Check for updates

RESEARCH ARTICLE

REVISED A genomic exploration of the early evolution of extant cats and their sabre-toothed relatives

[version 2; peer review: 2 approved]

Michael V Westbury[1], Ross Barnett[1], Marcela Sandoval-Velasco[1], Graham Gower[1], Filipe Garrett Vieira[1], Marc de Manuel[2], Anders J Hansen[1], Nobuyuki Yamaguchi[3], Lars Werdelin[4], Tomas Marques-Bonet[2,5-7], M Thomas P Gilbert[1,8*], Eline D Lorenzen[1*]

[1]The GLOBE Institute, University of Copenhagen, Øster Voldgade 5-7, Copenhagen, Denmark
[2]Institute of Evolutionary Biology (UPF-CSIC), PRBB, Dr. Aiguader 88, 08003 Barcelona, Spain
[3]Institute of Tropical Biodiversity and Sustainable Development, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia
[4]Department of Palaeobiology, Swedish Museum of Natural History, Box 50007, SE-104 05 Stockholm, Sweden
[5]Catalan Institution of Research and Advanced Studies (ICREA), Passeig de Lluís Companys, 23, 08010, Barcelona, Spain
[6]CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Baldiri i Reixac 4, 08028 Barcelona, Spain
[7]Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Edifici ICTA-ICP, c/ Columnes s/n, 08193 Cerdanyola del Vallès, Barcelona, Spain
[8]Department of Natural History, NTNU University Museum, Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim, Norway

* Equal contributors

v2 **First published:** 24 Mar 2021, **1**:25
https://doi.org/10.12688/openreseurope.13104.1
Latest published: 21 Jun 2021, **1**:25
https://doi.org/10.12688/openreseurope.13104.2

Abstract

Background: The evolutionary relationships of Felidae during their Early–Middle Miocene radiation is contentious. Although the early common ancestors have been subsumed under the grade-group *Pseudaelurus,* this group is thought to be paraphyletic, including the early ancestors of both modern cats and extinct sabretooths.
Methods: Here, we sequenced a draft nuclear genome of *Smilodon populator,* dated to 13,182 ± 90 cal BP, making this the oldest palaeogenome from South America to date, a region known to be problematic for ancient DNA preservation. We analysed this genome, together with genomes from other extinct and extant cats to investigate their phylogenetic relationships.
Results: We confirm a deep divergence (~20.65 Ma) within sabre-toothed cats. Through the analysis of both simulated and empirical data, we show a lack of gene flow between *Smilodon* and contemporary Felidae.
Conclusions: Given that some species traditionally assigned to *Pseudaelurus* originated in the Early Miocene ~20 Ma, this indicates

Open Peer Review

Approval Status ✓ ✓

	1	2
version 2 (revision) 21 Jun 2021		
version 1 24 Mar 2021	✓ view	✓ view

1. **Daniel Förster**, Leibniz Institute for Zoo and Wildlife Research (IZW), Berlin, Germany

2. **Manuel Salesa**, Museo Nacional de Ciencias Naturales-CSIC, Madrid, Spain

Any reports and responses or comments on the article can be found at the end of the article.

that some species of *Pseudaelurus* may be younger than the lineages they purportedly gave rise to, further supporting the hypothesis that *Pseudaelurus* was paraphyletic.

Keywords
Smilodon, ancient DNA, genomics, gene flow, Felidae, palaeogenome, phylogeny

This article is included in the Horizon 2020 gateway.

This article is included in the Evolution and Ecology gateway.

This article is included in the Genetics and Genomics gateway.

This article is included in the European Research Council (ERC) gateway.

This article is included in the Evolutionary Biology collection.

Corresponding author: Michael V Westbury (m.westbury@sund.ku.dk)

Author roles: Westbury MV: Conceptualization, Formal Analysis, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Barnett R**: Conceptualization, Investigation; **Sandoval-Velasco M**: Investigation; **Gower G**: Formal Analysis, Software; **Vieira FG**: Data Curation, Formal Analysis; **de Manuel M**: Formal Analysis; **Hansen AJ**: Funding Acquisition; **Yamaguchi N**: Writing – Review & Editing; **Werdelin L**: Writing – Review & Editing; **Marques-Bonet T**: Supervision; **Gilbert MTP**: Conceptualization, Funding Acquisition, Project Administration, Supervision, Writing – Review & Editing; **Lorenzen ED**: Funding Acquisition, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No [681396]), (project Extinction Genomics). This project has also received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. FP7-PEOPLE-2011-IEF-298820. The work was also supported by the Independent Research Fund Denmark | Natural Sciences, Forskningsprojekt 1, grant no. 8021-00218B to EDL.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2021 Westbury MV *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Westbury MV, Barnett R, Sandoval-Velasco M *et al.* **A genomic exploration of the early evolution of extant cats and their sabre-toothed relatives [version 2; peer review: 2 approved]** Open Research Europe 2021, **1**:25 https://doi.org/10.12688/openreseurope.13104.2

First published: 24 Mar 2021, **1**:25 https://doi.org/10.12688/openreseurope.13104.1

REVISED **Amendments from Version 1**

We have made some minor updates to the manuscript based on the reviewers' suggestions. This includes fixing some typos and adding two new tables to the main text from the extended data.

Any further responses from the reviewers can be found at the end of the article

Plain language summary

Here we sequenced the genome of the extinct sabre-toothed cat *Smilodon populator*. By comparing this genome to those of living cat species as well as another extinct sabre-toothed cat *Homotherium latidens,* we were able to assess their evolutionary relationships to one another. We show that not only are sabre-toothed and extant cats very divergent from one another, but there were also highly divergent species within sabre-toothed cats. This high level of divergence within sabre-toothed cats makes it difficult to place the common ancestor of all cat species based on just the fossil record. Moreover, we were able to show that it was very unlikely that sabre-toothed cats hybridised with any living cat species.

Introduction

Based on the fossil record, the origin of Felidae, colloquially referred to as cats, is well resolved to the Middle-Late Oligocene (~30–27 million years ago (Ma))[1]. However, the subsequent radiation of Felidae in the Early–Middle Miocene (~23–15 Ma) is less well understood, encompassing the evolution and radiation from the common ancestor of a number of species subsumed under the paraphyletic grade-group *Pseudaelurus*, which was widely distributed across Europe, Asia, and North America[1]. The putative paraphyly of *Pseudaelurus* is exemplified by the fact that it likely includes not only the early ancestors of modern cats (Felinae) but also those of the extinct sabretooths (Machairodontinae).

In contrast to the early radiation of Felidae, the phylogenetic relationships among extant cats are relatively well understood; extant cats are believed to share a most recent common ancestor in the Late Miocene (~11 Ma)[2]. However, despite this, there is uncertainty surrounding the constituents of the long-stem lineage of Felinae after its split from Machairodontinae (which extends through the Early and Middle Miocene) until its radiation in the Late Miocene. The evolutionary relationships within Machairodontinae are even less well understood, and increased knowledge of the evolutionary relationships both between Felinae and Machairodontinae, as well as within Machairodontinae itself, would provide important insights to help resolve the complex early radiation of Felidae.

