African and Asian leopards are highly differentiated at the genomic level

Johanna L.A. Paijmans*[1,2,3], Axel Barlow*[1,4], Matthew S. Becker[5], James A. Cahill[6], Joerns Fickel[1,7], Daniel W.G. Foerster7 , Katrin Gries[8], Stefanie Hartmann[1], Rasmus Worsøe Havmøller[9,10], Kirstin Henneberger[1], Christian Kern[11], Andrew C. Kitchener[12,13], Eline D. Lorenzen[8], Michael V. Westbury[1,8], Frieder Mayer[14], Stephen J. OBrien[15,16], Johanna von Seth[17,18,19], Mikkel-Holder S. Sinding[8], Göran Spong[20], Olga Uphyrkina[21], Bettina Wachter[6], Love Dalén[16,17,18], Jong Bhak[22,23,24], Andrea Manica[3], Michael Hofreiter[1]

[1] *Institute for Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany*
[2] *Department of Genetics & Genome Biology, University of Leicester, Leicester LE1 7RH, UK*
[3] *Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK*
[4] *School of Science and Technology, Nottingham Trent University, Clifton Lane, Nottingham NG11 8NS, UK*
[5] *Zambian Carnivore Programme, PO Box 80 Mfuwe, Eastern Province, Zambia*
[6] *Laboratory of Neurogenetics of Language, Rockefeller University, 1230 York Avenue, New York, NY, 10065, USA*
[7] *Leibniz Institute for Zoo- and Wildlife Research, Alfred-Kowalke-Str. 17, 10315 Berlin, Germany*
[8] *Der Grüne Zoo Wuppertal, Hubertusallee 30, 42117 Wuppertal, Germany*
[9] *GLOBE institute, University of Copenhagen, Oester Voldgade 5-7, 1350, Copenhagen K, Denmark*
[10] *Research and Collections, Natural History Museum of Denmark, University of Copenhagen, Universitetsparken 15, 2100 Copenhagen OE, Denmark*
[11] *Tierpark Berlin-Friedrichsfelde, Am Tierpark 125, 10319 Berlin, Germany*
[12] *Department of Natural Sciences, National Museums Scotland, Chambers Street, Edinburgh EH1 1JF, UK*
[13] *Institute of Geography, School of Geosciences. Drummond Street, University of Edinburgh EH8 9XP, UK*
[14] *Museum für Naturkunde - Leibniz–Institut für Evolutions- und Biodiversitätsforschung, Invalidenstraße 43, 10115 Berlin, Germany*
[15] *Laboratory of Genomics Diversity, Center for Computer Technologies, ITMO University, 49 Kronverkskiy Pr., St. Petersburg, 197101, Russian Federation*
[16] *Guy Harvey Oceanographic Center, Halmos College of Arts and Sciences, Nova Southeastern University, 8000 North Ocean Drive, Ft Lauderdale, Florida 33004 USA*
[17] *Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden*
[18] *Centre for Palaeogenetics, Svante Arrhenius väg 20C, 10691 Stockholm, Sweden*
[19] *Department of Zoology, Stockholm University, SE-10691 Stockholm, Sweden*
[20] *Wildlife, Fish, and Environmental Studies, Swedish University of Agricultural Sciences, 907 83 UMEA, SWEDEN*
[21] *East Asia Terrestrial Biodiversity Scientific Center, 159 Stoletiya Street, Vladivostok, 690022, Russia*
[22] *Korean Genomics Center (KOGIC), Ulsan National Institute of Science and Technology (UNIST), Ulsan, 44919, Republic of Korea*
[23] *Clinomics, UNIST, Ulsan, 44919, Republic of Korea*
[24] *Personal Genomics Institute, Genome Research Foundation, Cheongju, 28160, Republic of Korea*

Lead Contact: Johanna L.A. Paijmans, paijmans.jla@gmail.com

* These authors contributed equally to this manuscript

Keywords: leopards, *Panthera pardus,* genomes, historical samples, leopards, *Panthera pardus,* population genomics

Summary

Leopards are the only big cats still widely distributed across the continents of Africa and Asia. They occur in a wide range of habitats and are often found in close proximity to humans. But despite their ubiquity, leopard phylogeography and population history have not yet been studied with genomic tools. Here, we present population-genomic data from 26 modern and historical samples encompassing the vast geographical distribution of this species. We find that Asian leopards are broadly monophyletic with respect to African leopards across almost their entire nuclear genomes. This profound genetic pattern persists despite the animals' high potential mobility, and despite evidence of transfer of African alleles into Middle Eastern and Central Asian leopard populations within the last 100,000 years. Our results further suggest that Asian leopards originated from a single out-of-Africa dispersal event 500–600 thousand years ago, and are characterised by higher population structuring, stronger isolation-by-distance, and lower heterozygosity than African leopards. Taxonomic categories do not take into account the variability in depth of divergence among subspecies. The deep divergence between the African subspecies and Asian populations contrasts with the much shallower divergence amongst putative Asian subspecies. Reconciling genomic variation and taxonomy is likely to be a growing challenge in the genomics era.

Introduction

Leopards (*Panthera pardus*) are iconic big cats with the largest current distribution of all species within the genus *Panthera*. They are ecological generalists inhabiting semi-desert, savanna, rainforest and montane habitats and spanning an altitudinal range from sea level to 5,200 metres altitude [1,2]. Historically, leopards ranged throughout sub-Saharan and north Africa, and in Asia, from Turkey eastwards to south east Asia and the Russian Far East. Within the past hundred years, populations in many parts of this distribution have suffered declines and are becoming increasingly fragmented [2,3]. As a result, several leopard subspecies are now considered critically endangered or extinct [1,4,5].

A particular focus of genetic studies of leopards has been their subspecies taxonomy [1,2,6–9]. Although alternative taxonomies are used [2], here we follow the most recent taxonomy proposed by Kitchener et al. [5]. This taxonomy recognises eight extant subspecies with African leopards assigned to a single subspecies and the remaining seven subspecies defined for different Asian populations. The validity of these subspecies has been investigated using a range of molecular markers, including allozymes, microsatellites and mitochondrial sequences [1,2]. However, the extent to which genetic patterns inferred from these limited marker sets reflect broad patterns of variation across the nuclear genome as a whole remains unknown. This question is of further applied importance, as subspecies taxonomy currently provides a basis for leopard conservation planning and implementation. Thus, population-genomic investigations are vital to ensure that conservation efforts effectively safeguard genetic as well as taxonomic diversity in leopards.

A second focus of genetic studies of leopards has been their historical biogeography, which has primarily been inferred by analysis of their mitochondrial DNA [2,8,10]. In addition to Africa and Asia, leopards were also widespread in Europe during the Pleistocene, but went extinct during the Holocene [4,11]. The study of mitochondrial genome sequences suggested that both Asia and

Europe were colonised during a single out-of-Africa dispersal event some time during the Middle Pleistocene (400-700 Ka [thousands of years ago]; [2,10]). In contrast, previous mitochondrial and palaeontological studies have proposed multiple migration waves [4,8], and it has been suggested that the geographically isolated Javan leopard *P. p. melas* may represent a relict population from an earlier out-of-Africa migration event [8]. However, mitochondrial data provide limited power to differentiate among these dispersal hypotheses, or to determine the extent of post-colonisation gene flow between continents, due to the idiosyncratic nature of the evolution of single loci in populations.

In this study, we analyse genome data from 26 leopards, sampled across their current African and Asian distribution and representing almost all extant subspecies, in order to investigate the broad-scale genetic structure of leopards at the genomic level. We dissect the biogeographical history of the leopard in unprecedented detail, enabling us to make inferences about the timing and magnitude of the colonisation of Asia. We find evidence for a single out-of-Africa dispersal event, resulting in almost complete genome monophyly of Asian leopards; a pattern that is not reflected in current taxonomy, which applies equal rank to the African and all Asian subspecies. Our results should stimulate the debate on the interplay between taxonomy and conservation in the genomics era.

