and low heterozygosity threaten the gene pool of the critically endangered Asiatic and (Northwest) African cheetahs

We carried out inbreeding analyses using two different methods, described in Hanghoj et al. (2019) and Fumagalli et al. (2014). Both inferred the highest inbreeding coefficients to be present in samples of the two critically endangered subspecies, *A. j. venaticus* and *A. j. hecki* (Fig. 2D top panel using the method of Fumagalli et al. (2014) and Supplementary Fig. S4 using the method of Hanghoj et al. (2019)), though with slightly different intensities. *A. j. jubatus* showed slightly lower inbreeding coefficients than *A. j. soemmeringii or A. j. raineyi* (Fig. 2D top panel and Supplementary Fig. S4). While most of the *A. j. jubatus* individuals are captive-bred, we do not observe any differences between these and the wild individuals. We further calculated genome-wide heterozygosity for each of the modern samples (excluding the low quality museum samples of *A. j. hecki*), which resulted in a species mean of 0.00040 (range: 0.00020 to 0.00050; Fig. 2D bottom panel). *A. j. soemmeringii* showed a mean of 0.00040 (range: 0.00029 - 0.00050), *A. j. jubatus* a mean of 0.00043 (range: 0.00036 - 0.00050), *A. j. raineyi* a mean of 0.00046 (range: 0.00044 - 0.00048) and *A. j. venaticus* a mean of 0.00029 (range: 0.00020 - 0.00040). We performed relatedness analyses to assess the impact of relatedness on our analyses. We used two different methods, due to the effects of inbreeding on the analyses (Korneliussen and Molte (2015), Hanghoj et al. 2019). Most *A. j. jubatus* and *A. j. soemmeringii* showed relatedness patterns (coefficient of relatedness (r = k2 / 2 + K1) and k0) indicative of 2nd to 4th generation cousins (Supplementary Fig. S5). Some individuals showed sibling or parent-offspring (PO) relationships. Both methods were able to resolve relatedness for four parent-offspring trios (Supplementary Fig. S6). Comparisons between relatedness values (k2) using both methods can be found in Supplementary Fig. S7.

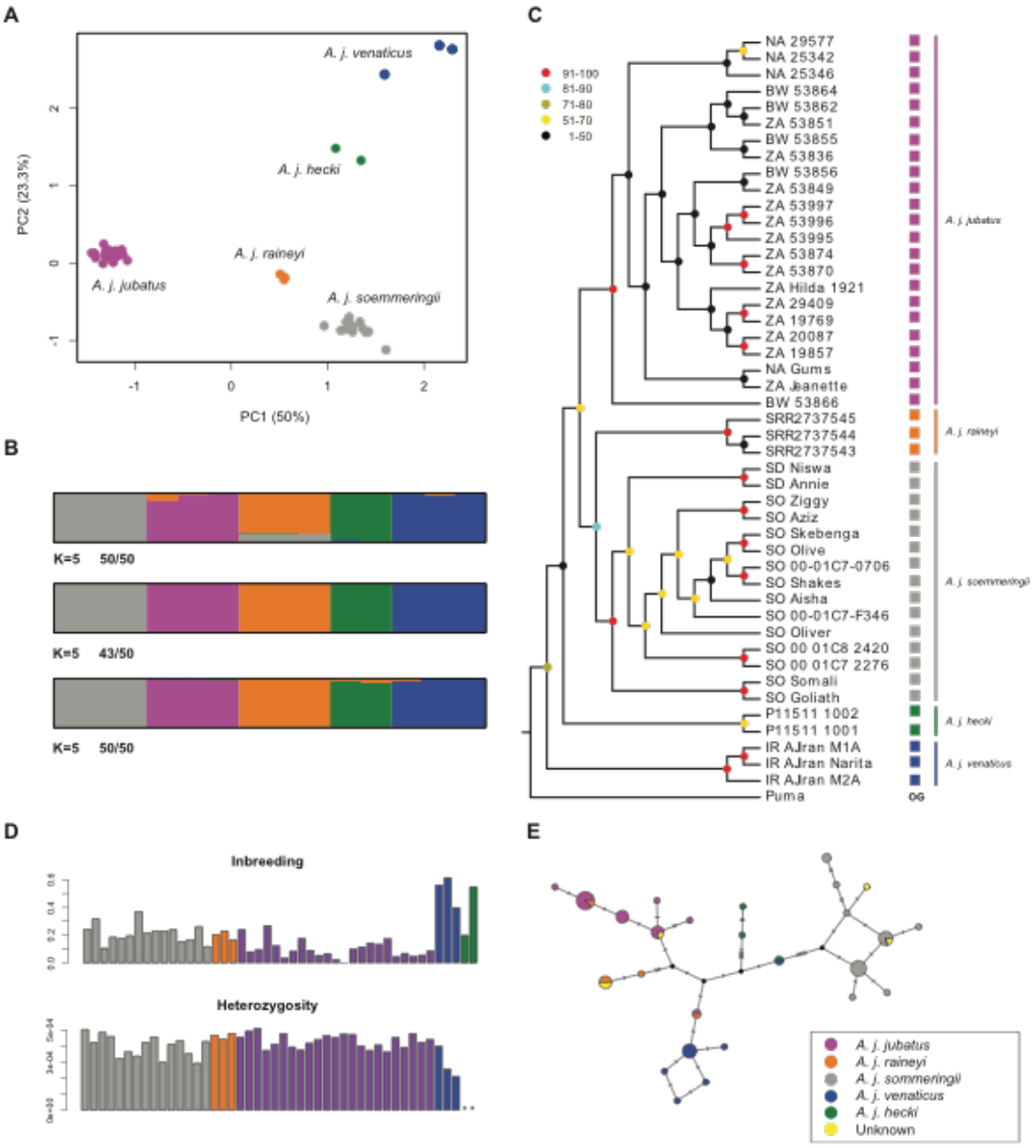

Figure 2: Population genomic analyses of genome-wide SNP data for 43 individuals (A-D) and mitochondrial DNA data from 58 individuals (E). (A) PCA analysis of population structure underlying the 3,743 filtered SNPs. The clustering corresponds to the morphological subspecies classification. Blue: *A. j. venaticus*, green: *A. j. hecki*, grey: *A. j. soemmeringii*, purple: *A. j. jubatus*, orange *A. j. raineyi.* (B) Admixture analyses for K=5 for the three taxon-replicates. Numbers indicate how many individual runs of the 50 replicates support this grouping. Colors as in Fig. 2A. (C) Phylogenetic relationships of representatives for the four cheetah subspecies. Colors as in Fig. 2A. (D) Genome-wide inbreeding and heterozygosity indicating high inbreeding in individuals of *A. j. venaticus* and *A. j. hecki* and low heterozygosity in individuals of *A. j. venaticus.* ** indicates that individuals of *A. j. hecki* were not used in the heterozygosity analysis. Colors as in Fig. 2A. (E) Median-joining haplotype network of 929bp mtDNA. Blue: *A. j. venaticus*, green: *A. j. hecki*, grey: *A. j. soemmeringii*, purple: *A. j. jubatus*, orange *A. j. raineyi* and yellow: unknown origin.