The last surviving members of the Machairodontinae belonged to the genus *Smilodon* (tribe Smilodontini). While once widespread across the continents of North America (*S. fatalis*) and South America (*S. populator*), the genus went extinct ~10 thousand years ago (kya)[3,4]. *Smilodon* are not known further north than southernmost Canada (~42 °N), and their range extended until the tip of the South American continent (~53 °S). Apart from a few anomalous mass occurrences in tar pits (e.g. Rancho La Brea, USA and Talara, Peru), *Smilodon* is relatively rare in the fossil record, which is not uncommon for an apex carnivore. Moreover, when it is present, there are only limited remains[5].

The evolutionary relationships of *Smilodon* to other extinct and extant large cats are not resolved. An early ancient DNA study using both mitochondrial and nuclear DNA placed *Smilodon* within Felinae[6]. However, this was later shown to likely reflect contaminant DNA from a domestic cat[7]. A later study using complete mitochondrial genomes found *Smilodon* to be highly divergent from both Felinae (with an estimated divergence of ~20 Ma), and another extinct Machairodontinae lineage (*Homotherium* (tribe Homotheriini)), diverging ~18 Ma[8]. However, conclusions based exclusively on a single maternally inherited locus can be biased by interspecific hybridisation and incomplete lineage sorting, which has been well-documented in living cats[9].

To date, only three studies present authentic DNA from *Smilodon*[8,10,11], all of which is mitochondrial DNA. Although there is a large collection from Rancho La Brea, retrieving endogenous DNA from the material is not considered feasible[7], greatly restricting the number of specimens available for DNA study. The relative absence of studies of *Smilodon* likely in part reflects their rarity in the fossil record outside of tar pits. In addition, as *Smilodon* are assumed to have been mixed wood-edge ambush predators[12], the majority of fossils are found in open-air sites with poor DNA preservation; cave sites (which afford better preservation) are rare. Finally, the detrimental effects of temperature at equatorial and near-equatorial latitudes further exacerbate DNA preservation.

To elucidate the evolutionary relationship of *Smilodon* to extant and extinct Felidae, and to gain insights into the early radiation of Felidae, we sequenced a draft nuclear genome of a single *Smilodon populator* individual from the Ultima Esperanza region of Chile. The specimen was radiocarbon dated to 13,182 ± 90 calibrated years before present (cal BP), and represents the oldest palaeogenome from South America, a region in which ancient DNA preservation is expected to be limited[13].

Results and discussion

Using an African lion assembly as the reference genome, we successfully mapped a draft nuclear genome of a single *Smilodon populator* individual to an average genome-wide coverage of ~0.7x, with an average read depth of ~2x. We achieved this using a combination of both target enrichment and shotgun sequencing. We performed target enrichment using baits created from either DNA extracted from a modern lion (whole-genome capture), or based on the exome annotations of the lion reference genome (exome capture). The discrepancy between genome-wide coverage and average read-depth likely reflects the use of captured data, and lack of a closely related reference genome. If we were to have a conspecific reference genome, we would expect a more even genome-wide coverage, more comparable to the read depth.

We sequenced approximately 800 million reads from four independently constructed libraries (Table 1). Investigations into mapped reads showed high levels of fragmentation (average read length <90bp for all libraries) and high rates of C-T transitions across the reads (Figure S1, *Extended data*[14]), both typical of authentic ancient DNA. Comparisons among libraries showed that the capture experiments did yield significant levels of endogenous DNA, but were not highly different from the shotgun data but had lower levels of complexity. Moreover, although the exome capture was able to increase the relative coverage of the exome, it also included a lot of non-coding whole genomic data (Table 1).

To investigate the topological placement of *Smilodon*, we computed a neighbour-joining (NJ) tree using genome-wide transversion pairwise distances, rooted using *Crocuta crocuta* (spotted hyena). Similar to the mitochondrial genome, we found *Smilodon* and *Homotherium* to be sister taxa, and Machairodontinae to be sister taxon to all extant cats (Figure S2, *Extended data*[14]). Support for this topology was high; we repeated this analysis independently for the 100 longest scaffolds (27.2Mb - 5.3Mb), and found identical placement of *Smilodon* for each scaffold.

We further tested the closer affinity of *Smilodon* and *Homotherium* to each other, relative to any living Felinae, using D-statistics[15] to test for topology. For this, we took advantage of the topological input requirement for the D-statistics test ([[[H1, H2], H3], Outgroup]). We computed the D-statistic with *Smilodon* and *Homotherium* as sister taxa (i.e. in the H1 and H2 positions) (Table S1, *Extended data*[14]), each species of Felinae as H3, and the spotted hyena as Outgroup, and compared this to the D-statistic and significance from 0 (Z-score) when placing *Smilodon* and *Homotherium* paraphyletically (i.e. in the H1 and H3 positions) (Table S2, *Extended data*[14]). We clearly see a much higher D-statistic (D= 0.54-0.67, Z= 286.6-442.5) when placing the sabre-toothed cats parapyletically, compared to when they are placed as sister taxa (D= 0.12-0.20, Z= 39.2-67.7). A high D-statistic could be interpreted as gene flow between either H1 and H3, or between H2 and H3. However, another possible explanation could be the incorrect input topology. Although paraphyly of the sabre-toothed cats followed by gene flow could explain the observed pattern, a more likely explanation for the higher D-statistic when placing the sabre-tooths paraphyletically, would be the monophyly of the sabre-tooths as seen in our NJ tree. This result further supports *Smilodon* and *Homotherium* as more closely related to one another, than either are to any extant cat species.

To further contextualise this relationship, we built a dated phylogenetic tree (Figure 1). However, due to the low coverage of our *Smilodon* genome, traditional methods for phylogeny dating are likely unsuitable. Therefore, we devised a method to overcome this by computing pairwise genetic drift distances using F-statistics (in our case F2) (Table S3, *Extended data*[14])[16]. The pairwise F2 statistics were built into an unrooted NJ tree (Figure S3, *Extended data*[14]). The relationships recovered in this tree were the same as those obtained using the transversional genetic distances.

To calibrate our phylogenetic tree, we estimated the average rate of genetic drift between all pairs of genera within Felinae based on the F2 results and previously calculated divergence dates[17]. We made a number of assumptions about our data, including (i) a strict molecular clock, (ii) constant population sizes through time, and (iii) that Felinae drift rates are similar to those in Machairodontinae.

Using the F2 statistics and average drift rate, we estimated a divergence date between Machairodontinae and Felinae of

Table 1. *Smilodon* mapping results. WGC - whole genome capture, SE - single end, PE - paired end.