Results

Intercontinental population structure of the leopard

We generated genome data from 23 leopard specimens, ranging from 3x to 40x genome sequence coverage, comprising five modern samples (blood or tissue) and 18 samples from archival collections (bone or preserved skin, henceforth referred to as 'historical samples') (Figure 1A, Table 1). Together with three previously published individuals [12], our dataset provides genome-level data from 14 African and 12 Asian leopards, representing 7 out of 8 currently recognised extant subspecies (all except *P. p. nimr*) [5]. Mitochondrial genome relationships among the sampled individuals are consistent with those previously recovered, displaying distinct African and Eurasian clades with 100% bootstrap support (Figure S1) [2,8,10].

We also investigated the population structure of leopards using a Principal Components Analysis (PCA) of 2.8 million filtered variable positions (Figure 1B). This analysis revealed strong differentiation of African and Asian leopards along Principal Component (PC) 1, representing almost 20% of the sampled genomic variation. PC2 is considerably smaller (approximately 5% of the variation) and separates leopards within Asia. Among African leopards, the individual from Morocco is also separated along this axis.

We investigated the extent of admixture among the African and Asian population clusters indicated by the PCA using the program NGSadmix [13] to assign, for each individual, the genomic fractions descending from two hypothesised ancestral populations (K=2 as the most likely value of K, further supported by the Δk method [14,15], see STAR Methods). Using this method, the genomes of all African and all but two Asian leopards were assigned to their respective population clusters (Figure 1C, Figure S2). The exceptions were individuals from the Palestine region and from Afghanistan,

which had estimated admixture fractions of 25% and 10% from the African population, respectively.

Effect of intercontinental structure and admixture on leopard genomes

To investigate the extent to which population structuring and admixture have shaped broad-scale patterns of genomic variation among African and Asian leopards, we divided their aligned genome sequences (autosomes only) into 2,206 non-overlapping one-megabase windows, calculated the maximum likelihood phylogeny of each, and generated a summarised Maximum Clade Credibility (MCC) tree annotated with clade frequencies indicating the proportion of genome windows that recover each clade in the tree (Figure 2). In >99% of the genome windows, Asian leopards are monophyletic (Figure 2A). In contrast, African leopards are monophyletic for only 36.9% of the genome windows. Although the percentage of genome windows that support an Asian monophyly reduces when smaller window sizes are used, as expected by incomplete lineage sorting and recombination (99.0%, 95.7%, 89.4% and 73.8% for 1Mb, 500kb, 250kb and 100kb respectively). However, even for the smallest window size tested (10x smaller, or 100Kb) almost 74% of windows support the monophyly of Asian leopards whereas only 14% support African monophyly. At this window size, a notable proportion of trees where Asian leopards are not monophyletic involve a repositioning of either the Palestinian of Afghan leopard, or both, outside of an otherwise intact Asian clade (14.8% of total trees). This suggests admixture with Africa indicated by the NGSadmix analysis is detectable at this smaller window scale, although it does not impact the general pattern of monophyly at the broader 1Mb scale. Overall, this broad pattern of whole-genome monophyly for Asian leopards is consistent with a single out-of-Africa dispersal event suggested previously by mitochondrial studies [2,10], although alternative explanations exist (such as a strong post-colonisation bottleneck, or allele "surfing" at the expansion front [16]).This pattern persists despite gene flow from Africa into Middle Eastern (Palestine region) and Central Asian (Afghanistan) populations.

Dating out-of-Africa dispersal and gene flow

We investigated the timing of leopard dispersal from Africa to Asia using Pairwise Sequentially Markovian Coalescent (PSMC) analysis of individuals with high sequence coverage (>16x, eight individuals). PSMC estimates population-size changes over time from the distribution of coalescence times of two alleles sampled along sections of the genome. For diverging populations, the point at which their respective PSMC curves bifurcate provides an estimate of the end of panmixia in the ancestral population. PSMC curves for African (Namibia, Tanzania, Zambia) and Asian (Nepal, China, Java) individuals start to diverge around 500-600 Ka, suggesting the initial out-of-Africa dispersal occurred around this time (Figure 3A). This date was calculated using the mutation rate previously used for other big cats [17] and later the leopard genome [12]. Although this divergence age may be overestimated if the founding population has not been sampled, it is notably consistent with recent estimates based on mitochondrial DNA (485-710 Ka; [10]; 471-825 Ka [2]).

We estimated the timing of the last gene flow event between Africa and Asia using an extension of

the PSMC method (hPSMC [18]), which involves PSMC analysis of a hybrid diploid genome generated from the haploidised genomes of an African and an Asian individual. When no more coalescent events occur between these two genomes, the population size inferred by PSMC will approach infinity, and can thus be used to date the end of gene flow, at least at the level of less than 0.1 migrants per generation [18]. Pairwise analysis of African and Asian high-coverage individuals produced estimates that are consistent with data simulated based on cessation of gene flow between 0-100 Ka (Figure 3B). Although the upper limit of this time range may represent an overestimate, because the Middle Eastern and Central Asian individuals were not included in this analysis due to their low sequence coverage, it indicates that African and Asian leopards have exchanged alleles relatively recently within the context of their 500-600 Ka initial divergence.

Intracontinental population structure of African and Asian leopards

The phylogenetic analysis of the 2,206 non-overlapping one megabase-genome windows also provides information on population structuring within African and Asian leopards. Among African leopards, phylogenetic analysis (Figure 2B) and PCA (Figure 4A) suggest three major groups: the Moroccan leopard, a group containing two West African leopards (from Gabon and Cameroon), and a group containing the remaining African leopards (including a third West African leopard from Equatorial Guinea). It should be noted that the recovered PCs all have similar loadings (Figure 4A inset figure), suggesting that the observed structuring along PC 1 does not reflect a substantially stronger structuring than subsequent PCs. NGSadmix analysis of only the African samples (most likely value of K is 2; Figure S3A) equally does not reveal strong structuring, and any tentative groupings would require further testing and additional African sampling. Within Asian leopards we also find three major groups, comprising central (Nepal, India, Sri Lanka), western (Palestine region, Afghanistan), and eastern clusters (Java, China, Amur, Korea, Vietnam, Thailand), with the latter two groups recovered in > 50% of phylogenies (Figure 2B). A PCA of Asian leopards further supports these population clusters, which separate along PCs 1 and 2 (Figure 4B). NGSadmix analysis of the Asian samples reveals similar geographical structuring (most likely value of K is 2; Figure S3B). Overall levels of structuring are lower for African leopards than for Asian leopards, as evidenced by generally lower frequencies of recovery for clades (mean clade frequency of 0.13 and 0.46 for Africa and Asia, respectively) (Figure 2B), and similar loadings for consecutive PCs (Figure 4A inset figure).

The position of the Moroccan leopard is difficult to resolve; of the phylogenies that did not return African leopards as monophyletic, a relatively large percentage (31%; or 20% of all phylogenies) place this individual either as sister taxon to all leopards (6.5%), or as sister to all Asian leopards (13.7%; Figure 2A). This frequency is particularly noteworthy, considering that no other topology was represented more than 3%. The morphological distinctiveness of the north African leopard has led to its assignment to a separate subspecies in the past (Barbary leopard, *P. p. panthera*; Schreber, 1777), although all African subspecies were later subsumed into a single subspecies, *P. p. pardus* [1].

We also examined the extent to which relationships among leopard genomes can be predicted by geographic distance (isolation-by-distance), using a linear regression analysis of pairwise genomic and geographical (Euclidean) distances. Although both African and Asian leopards show significant isolation-by-distance (Mantel test $R = 0.42$, $p = 0.02$ and $R = 0.71$, $p = 0.001$, respectively), the size of this effect is considerably lower for African leopards than Asian leopards (Standard Major Axis [SMA] regression = 0.24 and SMA regression = 0.36 for African and Asian leopards, respectively) (Figure 4C).