Modern and historic cheetah distribution

The analyses of mitochondrial DNA fragments of different sizes allowed us to investigate genetic population structure throughout most of the cheetahs' present and historic ranges, including parts of the distribution in which cheetahs are extinct today (Fig. 2E). We analyzed mtDNA fragments of 681bp from 134 individuals (Supplementary Fig. S8) and 929bp from 58 individuals (Fig. 2E).

A. j. venaticus - We sequenced 929bp of mtDNA from different countries of *A. j. venaticus*'s past and present distribution, including Iran (N=2), Jordan (N=1), Syria (N=2), Turkmenistan (N=2), Afghanistan (N=1) and India (N=2). All samples apart from one of the two Indian individual form a single cluster with samples from Iran (current distribution; Fig. 2E). Interestingly, this one Indian individual shares a haplotype with an individual from Chad (*A. j. hecki*, Northwest Africa; Fig. 2E). Surprisingly, an individual from Tanzania and one from Zimbabwe show similar haplotypes to *A. j. venaticus* individuals (Fig. 2E, Supplementary Fig. S8).

A. j. hecki - Our sampling of *A. j. hecki* includes four individuals (from Libya, Senegal, Western Sahara and Chad) for the medium and three individuals (all but the Libyan sample) for the long mitochondrial DNA fragments. Both data sets show multiple mutations separating this subspecies from the others (Fig. 2E and Supplementary Fig. S8). All but the sample from Chad form a separate phylogroup.

A. j. soemmeringii - All individuals assigned to *A. j. soemmeringii* form a single, clearly separated phylogroup in both mitochondrial datasets (Fig. 2E, Supplementary Fig. S8). Most notably, all individuals of this subspecies show a 3bp deletion (position: 12665-12667) in the mitochondrial ND5 gene, which has previously been described in Charruau et al. 2011.

A. j. jubatus - Haplotype network reconstructions show a separate phylogroup comprising all individuals belonging to *A. j. jubatus* (apart from the two samples discussed above; Fig. 2E, Supplementary Fig. S8). Even though individuals assigned to *A. j. raineyi* show a close relationship to members of *A. j. jubatus*, most of them make up a distinct phylogroup separated from the latter by five mutations for the longer mtDNA dataset (Fig. 2B).

A. j. raineyi - We included individuals from Tanzania, Kenia and Southern Somalia in the analyses. We found all but two individuals to form a distinct phylogroup, separated from *A. j. jubatus* by 5 mutations (in 929bp; Fig. 2E). One individual shared a haplotype with *A. j. jubatus* individuals, and the other one showed a haplotype closely related to *A.j . venaticus* (discussed above). Interestingly, we also found an individual from Ethiopia to share a haplotype with *A. j. raineyi.* In general, we found a close relationship between haplotypes of *A. j. raineyi* and *A. j. jubatus* (Fig. 2E, Supplementary Fig. S8).

Adaptive immune system diversity in cheetahs

We sequenced the MHC class II DRB exon 2 of 46 individuals (belonging to three out of the four subspecies), which resulted in 13 nucleotide and nine amino-acid (AA) haplotypes (Fig. 3A-B, Supplementary Table S2). The most common AA haplotype was *AcjuFLA-DBR*ha16* carried by

83% of all the individuals followed by AcjuFLA-DBR*ha17 and *AcjuFLA-DBR*ha20* present in 35% and 28%, respectively (Supplementary Table S2). We found *AcjuFLA-DBR*ha21* and *AcjuFLA-DBR*ha23* present only in *A. j. jubatus*. Out of the nine AA haplotypes, we found all to be present in *A. j. jubatus* (one was only present in *A. j. raineyi),* and five both in *A. j. soemmeringii* and *A. j. venaticus*. We found up to four different alleles to be present in single individuals (Supplementary Table S2). We inferred a haplotype diversity of 0.834 (standard deviation (std): 0.028), a nucleotide diversity (π) of 0.069 (std: 0.005) and an average of 16.2 nucleotide differences (k) for our sampling.

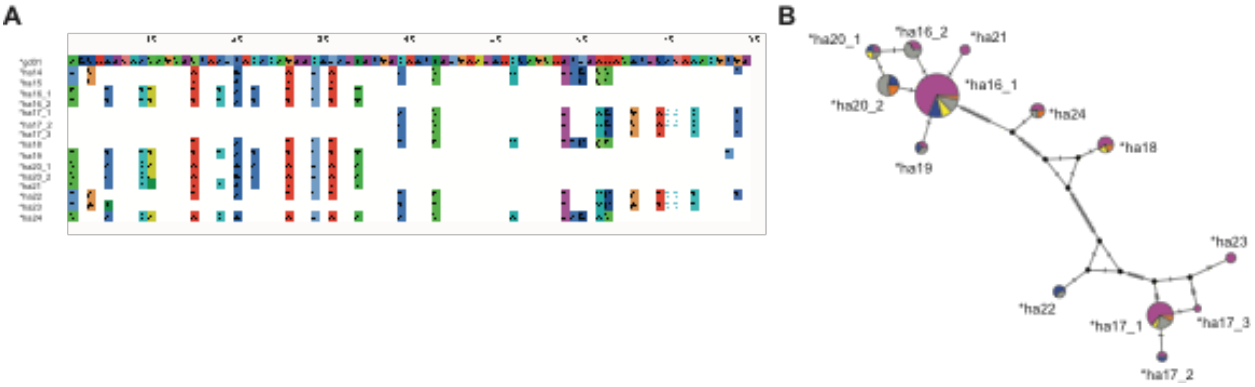

Figure 3: (A) Amino acid sequence and (B) a median joining network of the nucleotide sequences of the MHC II DBR exon 2. Here we abbreviated AcjuFLA-DBR*ha with *ha.