Library name	BU19_5	BU23_1	Smi745_1	Smi745_2
Library type	Exome capture	WGC	WGC/Shotgun mix	Shotgun
# SE reads	145,712,209	7,856,010	125,060,258	64,850,185
# PE reads	164,381,221	309,399,989	-	-
# PE reads merged	146,380,424	278,659,340	-	-
# Retained reads	313,124,333	342,949,720	121,988,548	60,427,680
# Mapped reads	6,070,092	6,474,904	5,191,598	7,136,656
Proportion duplicates	0.893	0.830	0.731	0.100
Endogenous proportion	0.019	0.019	0.043	0.118
Exome coverage	1.237	0.217	0.605	0.272
Genome coverage	0.171	0.234	0.139	0.179
Average read length (bp)	68.7	88.3	65.4	61.3

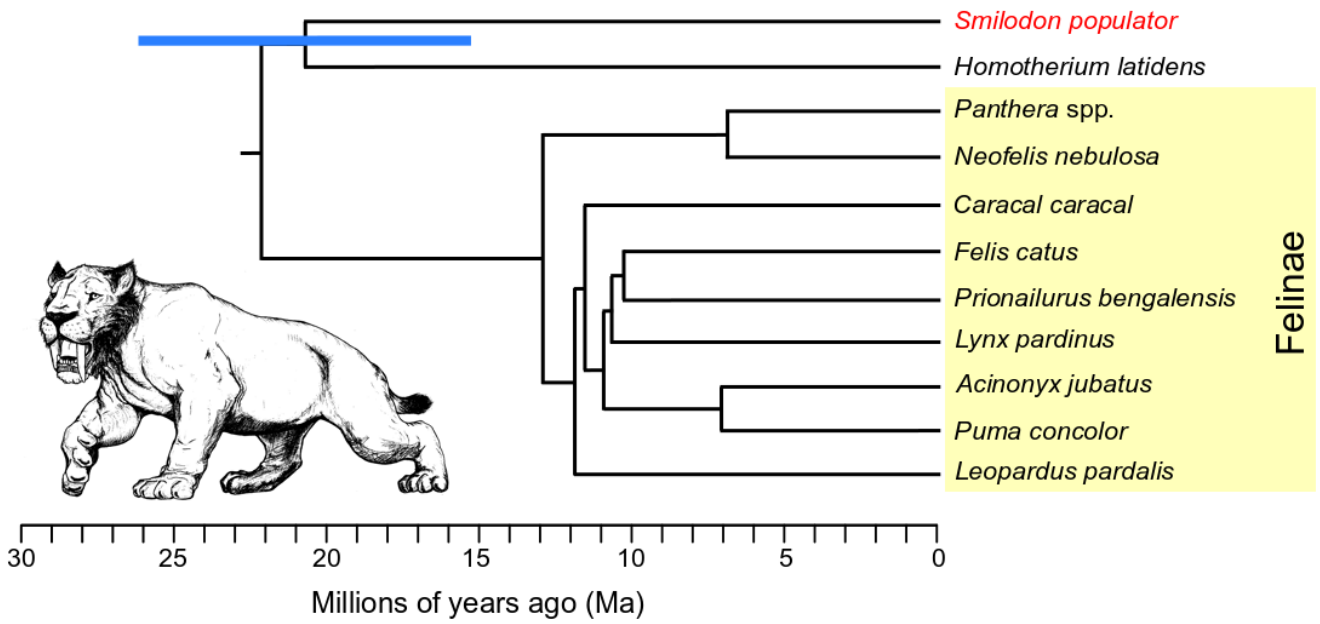


Figure 1. Dated Felidae phylogenetic tree based on genome-wide pairwise F2 statistics, calibrated using an average genetic drift rate estimated from within Felinae (yellow shading). Blue bar shows 95% confidence interval for the divergence between the *Smilodon* (red) *Homotherium* lineages, calculated from the F2 standard error. *Smilodon* illustration by Binia De Cahsan and included with permission.

~22.1 Ma (Figure 1), similar to the estimate of ~22.5 Ma calculated using a high-coverage exome dataset[17]. Although the correlation between these two results was not unexpected, as our tree was calibrated using the average within Felinae drift calculated using the divergence times of Barnett *et al.*[17], their similar divergence estimates suggest extrapolating Felinae drift rates to other subfamilies of Felidae is valid.

We further tested for the robustness of this method by both recalculating the divergence times within Felinae, and by downsampling two Felinae species (*Caracal caracal* - caracal and *Prionailurus bengalensis* - leopard cat) to ~0.7x. We recovered similar divergence estimates to those previously reported, and to those produced without downsampling, providing confidence in the use of this methodology for low-coverage genomes (Tables S3 and S4, *Extended data*[14]). Using this methodology, we estimated the divergence date of *Smilodon*/*Homotherium* to be ~20.65 Ma (95% CI 26.07-15.25 Ma) (Figure 1; Table S5, *Extended data*[14]). Furthermore, although based on low-coverage data and a number of data assumptions, the congruence of our results with previous mitogenome-based estimates from the same individuals[8], despite the use of different data and calibration methods, further adds confidence to our method of phylogenetic assessment.

Species traditionally assigned to *Pseudaelurus* originated in the Early Miocene, ~20 Ma. Given the deep divergence between Machairodontinae and Felinae, as well as within Machairodontinae, this indicates that some species of *Pseudaelurus* are younger than the lineages they purportedly gave rise to. This is strong support for the hypothesis that *Pseudaelurus* as a grade-group is paraphyletic. Furthermore, current phylogenies using well-known characters (such as length of upper canines and presence of serrations or crenulations) resolve the Smilodontini-Homotheriini divergence to only ~11-10 Ma, while older machairodonts constitute a stem lineage of uncertain relationship[18]. Therefore, given the problematic nature of machairodont phylogenetics, the new molecular information analysed here provides an important reference point for identifying morphologically Smilodontini or Homotheriini characters in specimens from >11 Ma, to help resolve machairodont phylogeny back to the early Miocene.

We next investigated the evolutionary relationships between *Smilodon* and Felinae by searching for signatures of gene flow, using two independent methods (f3[16] and D3[19]). The results were assessed using simulated data (Tables S6, S7, and S8, *Extended data*[14]). As lineages leading to *Smilodon* and *Homotherium* diverged relatively soon after the Machairodontinae/Felinae split, we were able to test for ancient gene flow (up to 20 Ma) between *Smilodon* or *Homotheirum* and stem Felinae. Ancient gene flow may have prevented the diverging lineages in early Felidae from accumulating obvious morphological differences, thus preventing the confident phylogenetic placement of lineages during this time period.

Moreover, we tested for more recent gene flow events between either *Smilodon* or *Homotherium* and the entire *Panthera* big cat lineage (all *Panthera* species grouped together), due to their potential size overlap, as well as with single species that

may have more recently met *Smilodon* due to spatial proximity (i.e. from South America: *Panthera onca* (jaguar), *Puma concolor* (puma), and *Leopardus pardalis* (ocelot)).

We did not find any indication of gene flow between any of the lineages tested, regardless of method (Tables S6 and S7, *Extended data*[14]). The lack of more recent gene flow events is not surprising, due to the relatively ancient divergence of Machairodontinae and Felinae. However, the lack of ancient gene flow signatures do not necessarily mean that gene flow was not present during the early divergence of these lineages. Rather, it may result from the inadequacy of the methods in uncovering such events.

We assessed the power of the D3 statistic to detect gene flow, which occurred at different times (20 Ma-50 kya), using simulated data. When using a simple demographic model of constant population size, mutation rate, and recombination rate on error-free simulated data, we detect significant levels of gene flow back to ~16 Ma (Table S8, *Extended data*[14]). However, empirical data do not always behave in such a simple manner, and a variety of factors may influence the results of the D3 statistic. These include, but are not restricted to: violations of the infinite sites model, different mutation rates across lineages, ancestral populations structure, and introgression with unsampled lineages[19]. However, our results suggest that D3 may be suitable for highly divergent lineages with ancient gene flow events and therefore violations of the infinite sites model may not be as problematic.