Finally, we compared genetic diversity in terms of average genome-wide heterozygosity. We restricted this analysis to individuals with > 10x sequencing coverage. Although we did include historical samples in this analysis, we interpret their precise heterozygosity estimates with a degree of caution as these can be affected by properties inherent to historical sequence data (e.g., error rates and low coverage, see Figure S5). Our analyses showed that African leopards are, on average, more than twice as diverse as Asian leopards (Figure 4D), with an estimated mean heterozygote frequency of 0.0026, and no obvious runs of reduced heterozygosity in any individual (Figure S4A). The historical sample from Durban (South Africa) has particularly high heterozygosity, outside of the range of the other African leopards (Figure 4D). Further analysis of both modern and historical samples from this region would be beneficial to investigate if this is a property unique to this genome data recovered from a historical specimen, or if this is inherent to the southern African leopard population. We also found a relatively high heterozygosity of the historical sample from Konde, Pemba Island (an island just north of Zanzibar), would suggest the population that this individual originates from, was not a small, isolated population. Given Pemba Island's proximity to the African mainland (Tanzania), it is conceivable that the island population maintained active gene flow with the mainland populations, although an alternative explanation could be that the sample provenance is incorrect, for example having been traded from the mainland. Asian leopards show a mean heterozygote frequency of 0.0010, which is highest in the historical Indian individual (Figure 4D) and lowest in the animal from the Henry Doorly Zoo and in the Amur leopards, which are from a critically endangered, small population [6,19,20] (Table 1). The zoo animal and the Amur leopards additionally show sequential windows of reduced heterozygosity, indicative of inbreeding (Figure S4B). We also find that individuals from small Asian populations (e.g., Java and Amur) do have much lower heterozygosity than those from larger populations (e.g., India and Nepal), which is consistent with previous results [6].

Discussion

African and Asian leopard genomes are highly distinct

Whole genome monophyly represents the extreme theoretical endpoint of population divergence. This phenomenon is most likely to occur in allopatric populations with small effective population sizes, deep divergence times, and low migration rates. Leopards, in contrast, have a large and continuous distribution, generalist ecology, and high dispersal potential [e.g., 21,22], and yet we find that the Asian leopard clade in the genome phylogeny is broadly monophyletic across almost all of the nuclear genome. Moreover, this phylogenomic pattern persisted despite evidence of transfer of alleles from Africa.

This pattern is even more striking when comparable analyses from other taxa are considered. At the smaller window size of 100 kb, Asian leopards are monphyletic in 74% of windows. At the same window size, previous studies have found brown bears (*Ursus arctos*) to be monophyletic with respect to their sister species, polar bears (*U. maritimus*), across only 66% of their aligned genomes (polar bears, in contrast, show 99% monophyly) [23]. Within the genus *Panthera*, analysis of the five currently recognised species indicates that as much as 35% of their aligned genome sequences failed to recover their correct evolutionary relationship [24]. Also noteworthy is the Middle Pleistocene (500-600 Ka) divergence time of African and Asian leopards indicated by the PSMC analyses. Simulation studies show that reaching reciprocal monophyly is expected to take millions of years for species with large effective population sizes (> 100,000) and moderate generation times (> 1 year) [25]. Therefore, the most plausible scenario is that leopards colonised Asia in a single out-of-Africa dispersal event, involving a small and closely related genetic subset of the African source population and resulting in a strong founder effect. Remarkably, the cohesion of the Asian clade as a whole has been retained for more than half a million years despite the transfer of alleles from Africa. Thus, genetic exchange among Asian leopard populations appears to be sufficient to counteract the tendency for admixed west Asian populations to be subsumed into the African clade.

The primary genetic division of African and Asian leopards is not reflected by their current subspecies taxonomy. Given this result, taxonomic changes could be justified under the criteria of separately evolving metapopulation lineages [26], as well as some other phylogenetic and genealogical species concepts [reviewed in e.g., 26]. However, this proposal contrasts strongly with the criteria used for species recognition in current IUCN felid taxonomy [5]. It is also not directly supported by morphological evidence [e.g., 27,28], and the 500–600 Ka divergence time of African and Asian leopards is considerably more recent than found among all other felid species [5,29]. Evidence for gene flow also argues against separate species recognition under the biological species concept [30], although genetic studies suggest interspecies admixture may have occurred frequently during the evolution of the Felidae [24,31]. It is therefore challenging to convey the broad scale pattern of genetic distinctiveness between African and Asian leopards within their current subspecies taxonomy due to a lack of intermediate taxonomic categories, as their hierarchical structure of genetic differentiation is not equally distributed across subspecies and populations.

Evolutionary history of African and Asian leopards

Our results consolidate several previous hypotheses on leopard evolution. The previously proposed African origin for leopards [2,32] is supported by the lack of whole-genome monophyly and higher genetic diversity of African leopards. Of particular interest in this regard is the Moroccan leopard, which we recovered as sister to either the Asian leopards or to all leopards in a disproportionally large number of phylogenies compared to other African leopards. Although an East African origin of leopards is suggested by the fossil record [32,33], our results could suggest that leopards in northwest Africa served as the source population for the colonisation of Asia, and potentially even as the origin for all modern leopards. A similar re-interpretation of the geographical origin from east to northwest Africa has also recently been proposed for hominids [34,35]. However, more in-depth analyses with improved sampling across north Africa and the Arabian Peninsula are required to

further test this hypothesis and place these genomic results into a more complete context.

Our results provide no evidence for multiple dispersal waves or relict Asian populations, given the high consistency of pairwise divergence dates estimated across all tested African and Asian leopard pairs (Fig 3C). Furthermore, the consensus genome phylogeny of Asian leopards reflects a series of nested east-to-west clades, suggesting the colonisation of southern Asia in a single expansion event (Figure 2B). The initial colonisation of Asia by leopards around 500-600 Ka also coincides with an important period of faunal exchange between Africa and Asia, following the last long period of higher aridity in Africa [36]. This period also included the initial out-of-Africa dispersal of several species [37], as well as increased pulses of human dispersal [38].

One unusual aspect of our results are the admixture proportions of the Middle Eastern and Central Asian individuals estimated using NGSadmix (up to 25% admixed with the African population) which contrasts with the extremely low frequency (14 out of 2,206, or 0.1%) of 1Mb sections of the genome where either or both of those individuals share more recent ancestry with an African leopard (Figure 2). Although the latter estimate increases substantially at smaller window sizes (3.6%, 8.5% and 14.8% for 500kb, 250kb and 100kb, respectively), it still fails to reach the magnitude of admixture suggested by NGSadmix. In this context, it is noteworthy that NGSadmix is a method based on single-nucleotide polymorphisms (SNPs) through genotype likelihood methods [13]. Therefore, the observed signal of admixture with African leopards must occur in relatively small segments of the genome, suggesting that the introgression of these segments took place relatively long ago and has subsequently been broken up by recombination. An alternative admixture scenario that could explain this discrepancy is admixture with a currently unsampled population that contributed SNPs shared with the African population on a genetic background of common ancestry with the Asian populations. A candidate for such a scenario is the Pleistocene European leopard, which was found to be sister to the Asian leopard clade in the mitochondrial phylogeny [10] (Figure S1). It may have possessed SNPs shared with the African population that were lost in the common ancestor of the Asian populations, but which were secondarily transferred to populations in the west of Asia through admixture. This could potentially also explain the absence of a signal of admixture in the leopards from Eritrea and Ethiopia (Figure 1C), despite their geographic proximity to Asia. Nuclear data from Pleistocene European leopards would be the best way to robustly test this hypothesis.