	A. j. jubatus	***A. j. venaticus***	***A. j. hecki***	***A. j. raineyi***
A. j. soemmeringii	0.19282	0.36045	0.39937	0.15660
A. j. jubatus		0.37502	0.41035	0.19232
A. j. venaticus			0.49696	0.47454
A. j. hecki				0.36941

Table 1: F_{ST} values between the five subspecies of cheetah.

Discussion

Conservation status and implications

Here, we present the first genome-wide nuclear DNA support for the distinction of cheetah subspecies. PCA-based clustering, the limited and for the most part lacking admixture between subspecies, phylogenetic analyses, and high F_{ST} values support five distinct groups, corresponding to the morphologically described subspecies: *A. j. jubatus*, *A. j. soemmeringii*, *A. j. venaticus*, *A. j. hecki* and *A. j. raineyi*. Although F_{ST} differentiation between cheetah subspecies has previously been suggested to be low (O'Brien et al. 2017), we show that genome-wide F_{ST} values (0.19 - 0.49) are similar if not slightly higher than that of other large felids such as tigers (0.11 - 0.43; Liu et al. 2018).

The cheetah as a species is classified as Vulnerable by the IUCN with a decreasing population trend (Durant et al. 2015). Today all remaining wild cheetahs are distributed across 33 fragmented subpopulations (Durant et al. 2017). Two subspecies, *A. j. venaticus* (in Iran) and *A. j. hecki* (northwest Africa) are listed as Critically Endangered by the IUCN and indeed we found the highest levels of inbreeding in these two subspecies. Furthermore, *A. j. venaticus* also showed very low genome-wide heterozygosity (*A. j. hecki* was excluded from this analysis, because of the low DNA quality). Of all the subspecies, *A. j. jubatus* showed the lowest level of inbreeding. Genome-wide heterozygosity was the highest in *A. j. jubatus* and *A. j. raineyi*. This is not surprising as *A. j. jubatus* makes up the largest continuous population of all cheetahs (Durant et al. 2017). We have to caution that most of our *A. j. jubatus* individuals (18 out of 25) originated in captivity. However, we did not observe any differences in the genome-wide heterozygosity estimates between our captive-bred and wild caught individuals. Recently, Durant et al. (2017) argued for the reclassification of the cheetah as Endangered based on their demographic, conservation and threat assessment. Our genetic analyses show that cheetahs have genome-wide heterozygosity values (0.0002 to 0.0005) lower than those of other endangered big cats, such as tigers (0.00049 to 0.00073, Cho et al. 2013; listed as Endangered) and lions (0.00048 to 0.00058, Cho et al. 2013; listed as Vulnerable), but higher than the Snow leopard (only one individual: 0.00023; Cho et al. 2013; listed as Vulnerable).

Cheetahs are threatened by armed and human-wildlife conflict, decline in prey, habitat fragmentation, illegal persecution and hunting, and illegal trade (Ray et al. 2005, Lindsey et al. 2011, Durant et al. 2014, Tricorache et al. 2018). We argue that our findings have several implications for the conservation of cheetah as well as highlighting the need for further genetic studies and monitoring.

(1) *Subspecies specific conservation strategies* - The current IUCN classification for cheetahs recognizes only four subspecies, after *A. j. raineyi* has recently been included into *A. j. jubatus* (Kitchener et al. 2017). This was based on the close relationships inferred from mitochondrial DNA data (Charruau et al. 2011). Here we show that this classification should be revisited. Based on nuclear SNP data, *A. j. raineyi* makes up its own group with a closer relationship to *A. j. soemmeringii* than to *A. j. jubatus*. This has important conservation implication as there are about 2,000 individuals of *A. j. raineyi*, about 4,300 of *A. j. jubatus* and only about 300 of *A. j. soemmeringii* left in the wild. Furthermore, our analyses indicate possible admixture between *A. j. raineyi* and *A. j. soemmeringii.*

(2) *Development of efficient range-wide genetic monitoring strategies* - We generated genome-wide data for all subspecies, which can function as a baseline for the development of reduced SNP sets, e. g. for genotyping using SNP arrays or real-time PCRs enabling more cost effective and large-scale monitoring. However, more samples of the two critically endangered subspecies *A. j. venaticus* and *A. j. hecki,* and *A. j. raineyi* will be needed to avoid ascertainment bias in the selected SNP sets.

(3) *The potential for genetic monitoring of illegal wildlife trade of cheetahs* - Northeast Africa is a poaching hot-spot for the illegal cheetah pet trade, mostly to the Gulf states (Nowell 2014, Tricorache et al. 2018). It is also likely the region with the greatest negative impact of illegal trade on wild populations of cheetahs (Nowell 2014). Individuals are likely transported to the Gulf via Somalia and Yemen. However, not much is known about the regional origins of these animals. Information from interdictions and interviews with traders suggest potential origins from opportunistic collections in ethnic Somali regions such as Ethiopia and Kenya (Nowell 2014). Interestingly, northeast Africa is the contact zone between the two subspecies *A. j. soemmeringii* and *A. j. raineyi* (currently listes as *A. j. jubatus*). Previous studies along with our current findings indicate the presence of *A. j. soemmeringii* in South Sudan, Ethiopia and Northern Somalia, and *A. j. jubatus* syn. *raineyi* in Kenya, Tanzania, Uganda and Southern Somalia (see also Charruau et al. 2011 and this study). Simple subspecies distinctions for illegally traded individuals and products will thus help us to quantify the respective proportion of the two subspecies in the trade, and ultimately the importance of different northeast African regions as potential countries of origin. This can form the base for targeted programs to reduce poaching and the illegal wildlife trade of cheetahs in these countries, and allow evidence-based release of confiscated animals into the wild. However, we have to caution that this might be complicated by possible admixture between the two, which could result in exchange of mitochondrial haplotypes. Indeed we found an haplotype assigned to *A. j. raineyi* in an individual from Ethiopia. Unfortunately, we do not have nuclear data for this individual to investigate its nuclear DNA subspecies signature. Reduced SNP set based technologies might help to overcome this potential issue.