Thus, although we are unable to exclude the possibility of ancient gene flow during the early radiation of Felidae, we are somewhat more confident that there was no more recent gene flow between either *Smilodon* or *Homotherium* and Felinae within the last 16 Ma years. If very ancient gene flow had occurred prior to 16 Ma, it is likely that recombination and genetic drift would have either highly fragmented the introgressed regions, or completely removed them from contemporary Felinae individuals. This would render the regions too small to be detected with current methods, which use genome-wide summary statistics or regions of phylogenetic incongruence with known evolutionary relationships, to infer gene flow.

Our study exemplifies how even a draft palaeogenome from an extinct species can provide important information into their evolutionary history. Through the sequencing of a single *Smilodon populator* genome, we provide insights into Felidae's early radiation in the Early–Middle Miocene (~23 - 15 Ma), which could not be uncovered using genetic data from extant species alone.

Methods

Sample information

Specimen ZMA20.042 is from the Naturalis museum, Leiden, the Netherlands, and was radiocarbon dated to 11,335±30 uncalibrated years before present (2-sigma range of 13,269-13,095 calibrated years before present) (Stafford: UCIAMS-142836). We calibrated the radiocarbon date using Calib v7.04[20] using the int13.14c calibration curve. It has been identified as a left tibia of *Smilodon populator* and is part of the Kruimel collection, an assortment of megafaunal remains recovered from the Última Esperanza region of Chile, most likely from the site of Cueva del Milodón. This and other specimens of *Smilodon* from the Kruimel Collection were described and figured by McDonald and Werdelin (2018); specimen ZMA20.042 is presented in Figure 4.6D[5]. We additionally included genomic information from another extinct sabre-toothed cat, 12 extant cat species, and the spotted hyena (Table 2).

Ancient DNA extraction, library preparation, and sequencing

Samples of cortical bone were taken from the long bone element (approx. 1 cm^3) using a Dremel drill, and reduced to powder in a Mikrodismembrator. Two independent DNA extractions were performed as described in Orlando *et al.*[21] in a dedicated ancient DNA laboratory, with negative controls. We built each DNA extract and negative control into genomic libraries using the NEB E6070 kit following a modified version of the protocol used by Vilstrup *et al.*[22]. Briefly, the extract (30 µl) was end-repaired and cleaned using a MinElute column, the collected flow-through was adapter-ligated and cleaned using a QiaQuick column, and the adapter fill-in reaction was performed on the flowthrough. To complete the library build, we performed a final incubation at 37°C (30 min) followed by inactivation overnight at -20°C.

Table 2. List of the additional species included in this study. Original sources and accession codes for all raw reads used in this study can be found in Table S9 (see *Extended data*[14]).

Species name:	Common name:
Felis catus	Domestic cat
Prionailurus bengalensis	Leopard cat
Lynx pardinus	Iberian lynx
Acinonyx jubatus	Cheetah
Caracal caracal	Caracal
Neofilis nebulosa	Clouded leopard
Panthera uncia	Snow leopard
Panthera onca	Jaguar
Panthera pardus	Leopard
Panthera leo	Lion
Crocuta crocuta	Spotted hyena
Homotherium latidens	Scimitar-toothed cat
Puma concolor	Puma
Leopardus pardalis	Ocelot

For each library, we performed two independent indexing PCR amplifications (Veriti thermal cycler, Applied Biosystems) in a 50 µl volume reaction, using 25 µl of library, 25 µl PCR master mix, and 12 cycles of PCR reactions. The final concentrations in the PCR master mix were 1.25 U AccuPrime™ Pfx DNA Polymerase (Invitrogen, Cat # 12344-024), 1x AccuPrime™ Pfx reaction mix (Invitrogen, Cat # 12344-024), 0.4 mg/ml BSA, 120 nM primer InPE, 120 nM of a multiplexing indexing primer containing a unique six-nucleotide index code (Illumina – sequences TGCAGG, CGATGA, GCGAGA, or CAGCAC). PCR cycling conditions consisted of an initial denaturation step at 95°C for 2 min, followed by 12 cycles of 95°C denaturation for 15 s, 60°C annealing for 30 s, and 68°C extension for 30 s, and a final extension step at 68°C for 7 min. Indexed libraries were checked for presence of DNA on a 2% agarose gel before purification using the QIAquick column system (Qiagen, Cat # 28104) and quantification on an Agilent 2100 BioAnalyzer. Quantified libraries were communally pooled in equimolar ratios and sequenced as 100 bp single-end reads on an Illumina HiSeq2000 platform at the Danish National High-Throughput Sequencing Centre and 100 bp paired-end reads on an Illumina Hiseq2000 at BGI Copenhagen.

Genome capture

We assessed the shotgun-sequenced libraries for endogenous content and selected the libraries with the highest levels of endogenous DNA for two sets of capture experiments. The first set used biotinylated RNA probes transcribed from fresh DNA extracted from a modern lion, for the purpose of enriching the entire nuclear genome (whole-genome capture). The second method used biotinylated RNA baits, assembled based on the exonic annotations of lion genomic data[23] (exome capture). Both types of baits were generated by Arbor Biosciences (Ann Arbor, MI, USA) and carried out using the myBaits target enrichment kit and the instructions described in manual V3.

After capture and cleanup, enriched libraries were re-amplified for further sequencing using either a Phusion polymerase (New England Biolabs, Cat # M0530S) or a KAPA HiFi HotStart polymerase (Roche, Cat # KK2801 07959052001) with primers IS5_reamp.P5 and IS6_reamp.P7 over 14 cycles[24]. We quantified the resultant enriched libraries on a TapeStation 2200 instrument and sequenced them on an Illumina Hiseq2000 at the Danish National High-throughput Sequencing Centre.

Data processing pipeline

Post-sequencing read processing of the *Smilodon populator* was performed using the PALEOMIX v1.2.5 pipeline[25]. Adaptor sequence removal and trimming of low-quality bases (BaseQ < 5 or Ns) was done with AdapterRemoval v2.0.0[26]. This step also removed all reads shorter than 30 bp in length or with more than 50 bp of missing data. Trimmed reads were mapped against an African lion reference genome (NCBI Accession JAAVKH000000000[27]) with BWA-MEM v0.7.5a[28], utilising default parameters. PCR duplicates were identified and filtered based on the 5'-end mapping coordinate using Picard v2.18.0. *GATK* v3.8.0[29] was used to perform an *indel* realignment step to adjust for increased error rates at the end of short reads in the presence of *indels*. In the absence of a curated dataset of *indels*, this step relied on a set of *indels* identified in the specific sample being processed. Read damage patterns were assessed and base quality scores recalibrated around read terminal damage patterns using mapDamage v2.0.5[30,31].