Our analyses also reveal substantially different population processes operating within African and Asian leopards. Asian leopards are characterised by high levels of structuring, strong isolation-by-distance and overall low heterozygosity, in contrast to African leopards where these patterns are diametrically opposed. The prominent population structure of Asian leopards could also be a relic of the initial dispersal into Asia, since population expansions can produce population structuring at neutral loci [16], although such patterns can be erased over time if levels of gene flow are high. Given the dispersal capacity of leopards and the time that has passed since their colonization of Eurasia, factors in addition to their initial population expansion may therefore also have contributed to the observed phylogeographical pattern. In particular, the much stronger structuring in Asia than in Africa may reflect differences in recent habitat loss and fragmentation, which have been much more severe in Asia (83-87% habitat reduction) compared to Africa (48-67%) [3]. However,

although relatively small, our dataset includes samples ranging from almost 200 years old to freshly collected, which all conform to the general pattern in divergence and heterozygosity we see between Africa and Asia, suggesting that these genetic patterns were present before the most recent encroachments by humans, and could thus be considered an intrinsic feature of these populations.

Conclusion

It could be expected that in well-studied groups, such as mammals, taxonomic consensus has been achieved. However, in contrast to this expectation, recent genomic studies have revealed unexpectedly strong population separations and ancient divergences that have led to the proposal of a number of new mammalian species, including red pandas [39], golden jackals [40], and orangutans [41]. Our analysis of Asian and African leopards revealed a striking pattern of almost complete genomic monophyly of Asian leopards, contrasting with a relatively recent divergence date of approximately 600 Ka and evidence for limited gene flow between the two continental groups. Thus, genome analysis may reveal subgroups that fulfil the criteria of species under some species concepts, but lack supportive evidence from other aspects of their biology. In microbiology, debate on the interplay between genomics and taxonomy is well underway [e.g., 42,43]. Given the continuing reduction in genome-sequencing costs, it seems likely that the field of vertebrate biology will face similar challenges in the near future.

Acknowledgements

This work was supported by the European Research Council (starting grant GeneFlow #310763 to M.H.). We also acknowledge sequencing support from the Swedish National Genomics Infrastructure (NGI) at the Science for Life Laboratory, which is supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation, and UPPMAX for access to computational infrastructure. L.D. and J.v.S. were funded by FORMAS (grant 2015-676). We further acknowledge Clinomics (Republic of Korea) for sequencing funds. We further would like to thank Andreas Wilting for sample contacts and manuscript feedback, André Stadler for sample contribution, Pepijn Kamminga and the Naturalis Biodiversity Center for sample contribution, and Sonja Heinrich for field dissection.

Author contributions

Conceptualisation, MH, AM, JB; Methodology, JLAP; Software, JLAP, SH; Formal Analysis, JLAP; Investigation, JLAP, RWH, KH, JvS, MHSS, GS, LD; Resources, MSB, JF, DWGF, KG, RWH, CK, ACK, EL, FM, GS, BW, LD; Discussion, JF, DWGF, RWH, ACK, EL, MVW, SJO, OU, JB, AM; Writing – Original Draft, JLAP, AB; Writing – Review and Editing, all authors have read, edited and approved the manuscript; Visualisation, JLAP; Supervision, MH; Project administration, JLAP; Funding acquisition, MH, JF, LD, JB.

Declaration of Interests

The authors declare no competing interests.

Figure legends

Figure 1: Leopard distribution and sample map, global leopard PCA & admixture. A) Map depicting all samples included in this study, numbers correspond to Table 1. Approximate species distribution is overlaid for reference (adapted from [1,2]). For this and all further figures, blue and red represent African (n=14) and Asian (n=12) samples, respectively. B) Principal Component Analysis (PCA) of genetic variation of all leopards based on 2.8 M variable positions. Axis labels include the percentage of variation explained by PC1 and PC2. The insets are scree plots showing the percentage of variation explained by each PC (PCs displayed in the main figure are shown in black). C) Admixture test based on genotype likelihood methods, using K=2. See also Figure S1 and S2.

Figure 2: Whole-genome phylogeny. A. Frequency of topology classes observed in the maximum likelihood (ML) phylogenies, calculated in 1Mb non-overlapping sliding windows along the reference genome: the left bar displays the percentage of trees that return African leopards as monophyletic in blue, and the percentage of the trees that return the Moroccan individual as sister lineage either to all leopards, or to Asian leopards. The right barplot shows the percentage of trees that return Asian leopards as monophyletic. Light grey portion of the barplot represents all topologies that occurred with a frequency of less than 3%. B. All 2,206 ML phylogenies with the Maximum Clade Credibility tree overlaid with node values indicating the clade frequency. The three-letter code indicates the subspecies each individual is assigned to based on geographical origin (following Kitchener et al. [5]): PAR = *P. p. pardus,* TUL = *P. p. tulliana,* FUS = *P. p. fusca*, KOT = *P. p. kotiya*, DEL = *P. p. delacouri*, ORI = *P. p. orientalis*, MEL = *P. p. melas.*

Figure 3: Dating out-of-Africa dispersal and gene flow using PSMC. A. Pairwise Sequential Markovian Coalescent (PSMC) of high-coverage leopard genomes (>16x, n = 8). Plotting was performed using a generation time of five years and a substitution rate of 1.1×10^{-9} per site. Ten bootstrap replicates were performed for each individual (grey). B. F1 hybrid PSMC to detect the end of geneflow between continents, calculated for high-coverage individuals. Simulated hPSMC with population coalescent times (i.e. the end of admixture) between 0 and 300 Ka are displayed in dotted lines, and the 1.5x and 10x pre-divergence population sizes in horizontal dashed lines.

Figure 4: PCA per continent, isolation-by-distance and heterozygosity. A-B: Principal Component Analysis of genetic variation of African (A, blue) and Asian (B, red) leopards, based on 2.4 M and 1.2 M variable positions after filtering, respectively. Axis labels include the percentage of variation explained by PC1 and PC2. The insets are scree plots showing the percentage of variation explained by each PC (PCs displayed in the main figure are shown in black). C: Correlation between the relative genomic distance (calculated using the identity-by-state [IBS] matrix), and the geographical distance, calculated for all combinations of African (blue) and Asian (red) samples, using random base sampling and removal of singletons to exclude errors. The significance (calculated using a Mantel test) and strength (calculated using Major Standard Axis [MSA]) are indicated for Africa and Asia. The shaded area indicates the upper and lower 2.5% jackknife confidence intervals. D: average frequency of heterozygous positions in 1 Mb windows for individuals with genome sequence coverage >10x. See also Figure S3-S5.

Table 1: Summarised details of leopard samples included in this study. See also Figure S1.

STAR Methods

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Johanna L.A. Paijmans (paijmans.jla@gmail.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Raw, unprocessed raw sequence data in fastq format have been deposited in the European Nucleotide Archive (ENA) under accession number PRJEB43565.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Historical samples were collected from collections at the Natural History Museum Berlin, Natural History Museum of Denmark (University of Copenhagen), Swedish Museum of Natural History and the National Museums Scotland. Where possible, the petrous bone or tooth cementum were sampled as these have been shown to be more likely to yield high endogenous DNA [44,45]. Alternative sampling involved other bones (turbinals or phalanges) or preserved skin. Two zoo animals, "Shinta" from Berlin Tierpark and "Bhagya" from Wuppertal Zoo, both in Germany, were sampled during routine veterinary interventions, either from a skin biopsy or blood. Bhagya was an individual wild-born in Nepal. Shinta was a captive born Javan leopard (*P. p. melas*) from a wild-born father and wild-born grandparents on the mother's side, documented in the International studbook for the Javan leopard (WAZA). Three additional samples were collected from carcasses or taken when collaring during the course of field research. All appropriate permits from the respective authorities were in place, and samples were transported with appropriate CITES permits (details available upon request).

Published sequences for two wild Amur leopards [12] were downloaded from ftp://biodisk.org/Distribute/Leopard/Rawdata/Amurleopard_resequencing1/ and ftp://biodisk.org/Distribute/Leopard/Rawdata/Amurleopard_resequencing2/. Sequence data from a Chinese leopard *P. p. japonensis* (here assigned to *P. p. orientalis*) from the Henry Doorly Zoo, Nebraska (no further provenance information available) was downloaded from the SRA (Acc Nr SRR5382750) using the SRA toolkit v2.8.1 (https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=software).