(4) *Environmental change and genetic diversity* - Environmental changes put stress on cheetahs throughout their whole range. Previous studies have shown the negative effects of stress on immune response and survival in small populations (Dhabhar 2014). Immunocompetence is an important factor for the survival of a species and thus in conservation biology. It is influenced by genetic factors such as the MHC and the environment in which individuals live in (Frankham et al. 2002). For more than two decades the cheetah has been a popular textbook example for a species with low genetic diversity, especially at MHC. Depleted immune gene diversity was previously supported by the cheetah's ability to accept reciprocal skin grafts from unrelated individuals (O'Brien et al. 1983) and by genetic analyses (e.g. O'Brien and S.J., Yuhki 1999). However, these findings have been debated after Castro-Prieto et al. (2011) and others (e.g. Drake et al. 2004) investigated allele diversity in a large sampling of cheetahs from the wild. Castro-Prieto et al. 2011 were able to detect a much higher genetic diversity within MHC I compared to previous studies, which they attributed to a higher sample size in their study (149 cheetahs from Namibia). However, they were not able to find any further MHC II-DRB alleles than the four previously described in Drake et al. 2004. On the contrary, our sampling of 46 individuals from three subspecies (including the currently not recognized *A. j. raineyi* subspecies) resulted in nine MHC II-

DRB alleles. While our results are encouraging, cheetahs still show MHC II-DRB diversities lower than other large felids, such as Bengal tigers (4 alleles in 16 individuals; *Panthera tigris tigris*; Pokorny et al. (2010)) or the Eurasian lynx in China (16 alleles in 13 individuals; *Lynx lynx*; Wang et al. (2009)), but higher than endangered canids such as Grey wolves in North America (17 alleles in 175 individuals; *Canis lupus*, Kennedy et al. (2007);) and African wild dogs (17 alleles in 368 individuals; *Lycaon pictus*; Marsden et al. (2009)).

Modern and historic cheetah distribution and its importance for conservation management

Evolutionary history

The oldest fossils assigned to the *Acinonyx* date to 3.5-3 mya and were discovered in East-Africa (Krausman and Morales. 2005). Modern cheetah appeared in Africa around 1.9 mya (Van Valkenburgh et al. 2018). The earliest fossils were discovered in South Africa, followed by slightly younger ones in Eastern Africa (Werdelin and Peigne, 2010). We found the earliest split in the phylogenetic analysis of our genome-wide SNP data to be that of the Asiatic and all the African subspecies (Fig. 2C). This split shows a bootstrap support of 74, The phylogenetic separation of the African subspecies showed much lower bootstrap support (48-90). Each African subspecies, however, made up its own monophyletic clade with bootstrap support of 70-100. Charrau et al. 2011 dated the split between Asian and African cheetahs to about 4,700 to 67,400 ya, so much younger than the fossil age of modern cheetahs. A possible explanation could be that the modern cheetah gene pool originated in Northern Africa and individuals expanded from there into Asia and the rest of Africa. This would also fit with the finding of an *A. j. venaticus* haplotype in Egypt described in Charrau et al. 2011. Alternatively, the presence of this haplotype could also be explained by a back migration into Africa from Western Asia. Genome-wide data or complete mtDNA genomes of samples from the North of Africa will be needed to build a more detailed picture about the evolutionary history of modern cheetahs.

Subspecies status of the northwest African cheetahs

The North African Region Cheetah Action Group (NARCAG; Belbachir 2007) recommended genetic studies to identify the subspecies status for the Saharan cheetah population of Algeria, to clarify whether these individuals belong to *A. j. venaticus* or *A. j. hecki*. We included one Algerian sample in the genome-wide SNP data analyses. All analyses place this individual into the *A. j. hecki* subspecies (Fig. 2A,C,E). Mitochondrial haplotype network analyses showed a separate phylogroup for *A. j. hecki* individuals (Libya, Senegal and Western Sahara; former range states). Interestingly, the individual from Chad fell outside this phylogroup, which highlights the need for further genomic investigations of cheetahs belonging to the Chad population.

Distribution of A. j. venaticus

Most individuals of *A. j. venaticus* make up a clearly distinct phylogroup in the mitochondrial DNA network analyses (Fig. 2E and Supplementary Fig. S8). These include individuals from Afghanistan (N_L=1 and N_M=1), India (N_L=1 and N_M=1), Iran (N_L=2 and N_M=11), Jordan (N_L=1 and N_M=2), Syria (N_L=2 and N_M=2) and Turkmenistan (N_L=2 and N_M=2) for the long (L) and medium (M) mtDNA fragment. Interestingly, one individual from India (N_L=1 and N_M=1), which was home to the *A. j. venaticus* subspecies showed a haplotype also found in Chad. It is well documented

that imports of tamed hunting cheetahs from northeastern (Pocock 1939) and eastern Africa (Divyabhanusinh 2007) into India and the Arabian Peninsula were a regular occurrence during the European colonial era, which could explain this finding. An individual from Tanzania and one from Zimbabwe showed similar haplotypes to *A. j. venaticus* individuals (Fig. 2E, Supplementary Fig. S8). More data, such as complete mitochondrial genomes or genome-wide SNP data will be necessary to investigate whether this is a real signal or only an artifact caused by the short length of the mtDNA fragments. Furthermore, we cannot rule out that these samples have been mislabeled sometime after their field collection. Lastly, we did not find any significant homology to nuclear mitochondrial DNA (NUMTs) copies in the published cheetah genome (Dobrynin et al. 2015).

Reintroductions of cheetahs in Asia

Several reintroduction strategies have been explored over the last years by former cheetah-range countries, including India (Ranjitsinh and Jhala 2010). Frequently, the reasons for animal reintroductions include conservation of the species as well as expanded tourism (Boast et al. 2018). Most reintroduction studies focus on the current carrying capacities of different potential areas for release. Ranjitsinh and Jhala (2010) found several Indian national parks to be suitable for reintroductions after extended preparation and resource investments (Ranjitsinh and Jhala 2010). However, genetic studies of regionally extinct individuals using museum collections are lacking, but would be crucial to assess past genetic structure in these regions and assign individuals to their respective subspecies. Unfortunately, our sampling only included two individuals from India for the two mtDNA fragments. While one individual clustered with a sample from Chad, the other one clustered with individuals assigned to *A. j. venaticus* (which is the suspected subspecies for cheetahs from India). The one Indian individual in Charruau et al. 2011 also showed an *A. j. venaticus* haplotype. In September 2009 Indian and International experts, at the Consultative Meeting in Gajner, suggested to introduce individuals from Africa to India (Ranjitsinh and Khala 2010), as the current wild populations of *A. j. venaticus* are highly threatened and only about 50-70 individuals remain in the wild. This was further supported by a small-scale multi-locus genetic analysis (O'Brien et al. 2017). In this study the authors argue that cheetah subspecies are very closely related and that genetic distances between Asian and African cheetah subspecies are equal to those within Africa, and suggested the introduction of African cheetahs to India. However, our genome-wide data shows that differentiation in cheetahs (average F_{ST} of 0.34 for cheetah subspecies) is higher than that found in other large endangered felids such as the tiger (0.27; Cho et al. 2013), and indicate a strong genome-wide differentiation of *A. j. venaticus* and the African subspecies. Based on our genome-wide data we argue against a release of African cheetahs in India, and for more genetic research to be carried out before a potential introduction of African cheetah subspecies, especially in the light of a substantial lack of information on regional adaptation in the different subspecies.