Data processing of the extant Felidae species, excluding *Puma concolor* and *Leopardus pardalis*, followed the same pipeline with the following minor adjustments; no minimum read length cut-off or missing data cut-off was applied during the adapter trimming step, and bases were not recalibrated using mapDamage. These steps were removed as the data from the extant species would not display the highly fragmented and damage patterns found in ancient DNA. The *Puma concolor* and *Leopardus pardalis* samples had Illumina adapter and short sequences trimmed using skewer 0.2.2[32] but followed the same protocol as the other extant species for the rest of the processing steps.

Neighbour-joining tree

To build a NJ phylogenetic tree, we computed an identity by state distance matrix considering only transversion differences using ANGSD v0.921[33], and specifying the following parameters; call the consensus base (-doIBS 2), minimum mapping and quality of 30 (-minmapq 30, -minq 30), only include a site if all individuals are covered (-minInd), remove secondary alignments (-remove_bads 1), only include reads that map to a single location (-uniqueonly 1), compute major and minor alleles based on genotype likelihoods (-domajorminor 1), remove transitions (-rmtrans 1), print all sites (-minminor 0), and use the GATK algorithm to compute genotype likelihoods (-GL 2) and only include scaffolds over 1Mb in length (-rf). After filtering, 221,350,529 sites remained.

We further checked for support of the genome-wide topology by building a distance matrix for the 100 longest scaffolds independently, resulting in 100 independent distance matrices based on 171,281 - 3,344,878 sites. The resultant distance matrices were converted into NJ trees using fastME v2.1.5[34].

D-statistics topology test

To investigate the closer relationship of *Smilodon* and *Homotherium* to each other relative to other extant cat species, we implemented a D-statistics test for topology[35]. Although D-statistics is most commonly used to find evidence of gene flow, it can also be used to test for phylogenetic relationships. This test takes advantage of the fact that D-statistics relies on a predefined four-taxon topology as input [[[H1,H2],H3],outgroup]. A high D-score is most commonly used to infer post-divergence gene flow, but it can also be caused by more recent common ancestry brought about by an incorrect predefined topology.

Taking the latter into account, we performed a number of D-statistics tests including the topologies: [[[*Smilodon*,

Homotherium], extant cat], *Crocuta crocuta*], [[[*Smilodon*, extant cat], *Homotherium*], *Crocuta crocuta*], and [[[*Homotherium*, extant cat], *Smilodon*], *Crocuta crocuta*]. In this test, 'extant cat' was replaced with each living Felidae species included in this study. We used ANGSD v0.921[33] to perform the D-statistics with the following parameters; -minmapq 30 -minq 30, -minind 15, -remove_bads 1, -uniqueonly 1, -domajorminor 1, -rmtrans 1, -GL 2, calculate D in a block size of 1Mb (-blocksize), and the spotted hyena as the ancestral/outgroup sequence (-anc).

Dated phylogeny

We computed the consensus haploid base calls (-dohaplocall 2) for all scaffolds greater than 1MB in length using ANGSD, specifying the following parameters; minimum mapping and quality of 30 (-minmapq 30, -minq 30), only include a site if all individuals are covered (minInd), remove secondary alignments (-remove_bads 1), only include reads that map to a single location (-uniqueonly 1), compute major and minor alleles based on genotype likelihoods (-domajorminor 1), remove transitions (-rmtrans 1), print all sites (-minminor 0), and use the GATK algorithm to compute genotype likelihoods (-GL 2). After filtering, 268,152,250 sites remained.

We converted the haploid call file into a PLINK file format using the haplo2plink command from the ANGSD toolsuite. Using the resultant PLINK file, we calculated F2 statistics[15] for each pairwise combination of our *Smilodon populator* individual, *Homotherium latidens*, and 12 Felinae species using the popstats.py script[36]. We used genetic drift calculated via F-statistics as opposed to absolute pairwise distance as it should be more suitable for ancient DNA data[37]. From these pairwise F2 comparisons, we built a distance matrix that was converted into a newick tree file using PHYLIP v3.696 neighbor[38] for visualisation.

From the distance matrix, we computed the Felinae average rate of genetic drift (F2) to be 0.000305 per million years. For this, we calculated the average F2 between pairwise comparisons of all genera within Felinae that had divergence time estimates available in Barnett *et al.*[17]. These included *Acinonyx*, *Caracal*, *Felis*, *Lynx*, *Neofelis*, *Panthera*, and *Prionailurus*, and resulted in 21 comparisons (Table S3, *Extended data*[14]). We used the formula of (F2 × 0.5)/divergence time to estimate the rate of genetic drift per million years for all 21 comparisons. We calculated the mean of these 21 comparisons, giving us the Felinae mean rate of genetic drift per million years. This mean rate was used in conjunction with the previously calculated average F2 to calibrate the tree including the *Smilodon populator*, *Homotherium latidens,* puma, and ocelot.

To test for the robustness of our method to low-coverage data, we downsampled both *Caracal* and *Prionailurus* to comparable coverage to *Smilodon populator* (~0.7x) using SAMtools v1.6[39], and recomputed the F2 statistics. We used the same average rate of genetic drift to estimate the divergence of these genera from their closest relatives (i.e. *Prionailurus* - *Felis*, and *Caracal* - *Acinonyx*/*Felis*/*Lynx*). For this analysis we assumed a known species tree and divergences within Felinae previously found in Barnett *et al.*[17], a genome-wide constant mutation/drift rate, and no variation in drift rates between lineages. Uncertainty around the divergence between *Smilodon* and *Homotherium* was calculated using 1.96 x the standard deviation on either side of the F2 statistic. The standard deviation was calculated using the standard error x $\sqrt{N}$ (where N = number of blocks using for jackknifing).

Assessing gene flow

We implemented two independent analyses to test for the presence of signs of gene flow between Machairodontinae and Felinae. We computed F3 statistics[16] to assess whether there were any signals of gene flow between a predefined triplet [[A,B],C] using the same PLINK file computed above for the F2 statistics. We used popstats.py[36] for five independent triplet combinations, which we present as ((A,B),C) in the subsequent text. We placed *Smilodon* and *Homotherium* in the A and B positions, while alternating C.

First, we investigated signs of very ancient gene flow between Machairodontinae and Felinae by specifying all extant Felinae as one population. Next, we looked for gene flow with the *Panthera* spp. big cats by specifying all *Panthera* as a single population. Finally, we computed F3 three times independently to test for signs of very recent gene flow between Machairodontinae and either of three extant cat species occupying South America, which may have come into contact with *Smilodon* based on geography (*Panthera onca* (jaguar), *Puma concolor* (puma), and *Leopardus pardalis* (ocelot)).

To complement this analysis, we also computed D3 statistics[18] on the same five triplet comparisons. For this, we also produced a haplocall file in ANGSD using the same filtering parameters as above, but using a random base call (-dohaplocall 1), as opposed to the consensus base call, while specifying the same additional parameters specified above. The resultant haploid output was then converted into a geno file and run through the popgenWindows.py python script. We ran the popgenWindows.py script using default parameters, specifying a window size of 1MB, and the minimum number of sites per window as 1kb. From this output, we calculated D3 for each window independently by applying the equation [[BC-AC]/[BC+AC]]. Mean values, standard deviations, and significance from 0 were measured in R v3.6.0 using the pnorm function[40].