Samples were assigned to subspecies based on their geographical origin. Due to the overlapping subspecies distributions in the region, assignment of the Palestinian sample ('MFN_MAM_056095') to either *P. p. tullania* or *P. p. nimr* is not straightforward. The original publication of this specimen reports the length of the specimen (skin) to be 2.6m, which is large – especially for a female [46]. Therefore, we assign this sample to *P. p. tullania,* rather than the smaller-bodied *P. p. nimr.* The Chinese leopard ('PP28') from the SRA is assigned here to *P. p.*

orientalis rather than *P. p. japonensis*, following [5] for taxonomic consistency in this manuscript, and is thus indicated as 'Zoo ORI' in Figures and Tables.

METHOD DETAILS

Lab procedures

For historical samples, all pre-PCR steps were performed in dedicated cleanroom facilities. Extraction was performed following a protocol optimised for the retrieval of short DNA fragments [47]. This procedure in brief: tissue lysis for bone or tooth samples (25-50 mg powdered) was performed in a 1 ml reaction mix containing 0.45 M EDTA and 0.25 mg/ml Proteinase K, and for skin samples in a 1 ml reaction mix containing 5 M guanidinium thiocyanate, 25 mM NaCl, 50 mM Tris-HCl, 20 mM EDTA, 1% Tween-20, 1% beta-mercaptoethanol, incubated overnight at 37 °C with rotation. For both tissue types, centrifugation was performed to pellet remaining tissue, and the supernatant combined with a volume of 13 ml of binding buffer (5 M guanidine hydrochloride, 40% isopropanol, 0.05% Tween-20, and 90 mM sodium acetate). Purification was performed using a Zymo-Spin V Column reservoir combined with a Qiagen MinElute column. Two wash steps were performed using PE Buffer (Qiagen), followed by a drying spin for 1 minute at 13,000 rpm. DNA was eluted twice, each using 12.5 mL TET buffer (10 mM Tris-HCl, 1 mM EDTA, 0.05% Tween-20), using a 10 minute incubation time.

Illumina sequencing libraries were constructed following a double-stranded library preparation protocol for the historical samples [10,48,49]. The procedure in brief: blunt-end repair of the extracted DNA was performed in 35 µl reactions containing 1x Buffer Tango, 100 µM each dNTP, 1 mM ATP, 0.5 U/µl T4 Polynucleotide Kinase, 0.1 U/µl T4 Polymerase and 25 µl template DNA. The reaction was incubated at 25 °C for 20 minutes, followed by an inactivation phase at 72 °C for 20 minutes. Double-stranded adapters were then ligated in a 60 µl reaction containing 1x T4 DNA Ligase Buffer, 5% (w/v) PEG-4000, 0.125 U/µl T4 DNA Ligase, 0.5 µM double-stranded adapter mix. To reduce the potential of adapter dimers, the 35 µl template (blunt-end adapter mixture) was mixed with the double-stranded adapter mix prior to adding the ligase mastermix. The reaction was incubated at 22 °C for 30 minutes. The resulting product was then purified using the Qiagen MinElute kit using 2x 10 µl elution volume. Adapter fill-in was performed in a 40 µl reaction containing 1x Thermopol buffer, 250 µM each dNTP, 0.3 U/µl Bst Polymerase Large Fragment and 20 µl template. The reaction was incubated at 37 °C for 20 minutes, followed by an inactivation phase at 80 °C for 20 minutes. To determine the appropriate number of indexing PCR cycle numbers, a quantitative PCR was performed in 10 µl reactions containing 1x SYBR Green qPCR master mix, 0.2 µM each of IS7 and IS8 amplification primers, and 1 µl of a 1:20 dilution of the unamplified library. Temperature profile for the qPCR was as follows: initial denaturation at 94 °C for 10 minutes, followed by 40 cycles (denaturation phase at 94 °C for 15 seconds, annealing phase at 60 °C for 30 seconds and extension phase at 72 °C for 60 seconds). The point of inflection on the qPCR curve was used as the optimal number of cycles for the indexing PCR, corrected for different reaction volumes and template amount in the indexing PCR. Indexing PCR was performed in 80 µl reactions containing 1x AccuPrime Pfx reaction mix, 0.75 µM each of P5 and P7 indexing primers, and 0.1U/µL AccuPrime Pfx Polymerase. Indexing primers with a 8 bp unique adapter sequence nested within the P5 and P7 Illumina adapters were used for each sample [adapted from 51]. Temperature profile for the indexing PCR was as follows: initial denaturation at 94 °C for 2

minutes, followed by the selected number of cycles (denaturation phase at 95 °C for 15 seconds, annealing phase at 60 °C for 30 seconds and extension phase at 68 °C for 60 seconds) followed by a final extension phase at 68 °C for 3 minutes. Resulting libraries were purified using the Qiagen MinElute kit, and quantified on a TapeStation 2200 instrument (Agilent) with D1000 screen tape and reagents, and a Qubit 2.0 instrument (Fisher) with the dsDNA HS Assay kit.
Test sequencing to assess the endogenous content for each sample was performed on the Illumina NextSeq 500 platform. Deep sequencing for selected samples was performed on the HiSeq X at SciLifeLab Stockholm, using a 100bp paired-end strategy.
For most modern samples, DNA was extracted using the HMW MagAttract kit from Qiagen. DNA was then sheared on a Covaris S220 to an estimated size of 500bp using a Covaris S220 microTUBE (130 µl volume). Double-stranded library preparation was performed following the same protocol as for the historical samples described above, but with a higher concentration of adapter mix (2.5 µM) during adapter ligation, and a different amplification polymerase (Herculase II polymerase) for the indexing PCR [50]. Indexing PCR was performed in 80 µl reactions containing 1x Herculase Buffer, 0.75 µM each of P5 and P7 indexing primers, 0.1 mg/ml BSA, 0.25 mM each dNTP, and 0.05 U/µL Herculase II Fusion DNA polymerase. Temperature profile for the indexing PCR was as follows: initial denaturation at 94 °C for 2 minutes, followed by the selected number of cycles (denaturation phase at 94 °C for 30 seconds, annealing phase at 60 °C for 45 seconds and extension phase at 72 °C for 45 seconds) followed by a final extension phase at 72 °C for 3 minutes. Size selection of the resulting libraries was then performed on the PippinPrep, selecting for fragments between 400-900 bp using the PippinPrep Cassette 1.5% w/EtBr. For the Javan leopard the libraries were prepared on the 10X Genomics Chromium Controller instrument, using the 'Chromium Genome Reagent Kit v1' in conjunction with the 'Chromium Genome Library, Gel Bead & Multiplex Kit' and the 'Chromium Genome Chip Kit' (https://www.10xgenomics.com/resources/user-guides/ - Manual Part Number: CG00022 Rev C). Samples P8506_116_GS (Tanzania) and P8506_117_GS (Zambia B) were extracted on a Qiagen robot with the 'Symphony Tissue Extraction kit'. Library preparation and sequencing on the HiSeq 2000 was performed at the National Genomics Infrastructure (NGI).

Sequence processing

Raw sequences were trimmed using Skewer v0.2.2 [52], with default parameters and a minimum length of 30bp, and merged using Flash v1.2.11 [53], with a maximum allowed overlap of 150 bp (-M) to account for the short fragment length of historical samples. For the Javan leopard a further 22 bp were trimmed off the start of the reads to remove Chromium 10x adapter sequences using FASTX-toolkit v0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/) prior to merging. Trimmed and merged reads were aligned to the domestic cat (*Felis catus*) reference genome (v6.2) [54], in order to avoid any potential biases that can arise if there has been uneven admixture between target and reference species, which has been shown to be particularly problematic for ancient or historical datasets [e.g., 55,56]. The Burrows-Wheeler Aligner 'mem' algorithm (BWA mem) v0.7.8 and samtools v1.3.1 [57,58] were used for mapping. Supplementary alignments and unmapped reads were removed, using the -F256 and -F4 flag, respectively. Reads with low mapping quality (<Q30) were removed using samtools. Duplicate reads were removed using samtools rmdup. Detailed sequence data recovery statistics for each sample are included in Table S1. For recovery of mitochondrial genomes, reads were mapped to a leopard mitogenome sequence available from

GenBank (Acc. Nr. KP202265) [31], using the same tools and parameters as described above, except for duplication removal: to avoid over-collapsing of duplicates for high-coverage mitogenomes, duplications were marked and removed taking both mapping coordinates into consideration (MarkDupsByStartEnd.jar: https://github.com/dariober/Java-cafe/tree/master/MarkDupsByStartEnd).