Materials and Methods

Voucher and individual identifiers for samples used in this study can be found in Supplementary Table S1. Samples were imported under the following CITES numbers: AT 16-E-0753, 16S-G006329CR, 15JP001990/TE, 11US761881/9, AT 15-E-1769, D79/DFF. Additionally, we trans-

ferred samples between CITES registered institutions (see Supplementary Table S3 for the institution names and their CITES registration code).

Laboratory Procedures Museum Samples (mtDNA)

DNA extraction - We followed the protocol of Rohland *et al.* (2010) for the DNA extraction from museum samples. To avoid DNA contamination we carried out all extractions in a dedicated laboratory for museum samples.

DNA amplification - We targeted two mitochondrial genes including 14 previously described diagnostic SNPs from Charruau *et al. (*2011) of the NADH-dehydrogenase subunit 5 (MT-ND5) and the control region (MT-CR). Four sets of primers yielding PCR products between 245 and 375 bp were used to amplify a final mtDNA fragment of 929 bp for 56 individuals and 679 bp for 134 individuals (for primer sequences see Supplementary Table S4). To avoid contamination, PCR reactions were prepared in a separate laboratory and the DNA extractions were added to the solutions in another room. Negative controls were included in each PCR reaction. The purified PCR products were sent for Sanger sequencing in both directions at Macrogen Europe Inc. (Amsterdam, Netherlands) and LCG Genomics (Berlin, Germany).

Laboratory Procedures Museum Samples (nuclear DNA)

DNA extraction, Library preparation and sequencing - We extracted the DNA of two museum samples of *A. j. hecki* using the DNA extraction protocol developed by Dabney et al. (2013). This protocol is optimized for retaining short DNA fragments typical for highly degraded historical samples. Illumina sequencing libraries were prepared following the double-indexing strategy of Meyer and Kircher 2010. The samples were pooled in equimolar amounts and sequenced on the Illumina (San Diego, CA, USA) HiSeqX platform at SciLife, Stockholm, Sweden.

Laboratory Procedures Modern Samples (nuclear DNA)

DNA extraction

We extracted DNA from 21 modern cheetah and one *Puma concolor* tissue sample using the Quiagen DNeasy Blood and Tissue kit (Qiagen, Venlo, Netherlands) and 32 diluted blood samples using the innuPREP Blood Kit (Analytik Jena AG, Jena, Germany).

Modern Samples - Double-digest RAD sequencing: library preparation and sequencing

Double-digest RAD sequencing was carried out for 53 samples, including the outgroup species *Puma concolor* by IGA Technologies, Udine, Italy. In brief, *in silico* analysis of the best combination of two restriction enzymes was carried out using simrad (Lepais and Weir 2014) and the Genbank aciJub1 cheetah genome assembly (GCA_001443585.1; Dobrynin et al. 2015). The double-digestion and library preparation was carried out following Peterson et al. 2012, using the SphI and HindIII enzymes. Sequencing was carried out on the HiSeq2500 instrument (Illumina, San Diego, CA, USA) using V4 chemistry and paired end 125bp mode.

Data processing and Analyses (nuclear DNA)

First, we assessed the raw read quality was using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Next, we mapped the read data against the Aci_jub_2 cheetah genome assembly (GCA_003709585.1; Dobrynin et al. 2015) using BWA mem (version 0.7.17-r1188; Li & Durbin, 2010) and processed the resulting mapping files using samtools (version 1.9; Li et al., 2009). We subsequently assessed the mapping quality using Qualimap (Okonechnikov et al. 2015). We carried out adapter-trimming and duplicate removal for the two museum samples using AdapterRemoval2 (Schubert et al. 2016) and Picard (https://broadinstitute.github.io/picard/), respectively. The two samples showed average coverages of 3.8x and 4.9x, respectively, after mapping and processing using bwa (version 0.7.17-r1188; Li & Durbin, 2010) and samtools (version 1.9; Li et al., 2009).

The resulting mapping files were then processed using ANGSD (Korneliussen et al. 2014), which was specifically developed for population genomic analyses of low coverage data. First, we carried out filtering using SNPcleaner (v 2.24; Fumagalli et al. 2014). To do so, we created a single diploid consensus sequence for all samples using samtools (version 1.9, Li et al. 2009). We then filtered the resulting vcf file for (a) the presence of no more than 25% of missing data for individual sites, (b) a maximum coverage of 5,000 to avoid calling sites in highly repetitive regions, and (c) a minimum coverage of 3x for each individual. We then extracted the first two columns of the resulting bed file to generate the filtered SNP files for the subsequent ANGSD analyses.

We included 43 individuals in the population genomic analyses. First, we carried out principal component analyses using ANGSD and pcangsd (Meisner and Albrechtsen 2018). Next, we looked for signatures of admixture using ANGSD and ngsAdmix (Skotte et al. 2013). In order to avoid sampling biases, due to the different numbers of samples per subspecies, we generated different sets of three randomly chosen individuals for each subspecies (except for *A. j. venaticus* and *A. j. hecki* where we only had three and two in our sampling, respectively). We further carried out a separate ngsAdmix analysis restricted to all individuals of the two subspecies *A. j. soemmerningii* and *A. j. jubatus*. We carried out 50 replicates for all ngsAdmix run ranging from k=2 to k=5. The results were analyzed and visualized using CLUPMACK (Kopelman et al. 2015). In order to investigate subspecies differentiation, we carried out F_{ST} analyses using ANGSD (realSFS). Phylogenetic analyses were carried out using a combination of ANGSD, ngsDist (Fumagalli et al. 2014) and FastME (Lefort et al. 2015). Inbreeding was investigated using ngsF (Fumagalli et al. 2014) and ngsRelate (v2; Hanghoj et al. 2019). We further used ngsRelate (v1: Korneliussen and Moltke 2015 and v2: Hanghoj et al. 2019) to carry out relatedness analyses. In order to check its performance, we first ran the tool using data from four parent-offspring trios. Lastly, we carried out heterozygosity analyses using ANGSD and realSFS (part of the ANGSD package). Here, we did not restrict our analyses to filtered sites, to be able to compare our estimates to published genome-wide heterozygosity values of other felids. Due to low coverage and quality, typical for degraded museum DNA, we removed the two museum samples from this analysis.