Evaluating D3 for detecting ancient gene flow

To evaluate the adequacy of the D3 method for detecting very ancient gene flow events (up to 20 Ma), we simulated three diploid sequences representing *Homotherium* (A), *Smilodon* (B), and a Felinae species (C), using msprime[41]. As input, we estimated the average transversion mutation rate within Felidae using the average pairwise distance between *Homotherium* and Felinae and divided this by two before multiplying it by the Machairodontinae/Felinae divergence time calculated above (22.1 Ma). We computed pairwise comparisons between *Homotherium* and each Felinae species included in the study using ANGSD and took the average pairwise distance from

these comparisons. We excluded the *Smilodon* from this calculation due to the lower quality of the genome.

To calculate the pairwise distances we used the consensus identity by state parameter (-doIBS 2) in ANGSD and applied the following filters; -minmapq 30, -minq 30, -minind 15, -unique-only 1, -remove_bads 1, -domajorminor 1, -rmtrans 1, -minminor 0, -makematrix 1, and only included scaffolds >1MB in length. This gave us an average pairwise distance of 0.01207, which we converted into an average transversion mutation rate of 2.730769e-10 per year. We converted this to a generational mutation rate using a generation time calculated for lions of 6.5 years[42]; if the generation time of the investigated species differed from that of lion, the mutation rate would adjust accordingly, and we therefore expect minimal impact of this parameter on the final results, especially as we ran the simulations specifying years as opposed to number of generations. For the recombination rate, we used a previously published sex-averaged recombination rate of 1.9 cM/Mb (1.9e-8)[43].

Using the above information, we ran five independent simulations, each consisting of 2,000 1 Mb windows with a constant effective population size of 40,000 individuals, a generation time of 6.5 years, the above-mentioned mutation and recombination rates, a 5% pulse of migration (m=0.05), and a different timing of the migration pulse for each of the five runs (20 Ma, 18 Ma, 17 Ma, 16 Ma, 15 Ma, 10 Ma, 5 Ma, 50 kya). The scripts for these simulation runs can be found on GitHub. We calculated the D3 statistic from this output using the tskit toolkit[44] using the d3.py script. Significance was calculated as it was done for the empirical data above.

Data availability

Underlying data

NCBI BioProject: Raw sequencing reads for the *Smilodon* individual. Accession number PRJNA691254; https://www.ncbi.nlm.nih.gov/bioproject/691254.

Extended data

Zenodo: A genomic exploration of the early evolution of extant cats and their sabre-toothed relatives - extended data. https://doi.org/10.5281/zenodo.4922450[14].

This project contains the following extended data:

- Westbury et al extended data.pdf (Supplementary Figures S1-S3 and Tables S1-S9)

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Acknowledgements

We thank the laboratory technicians of the Centre for GeoGenetics and the staff of the Danish National High-Throughput DNA Sequencing Centre for technical assistance. We thank Tom Stafford Jr and Stafford Research LLC for radiocarbon dating and discussion. We thank Jean-Marie Rouillard (Arbor Biosciences) for help with bait design. We would like to thank Binia De Cahsan for the illustration in Figure 1. Lastly, we would like to thank Reinier van Zelst and Caroline Pepermans Naturalis, Leiden, the Netherlands, for access to *Smilodon populator* samples.

References

1. Werdelin L, Yamaguchi N, Johnson WE, *et al.*: **Phylogeny and evolution of cats (Felidae).** *Biology and conservation of wild felids*. Oxford, 2010; 59–82.
Reference Source
2. Johnson WE, Eizirik E, Pecon-Slattery J, *et al.*: **The late Miocene radiation of modern Felidae: a genetic assessment.** *Science.* 2006; **311**(5757): 73–77.
PubMed Abstract | Publisher Full Text
3. Werdelin L, McDonald HG, Shaw CA: **Smilodon: The Iconic Sabertooth (JHU Press)**. 2018.
Reference Source
4. Kurtén B, Werdelin L: **Relationships between North and South American *Smilodon*.** *J Vert Paleontol.* 1990; **10**(2): 158–169.
Publisher Full Text
5. McDonald HG, Werdelin L: **The sabertooth cat *Smilodon populator* (Carnivora: Felidae), from Cueva del Milodón, Chile.** In *Smilodon: The Iconic Sabertooth*, L. Werdelin, H. G. McDonald, and C. A. and Shaw, eds. (Baltimore: John Hopkins University Press), 2018; 53–75.
Reference Source
6. Janczewski DN, Yuhki N, Gilbert DA, *et al.*: **Molecular phylogenetic inference from saber-toothed cat fossils of Rancho La Brea.** *Proc Natl Acad Sci U S A.* 1992; **89**(20): 9769–9773.
PubMed Abstract | Publisher Full Text | Free Full Text
7. Gold DA, Robinson J, Farrell AB, *et al.*: **Attempted DNA extraction from a Rancho La Brea Columbian mammoth (*Mammuthus columbi*): prospects for ancient DNA from asphalt deposits.** *Ecol Evol.* 2014; **4**(4): 329–336.
PubMed Abstract | Publisher Full Text | Free Full Text
8. Paijmans JLA, Barnett R, Gilbert MTP, *et al.*: **Evolutionary History of Saber-Toothed Cats Based on Ancient Mitogenomics.** *Curr Biol.* 2017; **27**(21): 3330–3336.e5.
PubMed Abstract | Publisher Full Text
9. Li G, Figueiró HV, Eizirik E, *et al.*: **Recombination-Aware Phylogenomics Reveals the Structured Genomic Landscape of Hybridizing Cat Species.** *Mol Biol Evol.* 2019; **36**(10): 2111–2126.
PubMed Abstract | Publisher Full Text | Free Full Text
10. Metcalf JL, Turney C, Barnett R, *et al.*: **Synergistic roles of climate warming and human occupation in Patagonian megafaunal extinctions during the Last Deglaciation.** *Sci Adv.* 2016; **2**(6): e1501682.
PubMed Abstract | Publisher Full Text | Free Full Text
11. Barnett R, Barnes I, Phillips MJ, *et al.*: **Evolution of the extinct Sabretooths and the American cheetah-like cat.** *Curr Biol.* 2005; **15**(15): R589–90.
PubMed Abstract | Publisher Full Text
12. Bocherens H, Cotte M, Bonini R, *et al.*: **Paleobiology of sabretooth cat *Smilodon populator* in the Pampean Region (Buenos Aires Province, Argentina) around the Last Glacial Maximum: Insights from carbon and nitrogen stable isotopes in bone collagen.** *Palaeogeogr Palaeoclimatol Palaeoecol.* 2016; **449**: 463–474.
Publisher Full Text
13. Hofreiter M, Paijmans JLA, Goodchild H, *et al.*: **The future of ancient DNA: Technical advances and conceptual shifts.** *Bioessays.* 2015; **37**(3): 284–293.
PubMed Abstract | Publisher Full Text
14. Westbury M, Barnett R, Sandoval-Velasco M, *et al.*: **A genomic exploration of the early evolution of extant cats and their sabre-toothed relatives - extended data.** [Data set]. *Zenodo.* 2021.
http://www.doi.org/10.5281/zenodo.4434076
15. Green RE, Krause J, Briggs AW, *et al.*: **A draft sequence of the Neandertal**

genome. *Science.* 2010; **328**(5979): 710–722.
PubMed Abstract | Publisher Full Text | Free Full Text