QUANTIFICATION AND STATISTICAL ANALYSIS

Population structure & admixture

Population-genomic analyses were performed using the ANGSD tool v0.914 [59], considering only autosomal chromosomes. First, we assessed the coverage, quality and error estimations using -doErrorEst and -doCounts in ANGSD, using the highest coverage sample as an 'error free' individual ('Bhagya', a zoo leopard wild-born in Nepal, ~40x coverage). For all ANGSD analyses, we applied a maximum global depth filter of the 95th percentile of the global coverage to remove areas with exceptionally high coverage (e.g., duplicate regions). Data with a base quality (-MinQ) and mapping quality (-minMapQ) less than 30 were removed. Due to the high levels of deamination in some of the historical samples (checked using mapDamage v2.0.7 [60] using default parameter settings and statistical estimation disabled), transversions were removed in any analyses that included the historical samples (-rmTrans 1). Positions with missing data in at least one individual were also removed. Where possible, singletons were removed by only taking variants into account that occurred in at least two individuals (i.e., minimum SNP frequencies of 2/number of individuals).

To recover population structure, Principal Component Analyses (PCA) were performed from the genotype-likelihood data (major/minor allele) using single-base sampling, using the previously described ANGSD filters. Eigen values were calculated from the resulting covariance matrix in R v3.5.2 [61]. Admixture proportions were calculated using NGSadmix [13] assuming two ancestral populations as prior (K=2). This assumption was further tested using a range of values for K, from 2 to 5, and selecting the most likely value based on 10 replicates of each using the Clumpak server [14] and the ΔK method of Evanno et al. [15] (Figure S2). K values above K=5 were also investigated but not considered further as they produced private groups for individual samples. NGSadmix was also performed for Asian and African samples separately (Figure S3), and the most likely value of K was estimated following the method above, which also yielded K=2 as most likely value for both sets of samples.

Genome-wide phylogeny

Pseudohaploid consensus sequences were generated for each individual in ANGSD, taking a consensus-base sampling approach (-doFasta 3) and restricting the analysis to the autosomes only. Maximum likelihood trees for each 1 Mb non-overlapping sliding window along the reference genome were then calculated, specifying the domestic cat (*F. catus*) as outgroup. Windows where any individual had >50% missing data were removed. Sequence data were then converted to a binary format to excluide transition sites and RaXML v8.2.10 [62] was used to calculate the phylogeny using a BINGAMMA substitution model. A maximum clade-credibility tree was generated using TreeAnnotater v1.8.2 as included in the BEAST package v1.8.2 [63]. It should be

noted that branch lengths in the resulting trees can be affected by methodological aspects, such as differences in lab protocols, although the topology has been shown to be robust against such artefacts [64]. Recovery and counting of topology classes was performed using custom Perl scripts. We tested the effect of window size by repeating the analysis also for smaller window sizes (500Kb, 250Kb and 100Kb) for the largest chromosome (~240Mb).

PSMC & hPSMC

For the Pairwise Sequential Markovian coalescent (PSMC) only high-coverage samples could be considered, resulting in a total of eight individuals for this analysis (three African and five Asian). Heterozygous positions were recovered using samtools, filtering data for low mapping (<30) or base quality (<30). Minimum and maximum depths were set at respectively half and double the average coverage of each sample. Only data for the autosomes were considered. PSMC v0.6.5-r67 [65] was used following the same strategy as used previously for the leopard genome [12], with a mutation rate of $1.1*10^{-9}$ substitutions/site/year [17] and a generation time of 5 years. For remaining PSMC parameters, we also used parameter values as for the leopard genome [12], which followed those used for great apes [66]: maximum numbers of iterations (-N) 25, maximum $2N_0$ coalescent time (-t) 15, initial theta/rho ratio (-r) 5, and parameter pattern (-p) 4+25*2+4+6. Results were plotted in R. Ten bootstrap replicates were performed for each individual, using random re-sampling with replacement [65].

For hPSMC [18], a F1 hybrid psmc-fasta sequence was generated for each possible combination of African and Asian individuals. The pre-divergence population size was estimated by taking the lowest recovered population size from the hPSMC output, i.e., approximately 60,000. For estimating the time of the end of geneflow, simulated hPSMC data were generated using ms [67], with the time of population divergence ranging from 0 to 300 Ka in 50 Ka intervals. PSMC was then calculated using the parameters as described above. The point at which the simulated data plot overlaps the real data within the range of 1.5x – 10x population size can be taken as indicative that admixture with >0.1 individuals per generation was still occurring [18].

Isolation-by-distance and heterozygosity

We tested the correlation between geographical distance and genomic distance within African and Asian leopards. As exact GPS coordinates were not available for these samples, locations were estimated based on the geographical information for each sample, and the distance was measured in kilometres using the Geographic Distance Matrix Generator (http://biodiversityinformatics.amnh.org/open_source/gdmg/download.php). The genomic distance within each group was calculated using random base sampling and singleton removal as described above (section 'Population clustering'). A Mantel test from the 'vegan' package v2.5-6 [68] in R was performed to estimate whether there is a significant relationship between pairwise genomic and Euclidean distances. A Standard Major Axis (SMA) regression analysis from the 'lmodel2' R package [69] was then used to recover the slope (i.e., the strength of the relationship). The significance of the slope was calculated by performing jackknife replicates, following a leave-one-individual-out strategy, resulting in 14 replicates for the African leopards and 12 for the Asian leopards, and recovering the 2.5 and 97.5 percentiles (Figure 4C).

Heterozygosity was calculated for medium- and high-coverage (>10x) leopards, to avoid biases from the increased error rates of the low-coverage historical samples, leaving 15 leopards for analyses (8 Asian and 7 African). We confirmed the presence of such biases by subsampling five of our historical samples to 10x, 5x and 2x (Figure S5). Allele frequencies were calculated for each individual in ANGSD, applying mapping and base-quality filters as described above, and using ½ and 2x the average sequence coverage as minimum and maximum depth filters. realSFS was then used to calculate the frequency of heterozygous positions in 1-Mb non-overlapping windows.

Mitogenome phylogeny

Mitogenome consensus sequences were generated from the mapped data using Geneious v7.0 (www.geneious.com), using a minimum sequence depth of 3x and a strict 90% majority rule for base calling. The resulting consensus sequences for each sample were combined with 24 additional leopard mitogenome sequences available from GenBank at the time of analysis [10,31], as well as other Felidae species, using the spotted linsang (*Prionodon pardicolor* NC_024569) [70] as outgroup, resulting in an alignment of 83 sequences. This alignment contained 24 previously published leopard mitogenomes (Accession numbers MH588611 – MH588632, NC_010641, KJ866876), as well as 35 additional felid mitogenomes (Accession numbers AY463959, FCU20753, HM589214, JF357967, JF357968, JF357969, JF357970, JF357973, JF357974, KC834784, KF297576, KF776494, KF892541, KF907306, KJ508412, KJ508413, KP202267, NC_005212, NC_008450, NC_010638, NC_010642, NC_014456, NC_014770, NC_016189, NC_016470, NC_018053, NC_022842, NC_024569, NC_028299, NC_028300, NC_028303, NC_028305, NC_028306, NC_028312, NC_028316). The sequences were aligned using ClustalW with default parameters [71], as implemented in Geneious. The control region, as well as any columns in the alignment that contained missing data, were removed, resulting in a final alignment of 10,192 bp. A maximum-likelihood tree was calculated, using RaxML-HPC v8.2.4 [62] on the CIPRES black box version, with default substitution model GTRCAT, using rapid bootstrapping and search for the best-scoring ML tree, selecting the spotted linsang as outgroup (*P. p. pardicolor*), on the CIPRES Science Gateway [72] (Figure S1).