Mitochondrial DNA data

For the mitochondrial DNA data we aligned all sequences using Codon Code Aligner v3.0.2 (Codon Code Corporation). We obtained the reference sequence for the cheetah mitochondrial genome from GenBank (accession number NC_005212.1; Burger *et al.* 2004)). We carried out parallel analyses on the two concatenated datasets that differed in the length of the fragment and the number of individuals (Supplementary Table S1). The first analysis comprised the largest mtDNA fragment of 929 bp amplified in 58 individuals. The second line of analysis incorporated the alignment of 78 individuals from Charruau *et al.* (2011) and 57 generated in this study. Median-joining networks were created using the freely available software tool, Popart (Leigh and Bryant 2015). Furthermore, we investigated the homology of the mtDNA fragments to known NUMTs in the published cheetah genome (GCA_003709585.1, Dobrynin et al .2015) using Blast.

MHC DNA data

DNA extraction, Amplification and Sequencing

We used the Qiagen DNeasy Blood and Tissue kit for DNA extraction from hair and tissue samples, and the VWR PeqGold™ Tissue DNA Mini Kit Plus for blood samples. We carried out Polymerase Chain Reaction (PCR) as described in Castro-Prieto et al. (2011) using the primers: DRB_SL-F (GCGTCAGTGTCTTCCAGGAG) and DRB_SL-R (GGGACCCAGTCTCTGTCTCA). Indexing, multiplexing and sequencing was carried out following the Illumina Nextera XT DNA LibraryPrep reference guide. Sequencing was carried out on an Illumina MiSeq (2x250bp). We then mapped the Illumina sequencing reads against a reference Sanger sequence using BWA version 0.7.11 (Li & Durbin, 2010) and further processed the mapping file using samtools version 1.3.1 (Li et al., 2009). We called variants using Picard v. 2.8.2 (http://broadinstitute.github.io/picard) and GATK v.3.1.8 (McKenna et al., 2010). The main alleles were phased using the FastaAlternateReferenceMaker command in GATK, based on a minimum of 6 reads per sample, and verified manually by visualization of the coordinate-sorted BAM files using IGViewer (James et al., 2017). We estimated haplotype diversity (Hd) and nucleotide diversity (π) using DNAsp (Librado and Rozas et al., 2009).

Acknowledgements

We thank the curators of the Natural History Museums of London, Vienna, Berlin, Stuttgart and Senckenberg-Frankfurt, the Alexander Koenig Research Museum, the Harrison Institute, the Royal Museum of Scotland and the Royal Museum for Central Africa (Tervuren). We appreciate the support from the veterinarians and keepers of DECAN Rescue Centre Djibouti, Vienna Zoo, Salzburg Zoo, La Palmyre Zoo and the Zoological Garden of Montpellier. The authors further acknowledge support from the National Genomics Infrastructure in Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council, and SNIC/Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with massively parallel sequencing and access to the UPPMAX computational infrastructure. The authors also thank Henrique Leitão and Thijs Hofstede for help processing the

museum specimens in the aDNA lab. PB acknowledges funding from the Austrian Science Foundation FWF P29623-B25. We thank Christine and Urs Breitenmoser, Sven Winter and Menno de-Jong for their helpful comments on the manuscript.

References

Belbachir, F. 2007. Les grands questions relative a la conservation des grands felins d'Algerie: cas du guepard et du leopard. In: R. Berzins and F. Belbachir (eds), *Compte-rendu de la deuxième réunion de l'Observatoire du Guépard en Régions d'Afrique du Nord (OGRAN), 20-25 Novembre 2006, Tamanrasset, Algé*, pp. 8-10. Société Zoologique de Paris (SZP), Paris, France.

Belbachir, F. 2008. *Acinonyx jubatus ssp. hecki.* The IUCN Red List of Threatened Species 2008: e.T221A13035738. http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T221A13035738.en

Boast, L. K., Chelysheva, E. V., van der Merwe, V., Schmidt-Küntzel, A., Walker, E. H., Cilliers, D., ... & Marker, L. 2018. Cheetah translocation and reintroduction programs: past, present, and future. *Cheetahs: biology and conservation. Academic Press, London*, 275-289.

Castro-Prieto A, Wachter B, Sommer S. 2011. Cheetah paradigm revisited: MHC diversity in the world's largest free-ranging population. Mol Biol Evol. 28(4):1455–68.

Charruau, P., Fernandes, C., OROZCO−terWENGEL, P., Peters, J., Hunter, L., Ziaie, H., Jourabchian, A., Jowkar, H., Schaller, G., Ostrowski, S. and Vercammen, P., 2011. Phylogeography, genetic structure and population divergence time of cheetahs in Africa and Asia: evidence for long−term geographic isolates. *Molecular Ecology*, *20*(4), pp.706-724.

Cho, Y. S., Hu, L., Hou, H., Lee, H., Xu, J., Kwon, S., ... & Shin, Y. A. 2013 The tiger genome and comparative analysis with lion and snow leopard genomes. *Nature communications*, *4*, 2433.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M. T., Weihmann, A., Nickel, B., ... & Meyer, M. 2013 Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences*, *110*(39), 15758-15763.

Dhabhar, F.S., 2014. Effects of stress on immune function: the good, the bad, and the beautiful. Immunologic research, 58(2-3), pp.193-210.

Divyabhanusinh C. The End of a Trail. 3rd edn. Oxford: Oxford University Press; 2007. p. 307.

Dobrynin P, Liu S, Tamazian G, Xiong Z, Yurchenko AA, Krasheninnikova K, Kliver S, Schmidt-Küntzel A, Koepfli KP, Johnson W, Kuderna LF. Genomic legacy of the African cheetah, Acinonyx jubatus. Genome biology. 2015 Dec;16(1):277.

Drake, G.J.C., Kennedy, L.J., Auty, H.K., Ryvar, R., Ollier, W.E.R., Kitchener, A.C., Freeman, A.R. and Radford, A.D., 2004. The use of reference strand−mediated conformational analysis for the study of cheetah (Acinonyx jubatus) feline leucocyte antigen class II DRB polymorphisms. *Molecular Ecology*, *13*(1), pp.221-229.