16. Reich D, Thangaraj K, Patterson N, *et al.*: **Reconstructing Indian population history.** *Nature.* 2009; **461**(7263): 489–494.
PubMed Abstract | Publisher Full Text | Free Full Text

17. Barnett R, Westbury MV, Sandoval-Velasco M, *et al.*: **Genomic adaptations and evolutionary history of the extinct scimitar-toothed cat, *Homotherium latidens*.** *Curr Biol.* 2020; **30**(24): 5018–5025.e5.
PubMed Abstract | Publisher Full Text | Free Full Text

18. Werdelin L, Flink T: **The phylogenetic context of *Smilodon*.** In *Smilodon: The Iconic Sabertooth*, L. Werdelin, H. G. McDonald, and C. A. and Shaw, eds. (Baltimore: John Hopkins University Press), 2018; 14–29.
Reference Source

19. Hahn MW, Hibbins MS: **A Three-Sample Test for Introgression.** *Mol Biol Evol.* 2019; **36**(12): 2878–2882.
PubMed Abstract | Publisher Full Text

20. Reimer PJ, Bard E, Bayliss A, *et al.*: **IntCal13 and Marine13 Radiocarbon Age Calibration Curves 0–50,000 Years cal BP.** *Radiocarbon.* 2013; **55**(4): 1869–1887.
Publisher Full Text

21. Orlando L, Ginolhac A, Zhang G, *et al.*: **Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse.** *Nature.* 2013; **499**(7456): 74–78.
PubMed Abstract | Publisher Full Text

22. Vilstrup JT, Seguin-Orlando A, Stiller M, *et al.*: **Mitochondrial phylogenomics of modern and ancient equids.** *PLoS One.* 2013; **8**(2): e55950.
PubMed Abstract | Publisher Full Text | Free Full Text

23. Cho YS, Hu L, Hou H, *et al.*: **The tiger genome and comparative analysis with lion and snow leopard genomes.** *Nat Commun.* 2013; **4**: 2433.
PubMed Abstract | Publisher Full Text | Free Full Text

24. Meyer M, Kircher M: **Illumina sequencing library preparation for highly multiplexed target capture and sequencing.** *Cold Spring Harb Protoc.* 2010; **2010**(6): pdb.prot5448.
PubMed Abstract | Publisher Full Text

25. Schubert M, Ermini L, Der Sarkissian C, *et al.*: **Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX.** *Nat Protoc.* 2014; **9**(5): 1056–1082.
PubMed Abstract | Publisher Full Text

26. Schubert M, Lindgreen S, Orlando L: **AdapterRemoval v2: rapid adapter trimming, identification, and read merging.** *BMC Res Notes.* 2016; **9**: 88.
PubMed Abstract | Publisher Full Text | Free Full Text

27. de Manuel M, Barnett R, Sandoval-Velasco M, *et al.*: **The evolutionary history of extinct and living lions.** *Proc Natl Acad Sci U S A.* 2020; **117**(20): 10927–10934.
PubMed Abstract | Publisher Full Text | Free Full Text

28. Li H, Durbin R: **Fast and accurate short read alignment with Burrows-Wheeler transform.** *Bioinformatics.* 2009; **25**(14): 1754–1760.
PubMed Abstract | Publisher Full Text | Free Full Text

29. McKenna A, Hanna M, Banks E, *et al.*: **The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data.** *Genome Res.* 2010; **20**(9): 1297–1303.
PubMed Abstract | Publisher Full Text | Free Full Text

30. Ginolhac A, Rasmussen M, Gilbert MTP, *et al.*: **mapDamage: testing for damage patterns in ancient DNA sequences.** *Bioinformatics.* 2011; **27**(15): 2153–2155.
PubMed Abstract | Publisher Full Text

31. Jónsson H, Ginolhac A, Schubert M, *et al.*: **mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters.** *Bioinformatics.* 2013; **29**(13): 1682–1684.
PubMed Abstract | Publisher Full Text | Free Full Text

32. Jiang H, Lei R, Ding SW, *et al.*: **Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads.** *BMC Bioinformatics.* 2014; **15**: 182.
PubMed Abstract | Publisher Full Text | Free Full Text

33. Korneliussen TS, Albrechtsen A, Nielsen R: **ANGSD: Analysis of Next Generation Sequencing Data.** *BMC Bioinformatics.* 2014; **15**(1): 356.
PubMed Abstract | Publisher Full Text | Free Full Text

34. Lefort V, Desper R, Gascuel O: **FastME 2.0: A Comprehensive, Accurate, and Fast Distance-Based Phylogeny Inference Program.** *Mol Biol Evol.* 2015; **32**(10): 2798–2800.
PubMed Abstract | Publisher Full Text | Free Full Text

35. Westbury MV, Hartmann S, Barlow A, *et al.*: **Extended and Continuous Decline in Effective Population Size Results in Low Genomic Diversity in the World's Rarest Hyena Species, the Brown Hyena.** *Mol Biol Evol.* 2018; **35**(5): 1225–1237.
PubMed Abstract | Publisher Full Text | Free Full Text

36. Skoglund P, Mallick S, Bortolini MC, *et al.*: **Genetic evidence for two founding populations of the Americas.** *Nature.* 2015; **525**(7567): 104–108.
PubMed Abstract | Publisher Full Text | Free Full Text

37. Harris AM, DeGiorgio M: **Admixture and Ancestry Inference from Ancient and Modern Samples through Measures of Population Genetic Drift.** *Hum Biol.* 2017; **89**(1): 21–46.
PubMed Abstract | Publisher Full Text

38. Felsenstein J: **PHYLIP (Phylogeny Inference Package) version 3.6**. 2005.
Reference Source

39. Li H, Handsaker B, Wysoker A, *et al.*: **The Sequence Alignment/Map format and SAMtools.** *Bioinformatics.* 2009; **25**(16): 2078–2079.
PubMed Abstract | Publisher Full Text | Free Full Text

40. R Core Team: **R: A language and environment for statistical computing**. 2013.
Reference Source

41. Kelleher J, Etheridge AM, McVean G: **Efficient Coalescent Simulation and Genealogical Analysis for Large Sample Sizes.** *PLoS Comput Biol.* 2016; **12**(5): e1004842.
PubMed Abstract | Publisher Full Text | Free Full Text

42. Black SA, Fellous A, Yamaguchi N, *et al.*: **Examining the extinction of the Barbary lion and its implications for felid conservation.** *PLoS One.* 2013; **8**(4): e60174.
PubMed Abstract | Publisher Full Text | Free Full Text

43. Li G, Hillier LW, Grahn RA, *et al.*: **A High-Resolution SNP Array-Based Linkage Map Anchors a New Domestic Cat Draft Genome Assembly and Provides Detailed Patterns of Recombination.** *G3 (Bethesda).* 2016; **6**(6): 1607–1616.
PubMed Abstract | Publisher Full Text | Free Full Text

44. Ralph P, Thornton K, Kelleher J: **Efficiently Summarizing Relationships in Large Samples: A General Duality Between Statistics of Genealogies and Genomes.** *Genetics.* 2020; **215**(3): 779–797.
PubMed Abstract | Publisher Full Text | Free Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 1

Reviewer Report 19 April 2021

https://doi.org/10.21956/openreseurope.14174.r26653

© **2021 Salesa M.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Manuel Salesa**
[1] Departamento de Paleobiología, Museo Nacional de Ciencias Naturales-CSIC, Madrid, Spain
[2] Departamento de Paleobiología, Museo Nacional de Ciencias Naturales-CSIC, Madrid, Spain

In the present work, the authors show the results of the sequencing of a draft nuclear genome of the largest sabre-toothed felid, *Smilodon populator*, with an age of around 13,000 years ago, making this the oldest palaeogenome from South America to date, a region known to be problematic for ancient DNA preservation. The analysis of this genome, together with genomes from other extinct and extant cats, allows the authors to suggest a series of very interesting phylogenetic proposals. Among them, they confirm a deep divergence (around 20.65 Ma) within Machairodontinae, something that was evident from the fossil record, but never stated from palaeogenomics.