References

1. Miththapala, S., Seidensticker, J., and O'Brien, S.J. (1996). Phylogeographic Subspecies Recognition in Leopards (*Panthera pardus*): Molecular Genetic Variation. Conservation Biology *10*, 1115–1132.

2. Uphyrkina, O., Johnson, W.E., Quigley, H., Miquelle, D., Marker, L., Bush, M., and O'Brien, S.J. (2001). Phylogenetics, genome diversity and origin of modern leopard, *Panthera pardus*. Molecular Ecology *10*, 2617–2633.

3. Jacobson, A.P., Gerngross, P., Lemeris Jr., J.R., Schoonover, R.F., Anco, C., Breitenmoser-Würsten, C., Durant, S.M., Farhadinia, M.S., Henschel, P., Kamler, J.F., *et al.* (2016). Leopard (*Panthera pardus*) status, distribution, and the research efforts across its range. PeerJ *4*, e1974.

4. Diedrich, C.G. (2013). Late Pleistocene leopards across Europe – northernmost European German population, highest elevated records in the Swiss Alps, complete skeletons in the Bosnia Herzegowina Dinarids and comparison to the Ice Age cave art. Quaternary Science Reviews *76*, 167–193.

5. Kitchener, A.C., Breitenmoser-Würsten, C., Eizirik, E., Gentry, A., Werdelin, L., Wilting, A., Yamaguchi, N., Abramov, A.V., Christiansen, P., Driscoll, C., *et al.* (2017). A revised taxonomy of the Felidae : The final report of the Cat Classification Task Force of the IUCN Cat Specialist Group. Available at: http://repository.si.edu/xmlui/handle/10088/32616.

6. Uphyrkina, O., Miquelle, D., Quigley, H., Driscoll, C., and O'Brien, S.J. (2002). Conservation Genetics of the Far Eastern Leopard (*Panthera pardus orientalis*). J Hered *93*, 303–311.

7. Dou, H., Feng, L., Xiao, W., and Wang, T. (2016). The complete mitochondrial genome of the North Chinese Leopard (*Panthera pardus japonensis*). Mitochondrial DNA *27*, 1167–1168.

8. Wilting, A., Patel, R., Pfestorf, H., Kern, C., Sultan, K., Ario, A., Peñaloza, F., Kramer-Schadt, S., Radchuk, V., Foerster, D.W., *et al.* (2016). Evolutionary history and conservation significance of the Javan leopard *Panthera pardus melas*. J Zool *299*, 239–250.

9. Ebrahimi, A., Farashi, A., and Rashki, A. (2017). Habitat suitability of Persian leopard *Panthera pardus saxicolor* in Iran in future. Environ Earth Sci *76*, 697.

10. Paijmans, J.L.A., Barlow, A., Förster, D.W., Henneberger, K., Meyer, M., Nickel, B., Nagel, D., Worsøe Havmøller, R., Baryshnikov, G.F., Joger, U., *et al.* (2018). Historical biogeography of the leopard (*Panthera pardus*) and its extinct Eurasian populations. BMC Evolutionary Biology *18*, 156.

11. Sommer, R.S., and Benecke, N. (2006). Late Pleistocene and Holocene development of the felid fauna (Felidae) of Europe: a review. Journal of Zoology *269*, 7–19.

12. Kim, S., Cho, Y.S., Kim, H.-M., Chung, O., Kim, H., Jho, S., Seomun, H., Kim, J., Bang, W.Y., Kim, C., *et al.* (2016). Comparison of carnivore, omnivore, and herbivore mammalian genomes with a new leopard assembly. Genome Biology *17*, 211.

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

14. Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., and Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Molecular Ecology Resources *15*, 1179–1191.

15. Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology *14*, 2611–2620.

16. Excoffier, L., and Ray, N. (2008). Surfing during population expansions promotes genetic revolutions and structuration. Trends in Ecology & Evolution *23*, 347–351.

17. Cho, Y.S., Hu, L., Hou, H., Lee, H., Xu, J., Kwon, S., Oh, S., Kim, H.-M., Jho, S., Kim, S., *et al.* (2013). The tiger genome and comparative analysis with lion and snow leopard genomes. Nature Communications *4*, 2433.

18. Cahill, J.A., Soares, A.E.R., Green, R.E., and Shapiro, B. (2016). Inferring species divergence times using pairwise sequential Markovian coalescent modelling and low-coverage genomic data. Phil. Trans. R. Soc. B *371*, 20150138.

19. Sugimoto, T., Aramilev, V.V., Kerley, L.L., Nagata, J., Miquelle, D.G., and McCullough, D.R. (2014). Noninvasive genetic analyses for estimating population size and genetic diversity of the remaining Far Eastern leopard (*Panthera pardus orientalis*) population. Conserv Genet *15*, 521–532.

20. Jackson, P., and Nowell, K. (2016). *Panthera pardus ssp. orientalis*. IUCN Red List of Threatened Species. Available at: https://www.iucnredlist.org/en.

21. Seidensticker, J., and Lumpkin, S. (1991). Great Cats (Emmaus, Pa. : New York: Rodale Pr).

22. Fattebert, J., Dickerson, T., Balme, G., Slotow, R., and Hunter, L. (2013). Long-Distance Natal Dispersal in Leopard Reveals Potential for a Three-Country Metapopulation. South African Journal of Wildlife Research *43*, 61–67.

23. Barlow, A., Cahill, J.A., Hartmann, S., Theunert, C., Xenikoudakis, G., Fortes, G.G., Paijmans, J.L.A., Rabeder, G., Frischauf, C., Grandal-d'Anglade, A., *et al.* (2018). Partial genomic survival of cave bears in living brown bears. Nature Ecology & Evolution *2*, 1563.

24. Figueiró, H.V., Li, G., Trindade, F.J., Assis, J., Pais, F., Fernandes, G., Santos, S.H.D., Hughes, G.M., Komissarov, A., Antunes, A., *et al.* (2017). Genome-wide signatures of complex introgression and adaptive evolution in the big cats. Science Advances *3*, e1700299.

25. Hudson, R.R., and Coyne, J.A. (2002). Mathematical Consequences of the Genealogical Species Concept. Evolution *56*, 1557–1565.

26. De Queiroz, K. (2007). Species Concepts and Species Delimitation. Syst Biol *56*, 879–886.

27. Pocock, R.I. (1930). The Panthers and Ounces of Asia. Journal of the Bombay Natural History Society *34*, 64–82.

28. Meijaard, E. (2004). Biogeographic History of the Javan Leopard Panthera pardus Based on a Craniometric Analysis. Journal of Mammalogy *85*, 302–310.

29. Johnson, W.E., Eizirik, E., Pecon-Slattery, J., Murphy, W.J., Antunes, A., Teeling, E., and

O'Brien, S.J. (2006). The Late Miocene Radiation of Modern Felidae: A Genetic Assessment. Science *311*, 73–77.

30. O'brien, S.J., and Mayr, E. (1991). Bureaucratic Mischief: Recognizing Endangered Species and Subspecies. Science *251*, 1187–1188.

31. Li, G., Davis, B.W., Eizirik, E., and Murphy, W.J. (2016). Phylogenomic evidence for ancient hybridization in the genomes of living cats (Felidae). Genome Research *26*, 1–11.

32. Werdelin, L., Yamaguchi, N., Johnson, W.E., and O'Brien, S.J. (2010). Phylogeny and evolution of cats (Felidae). Biology and conservation of wild felids. Oxford, 59–82.

33. Turner, A., and Antón, M. (1997). The Big Cats and Their Fossil Relatives: An Illustrated Guide to Their Evolution and Natural History (Columbia University Press).