Durant, S.M., Mitchell, N., Groom, R., Pettorelli, N., Ipavec, A., Jacobson, A.P., Woodroffe, R., Böhm, M., Hunter, L.T., Becker, M.S. and Broekhuis, F., 2017. The global decline of cheetah Acinonyx jubatus and what it means for conservation. *Proceedings of the National Academy of Sciences*, *114*(3), pp.528-533.

Durant, S., Mitchell, N., Ipavec, A. & Groom, R. 2015. *Acinonyx jubatus*. *The IUCN Red List of Threatened Species* 2015: e.T219A50649567. http://dx.doi.org/10.2305/IUCN.UK.2015-4.RLTS.T219A50649567.en. Downloaded on 04 August 2019.

Durant, S.M., Wacher, T., Bashir, S., Woodroffe, R.D., De Ornellas, P., Ransom, C., Newby, J., Abáigar, T., Abdelgadir, M., El Alqamy, H. and Baillie, J., 2014. Fiddling in biodiversity hotspots while deserts burn? Collapse of the S ahara's megafauna. *Diversity and Distributions*, *20*(1), pp. 114-122.

Farhadinia, M. S., Akbari, H., Eslami, M., & Adibi, M. A. 2016. A review of ecology and conservation status of Asiatic cheetah in Iran. Cat News Special Issue Iran, 10, 18-26.

Frankham, R., Briscoe, D.A. and Ballou, J.D., 2002. *Introduction to conservation genetics*. Cambridge university press.

Fumagalli, M., Vieira, F. G., Linderoth, T., & Nielsen, R. 2014 ngsTools: methods for population genetics analyses from next-generation sequencing data. *Bioinformatics*, *30*(10), 1486-1487.

Gansauge, M. T., & Meyer, M. 2014 Selective enrichment of damaged DNA molecules for ancient genome sequencing. *Genome research*, *24*(9), 1543-1549.

Hanghøj, K., Moltke, I., Andersen, P. A., Manica, A., & Korneliussen, T. S. 2019. Fast and accurate relatedness estimation from high-throughput sequencing data in the presence of inbreeding. *GigaScience*, *8*(5), giz034.

Iranian Cheetah Society. 2013. More than 40% of Cheetahs Killed on Roads in Iran. Available at: http://www.wildlife.ir/en/2013/01/14/more-than-40-of-cheetahs-killed-on-roads-in-iran/.

James T. Robinson, Helga Thorvaldsdóttir, Aaron M. Wenger, Ahmet Zehir, Jill P. Mesirov. 2017. Variant Review with the Integrative Genomics Viewer (IGV). Cancer Research 77(21) 31-34.

Jowkar, H., Hunter, L., Ziaie, H., Marker, L., Breitenmoser-Wursten, C. & Durant, S. 2008. *Acinonyx jubatus ssp. venaticus*. *The IUCN Red List of Threatened Species 2008*: e.T220A13035342. http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T220A13035342.en

Kennedy LJ, Angles JM, Barnes A, Carmichael LF, Radford AD, Ollier WER, Happ GM. 2007. DLA-DRBI, DQAI, and DQBI alleles and haplotypes in North American grey wolves. J Hered. 98: 491–499.

Kircher M, Sawyer S, Meyer M. 2012. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Res 40: 1–8.

Kitchener, A.C., Breitenmoser-Würsten, C., Eizirik, E., Gentry, A., Werdelin, L., Wilting, A., Yamaguchi, N., Abramov, A.V., Christiansen, P., Driscoll, C. and Duckworth, J.W., 2017. A revised taxonomy of the Felidae: The final report of the Cat Classification Task Force of the IUCN Cat Specialist Group. Cat News.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular ecology resources*, *15*(5), 1179-1191.

Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: analysis of next generation sequencing data. BMC bioinformatics. Dec;15(1):356.

Korneliussen, T. S., & Moltke, I. 2015. NgsRelate: a software tool for estimating pairwise relatedness from next-generation sequencing data. *Bioinformatics*, *31*(24), 4009-4011.

Krausman, P.R. and Morales, S.M., 2005. Acinonyx jubatus. *Mammalian Species*, *2005*(771), pp.1-6.

Lefort, V., Desper, R., & Gascuel, O. 2015. FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Molecular biology and evolution*, *32*(10), 2798-2800.

Leigh, J.W. and Bryant, D., 2015. popart: full−feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), pp.1110-1116.

Lepais O, Weir JT. 2014. Sim RAD: an R package for simulation−based prediction of the number of loci expected in RAD seq and similar genotyping by sequencing approaches. Molecular Ecology Resources. 14(6):1314-21.

Li, H., & Durbin, R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, *26*(5), 589-595.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics. 15;25(16):2078-9.

Librado, P., & Rozas, J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*(11), 1451-1452.

Lindsey, P.A., Romanach, S.S., Tambling, C.J., Chartier, K. and Groom, R., 2011. Ecological and financial impacts of illegal bushmeat trade in Zimbabwe. *Oryx*, *45*(1), pp.96-111.

Marker, L.L., Dickman, A.J., Mills, M.G., Jeo, R.M. and Macdonald, D.W., 2008. Spatial ecology of cheetahs on north−central Namibian farmlands. *Journal of Zoology*, *274*(3), pp.226-238.

Marsden CD, Mable BK, Woodroffe R, Rasmussen GSA, Cleaveland S, McNutt JW, Emmanuel M, Thomas R, Kennedy LJ. 2009. Highly endangered African wild dogs (Lycaon pictus) lack variation at the major histocompatibility complex. J Hered. 100:S54–S65.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... & DePristo, M. A. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, *20*(9), 1297-1303.

Meisner, J. & Albrechtsen, A. 2018. Inferring Population Structure and Admixture Proportions in Low Depth NGS Data. bioRxiv, 302463

Meyer, M. and Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, *2010*(6), pp.pdb-prot5448.

Nowell, K., 2014. An assessment of the conservation impacts of legal and illegal trade in cheetahs Acinonyx jubatus. In *Convention on internation trade in endangered species of wild fauna and flora* (pp. 1-54

O'Brien, S. J., Johnson, W. E., Driscoll, C. A., Dobrynin, P., & Marker, L. 2017. Conservation genetics of the cheetah: lessons learned and new opportunities. *Journal of Heredity*, *108*(6), 671-677.

O'Brien, S.J., Yuhki, N. 1999. Comparative genome organization of the major histocompatibility complex: lessons from the Felidae. Immunological Reviews. Vol. 167: 133-144

O'Brien SJ, Wildt DE, Goldman D, Merril CR, Bush M. 1983. The cheetah is depauperate in genetic variation. Science. 221(4609):459–62.