34. Richter, D., Grün, R., Joannes-Boyau, R., Steele, T.E., Amani, F., Rué, M., Fernandes, P., Raynal, J.-P., Geraads, D., Ben-Ncer, A., *et al.* (2017). The age of the hominin fossils from Jebel Irhoud, Morocco, and the origins of the Middle Stone Age. Nature *546*, 293–296.

35. Hublin, J.-J., Ben-Ncer, A., Bailey, S.E., Freidline, S.E., Neubauer, S., Skinner, M.M., Bergmann, I., Le Cabec, A., Benazzi, S., Harvati, K., *et al.* (2017). New fossils from Jebel Irhoud, Morocco and the pan-African origin of *Homo sapiens*. Nature *546*, 289–292.

36. deMenocal, P.B. (2004). African climate change and faunal evolution during the Pliocene–Pleistocene. Earth and Planetary Science Letters *220*, 3–24.

37. Martínez-Navarro, B. (2010). Early Pleistocene Faunas of Eurasia and Hominin Dispersals. In Out of Africa I: The First Hominin Colonization of Eurasia Vertebrate Paleobiology and Paleoanthropology., J. G. Fleagle, J. J. Shea, F. E. Grine, A. L. Baden, and R. E. Leakey, eds. (Dordrecht: Springer Netherlands), pp. 207–224. Available at: https://doi.org/10.1007/978-90-481-9036-2_13.

38. Abbate, E., and Sagri, M. (2012). Early to Middle Pleistocene Homo dispersals from Africa to Eurasia: Geological, climatic and environmental constraints. Quaternary International *267*, 3–19.

39. Hu, Y., Wu, Q., Ma, S., Ma, T., Shan, L., Wang, X., Nie, Y., Ning, Z., Yan, L., Xiu, Y., *et al.* (2017). Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. PNAS *114*, 1081–1086.

40. Koepfli, K.-P., Pollinger, J., Godinho, R., Robinson, J., Lea, A., Hendricks, S., Schweizer, R.M., Thalmann, O., Silva, P., Fan, Z., *et al.* (2015). Genome-wide Evidence Reveals that African and Eurasian Golden Jackals Are Distinct Species. Current Biology *25*, 2158–2165.

41. Nater, A., Mattle-Greminger, M.P., Nurcahyo, A., Nowak, M.G., de Manuel, M., Desai, T., Groves, C., Pybus, M., Sonay, T.B., Roos, C., *et al.* (2017). Morphometric, Behavioral, and Genomic Evidence for a New Orangutan Species. Current Biology *27*, 3487-3498.e10.

42. Hugenholtz, P., Skarshewski, A., and Parks, D.H. (2016). Genome-Based Microbial Taxonomy Coming of Age. Cold Spring Harb Perspect Biol *8*, a018085.

43. Mahato, N.K., Gupta, V., Singh, P., Kumari, R., Verma, H., Tripathi, C., Rani, P., Sharma,

A., Singhvi, N., Sood, U., *et al.* (2017). Microbial taxonomy in the era of OMICS: application of DNA sequences, computational tools and techniques. Antonie van Leeuwenhoek *110*, 1357–1371.

44. Damgaard, P.B., Margaryan, A., Schroeder, H., Orlando, L., Willerslev, E., and Allentoft, M.E. (2015). Improving access to endogenous DNA in ancient bones and teeth. Sci Rep *5*, 1–12.

45. Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., Gerritsen, F., Moiseyev, V., Gromov, A., Raczky, P., *et al.* (2015). Optimal Ancient DNA Yields from the Inner Ear Part of the Human Petrous Bone. PLoS ONE *10*, e0129102.

46. Kumerloeve, H. (1972). Zum Stand des Vorkommens von Panthera pardus tulliana Valenciennes 1856 in Kleinasien. Mit Hinweisen auf Syrien und andere Nachbargebiete. Zoologische Gart, 4–22. Available at: https://eurekamag.com/research/024/157/024157281.php.

47. Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., *et al.* (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. PNAS *110*, 15758–15763.

48. Meyer, M., and Kircher, M. (2010). Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. Cold Spring Harbor Protocols *2010*, doi:10.1101/pdb.prot5448.

49. Henneberger, K., Barlow, A., and Paijmans, J.L.A. (2019). Double-Stranded Library Preparation for Ancient and Other Degraded Samples. In Ancient DNA, B. Shapiro, A. Barlow, P. D. Heintzman, M. Hofreiter, J. L. A. Paijmans, and A. E. R. Soares, eds. (New York, NY: Springer New York), pp. 65–73. Available at: http://link.springer.com/10.1007/978-1-4939-9176-1_8.

50. Dabney, J., and Meyer, M. (2012). Length and GC-biases during sequencing library amplification: A comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. BioTechniques *52*, 87–94.

51. Gansauge, M.-T., and Meyer, M. (2013). Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. Nat. Protocols *8*, 737–748.

52. Jiang, H., Lei, R., Ding, S.-W., and Zhu, S. (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics *15*, 182.

53. Magoč, T., and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics *27*, 2957–2963.

54. Pontius, J.U., Mullikin, J.C., Smith, D.R., Agencourt Sequencing Team, Lindblad-Toh, K., Gnerre, S., Clamp, M., Chang, J., Stephens, R., Neelam, B., *et al.* (2007). Initial sequence and comparative analysis of the cat genome. Genome Res. *17*, 1675–1689.

55. Sheng, G.-L., Basler, N., Ji, X.-P., Paijmans, J.L.A., Alberti, F., Preick, M., Hartmann, S., Westbury, M.V., Yuan, J.-X., Jablonski, N.G., *et al.* (2019). Paleogenome Reveals Genetic Contribution of Extinct Giant Panda to Extant Populations. Current Biology *29*, 1695-1700.e6.

56. Günther, T., and Nettelblad, C. (2018). The presence and impact of reference bias on

population genomic studies of prehistoric human populations. bioRxiv, 487983.

57. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

58. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

59. Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics *15*, 356.

60. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., and Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics *29*, 1682–1684.

61. R Core Team (2013). R: A language and environment for statistical computing Available at: http://www.R-project.org/.

62. Stamatakis, A. (2014). RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. Bioinformatics, btu033.

63. Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., and Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol *4*, vey016.

64. Barlow, A., Hartmann, S., Gonzalez, J., Hofreiter, M., and Paijmans, J.L.A. (2020). Consensify: A Method for Generating Pseudohaploid Genome Sequences from Palaeogenomic Datasets with Reduced Error Rates. Genes *11*, 50.

65. Li, H., and Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. Nature *475*, 493–496.

66. Prado-Martinez, J., Sudmant, P.H., Kidd, J.M., Li, H., Kelley, J.L., Lorente-Galdos, B., Veeramah, K.R., Woerner, A.E., O'Connor, T.D., Santpere, G., *et al.* (2013). Great ape genetic diversity and population history. Nature *499*, 471–475.

67. Hudson, R.R. (2002). Generating samples under a Wright–Fisher neutral model of genetic variation. Bioinformatics *18*, 337–338.

68. Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M.H.H., Oksanen, M.J., and Suggests, M. (2007). The vegan package. Community ecology package *10*, 719.

69. Legendre, P. (2018). lmodel2 package | R Documentation. Available at: https://www.rdocumentation.org/packages/lmodel2/versions/1.7-3.

70. Hassanin, A., and Veron, G. (2016). The complete mitochondrial genome of the Spotted Linsang, *Prionodon pardicolor*, the first representative from the family Prionodontidae (Mammalia, Carnivora). Mitochondrial DNA Part A *27*, 912–913.

71. Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., *et al.* (2007). Clustal W and Clustal X

version 2.0. Bioinformatics *23*, 2947–2948.

72. Miller, M.A., Pfeiffer, W., and Schwartz, T. (2011). The CIPRES Science Gateway: A Community Resource for Phylogenetic Analyses. In Proceedings of the 2011 TeraGrid Conference: Extreme Digital Discovery TG '11. (New York, NY, USA: ACM), p. 41:1-41:8. Available at: http://doi.acm.org/10.1145/2016741.2016785.