O'Brien, S.J., Wildt, D.E., Bush, M., Caro, T.M., FitzGibbon, C., Aggundey, I. and Leakey, R.E., 1987. East African cheetahs: evidence for two population bottlenecks?. *Proceedings of the National Academy of Sciences*, *84*(2), pp.508-511

Okonechnikov K, Conesa A, García-Alcalde F. 2015 . Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics. 32(2):292-4.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S. and Hoekstra, H.E., 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS one*, *7*(5).

Pocock RI. 1939. The Fauna of British India, Including Ceylon and Burma. London: Taylor and Francis; p. 549

Pokorny I, Sharma R, Goyal SP, Mishra S, Tiedemann R. 2010. MHC class I and MHC class II DRB gene variability in wild and captive Bengal tigers (Panthera tigris tigris Immunogenetics 62:667–679.

Ranjitsinh, M. K., & Jhala, Y. V. 2010. Assessing the potential for reintroducing the cheetah in India. *Wildlife Trust of India, Noida, & the Wildlife Institute of India, Dehradun, TR2010/001*.

Ray, J.C., Hunter, L. and Zigouris, J., 2005. *Setting conservation and research priorities for larger African carnivores* (Vol. 24 New York: Wildlife Conservation Society.

Rohland, N., Siedel, H., & Hofreiter, M. 2010. A rapid column−based ancient DNA extraction method for increased sample throughput. *Molecular ecology resources*, *10*(4), 677-683.

Schmidt-Kuentzel, A., Dalton, D.L., Menotti-Raymond, M., Fabiano, E., Charruau, P., Johnson, W.E., Sommer, S., Marker, L., Kotze, A. and O'Brien, S.J., 2018. Conservation genetics of the cheetah: Genetic history and implications for conservation. *Cheetahs: Biology and Conservation: Biodiversity of the World: Conservation From Genes to Landscapes*, pp.71-92.

Schubert M, Lindgreen S, Orlando L. 2016. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC research notes. 9(1):88.

Skotte, L., Korneliussen, T. S., & Albrechtsen, A. 2013. Estimating individual admixture proportions from next generation sequencing data. *Genetics*, *195*(3), 693-702.

Tricorache, P., Nowell, K., Wirth, G., Mitchell, N., Boast, L. K., & Marker, L. 2018. Pets and Pelts: Understanding and Combating Poaching and Trafficking in Cheetahs. *Biodiversity of the World—Cheetahs: Biology and Conservation. 1st ed. San Diego: Elsevier*, 191-205.

Van Valkenburgh, B., Pang, B., Cherin, M., Rook, L. 2018. The Cheetah: Evolutionary History and Paleoecology. *Cheetahs: Biology and Conservation: Biodiversity of the World: Conservation From Genes to Landscapes*, pp.25-31.

Wang X, Wei K, Zhang Z, Xu X, Zhang W, Shen F, Zhang L, Yue B. 2009. Major histocompatibility complex class II DRB exon-2 diversity of the Eurasian lynx (Lynx lynx) in China. J Nat Hist. 43:245–257.

Weise, F.J., Lemeris Jr, J.R., Munro, S.J., Bowden, A., Venter, C., van Vuuren, M. and van Vuuren, R.J., 2015. Cheetahs (Acinonyx jubatus) running the gauntlet: An evaluation of translocations into free-range environments in Namibia. *PeerJ*, *3*, p.e1346.

Werdelin, L., Peigné, S., 2010 Carnivora. 603-657.

Werle E, Schneider C, Renner M, Volker M, Fiehn W, 1994. Convenient single-step, one tube purification of PCR products for direct sequencing. Nucleic Acids Res 22: 4354–4355.

SUPPLEMENTARY FIGURES AND TABLES

Supplementary Figure S1: PCA analyses for the genome-wide SNP data (PC1 versus PC3). Blue: *A. j. venaticus*, green: *A. j. hecki*, grey: *A. j. soemmeringii,* purple: *A. j. jubatus,* orange: *A. j. raineyi*.

Supplementary Figure S2: Admixture results for the three independent replicate runs. Results are shown for 50 replicates of K=2, K=3, K=4 and K=5. Each run used different individuals for *A. j. jubatus* and *A. j. soemmeringii*. Only groupings supported by more than four replicates are shown.

Supplementary Figure S3: Admixture results for the *A. j. jubatus* and *A. j. soemmeringii* dataset. Results are shown for 50 replicates of K=2, K=3, K=4 and K=5. Only groupings supported by more than four replicates are shown.

Supplementary Figure S4: Inbreeding values calculated using the method of Hanghoj et al. (2019). Blue: *A. j. venaticus*, green: *A. j. hecki*, grey: *A. j. soemmeringii,* purple: *A. j. jubatus,* orange: *A. j. raineyi*.

Supplementary Figure S5: Plot showing K0 versus the coefficient of relatedness (r) for *A. j. jubatus* and *A. j. soemmeringii.* The expected relatedness are shown for different areas in the plot. PO…parent offspring, Sib…sibling, 2nd…second generation cousins, 3rd…third generation cousins, 4th…fourth generation cousins and UR…unrelated. Grey: *A. j. soemmeringii* and purple: *A. j. jubatus.*

Supplementary Figure S6: Relatedness (K2) of four parent-offspring trios. Top panel: relatedness inferred without accounting for inbreeding, and bottom panel relatedness inferred after correcting for inbreeding. Dark blue indicates a close relationship and a white no relationship between individual pairs.

Supplementary Figure S7: Relatedness (K2) found in the 43 cheetah samples. Top panel: relatedness inferred without accounting for inbreeding, and bottom panel relatedness inferred after correcting for inbreeding. Dark blue indicates a close relationship and a white no relationship between individual pairs.

Supplementary Figure S8: Median-joining haploptype network of 679 bp mitochondrial DNA. Blue: *A. j. venaticus*, green: *A. j. hecki*, grey: *A. j. soemmeringii,* purple: *A. j. jubatus,* orange: *A. j. raineyi*.

Table S1: Samples used in this study.

Table S2: MHC II DBR exon 2 haplotypes and their frequency in the different subspecies.

Table S3: CITES registered institutions and their registration numbers. Samples were transported between institutions under the listed CITES registration numbers.

Table S4: Characteristics of the primers employed in this study. aMT indicates the primers used for highly fragmented museum samples.