I haven't found major concerns about this work. The methodology seems impeccable, the figure is informative, the taxonomic citations are correct, and the discussion is well supported by the data. The style of the text is clear, and the manuscript is correctly written. I just detected a grammatical mistake when citing *S. populator* on page 3, which is cited as "*popular*". Finally, I find the subject treated in the manuscript suitable for its indexing on *Open Research Europe.*

MINOR COMMENTS

Although fossils of Early and Middle Miocene felids are known since the XIX century, our understanding of their phylogenetic relationships has not changed greatly over this time. This was in part due to the fact that those fossils of primitive felids, mostly fragmentary dental and postcranial material, show a very similar morphology to that of extant species, and only sabre-toothed taxa exhibit evident differences in their elongated and flattened upper canines or their huge carnassial teeth. In the present paper, Westbury *et al*. have sequenced the genome of one of the most iconic sabre-toothed felids, the giant *Smilodon populator*, and then they compared it to those of extant species of Felidae, as well as to that of another extinct sabre-toothed species, the relatively slender and cursorial *Homotherium latidens*. The results were expectable but very interesting: they not only show the strong divergence between Machairodontinae (sabre-toothed

felids) and Felinae (extant felids), but also between the two analysed species of sabre-toothed felids, *S. populator* and *H. latidens*. These two derived forms were the last sabre-toothed felids that were present in the faunas, and both were extinct as recently as around 28,000–11,000 years. *Smilodon* and *Homotherium* illustrate two very different ways to be a sabre-toothed felid: *S. populator*, included in a lineage known as Smilodontini, was a strong-built animal, with short and powerful limbs and very elongated upper canines, whereas *H. latidens* was a more slender predator, with relatively long limbs (the hind limbs much shorter than the forelimbs) and shorter canines. Thus, the divergence found by Westbury *et al*. fits well with the idea that the sabre-toothed felids developed at least two different morphotypes very early in their evolution, a robust, forest-dweller type illustrated by the Middle Miocene *Pseudaelurus quadridentatus* and the Late Miocene *Promegantereon ogygia*, and a slenderer model, probably occupying more open environments, whose first well-known representative is the Late Miocene *Machairodus aphanistus*. These two lineages evolved in parallel, refining each of the morphotypes, occupying different habitats thanks to different locomotor adaptations, thus allowing the presence in the faunas of two large predators through habitat segregation.

Concerning the early divergence between sabre-toothed felids and felines, the present work by Westbury *et al*. points towards a date of around 22.1–22.5 Ma. This is interesting, as we have little data on the felids of that moment, but all the available fossils suggest that the split between felines and macairodontines occurred at that time. There are several European localities with fossils slightly younger than that age (around 18 Ma) that can be identified as early felines, as their dentitions exhibit short canines of a rounded section, although they also show some very primitive features such as the presence of m2; and besides these forms, there are other fossils that show an incipient development of the macairodontine model illustrated by the laterally flattened upper canines of *Pseudaelurus quadridentatus*. Now seems clear that this latter genus should be considered as the first known sabre-toothed felid, as it has laterally flattened upper canines, a relatively robust first metacarpal, and an overall postcranial anatomy very similar to that of the younger *Promegantereon ogygia* (Beaumont, 1975[1]; Salesa *et al*., 2010[2]). Nevertheless, as the present work by Westbury *et al*. remarks, the genus *Pseudaelurus* has traditionally included felids with evident differences in their dental morphology, both from North America and Eurasia; the paraphyly of this genus was already suggested by Werdelin *et al.* (2010)[3] but this new work reinforces this idea.

And finally, it is very interesting that, although Westbury *et al.* remark that they were unable to exclude the possibility of ancient gene flow during the early radiation of Felidae, they are somewhat more confident that there was no more recent gene flow between either *Smilodon* or *Homotherium* and Felinae within the last 16 Ma. Felinae and Machairodontinae developed a body plan devoted to preying upon different kinds of prey, and their dental and skeletal morphology suggests an early independent evolution of each model. This absence of gene flow also suggests the existence of several ethological and physiological differences between felines and sabre-toothed felids whose trace unfortunately vanished when the latter disappeared.

References

1. de Beaumont G: Recherches sur les Félidés (Mammifères, Carnivores) du Pliocène inférieur des sables à Dinotherium des environs d'Eppelsheim (Rheinhessen). *Archives des Sciences Physiques et Naturelles, Genève*. **28**: 369-405
2. SALESA M, ANTÓN M, TURNER A, ALCALÁ L, et al.: Systematic revision of the Late Miocene sabre-toothed felid Paramachaerodus in Spain. *Palaeontology*. 2010; **53** (6): 1369-1391 Publisher Full Text

3. Werdelin L, Yamaguchi N, Johnson WE, O'Brien SJ: Phylogeny and evolution of cats (Felidae). *Biology and Conservation of Wild Felids - Oxford University Press*. 2010. 59-82

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Palaeobiology, Systematics, Functional anatomy

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 10 Jun 2021

Michael Westbury

We would like to thank Manuel Salesa for praising our work and giving a very interesting summary of the results within the context of prior literature. We believe this additional information published alongside our manuscript to be helpful to the reader to better understand the background of the study. We have also corrected the mistake when citing S. populator on page 3, which was cited as "popular".

Competing Interests: No competing interests were disclosed.

Reviewer Report 08 April 2021

https://doi.org/10.21956/openreseurope.14174.r26645

© 2021 Förster D. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Daniel Förster
[1] Evolutionary Genetics Department, Leibniz Institute for Zoo and Wildlife Research (IZW), Berlin, Germany
[2] Evolutionary Genetics Department, Leibniz Institute for Zoo and Wildlife Research (IZW), Berlin, Germany

Westbury *et al.* use nuclear data of a *Smilodon populator* individual to resolve the phylogenetic relationship of this species with another extinct sabre-toothed cat (*Homotherium latidens*), and of the Machairodontinae (here, Smilodon and Homotherium) with the extant cats (Felinae). They illuminate the deeper divergences within this Carnivoran family, i.e. the Felidae.

The study confirms that *Smilodon* and *Homotherium* are sister lineages, previously suggested by mtDNA data, and, as the authors point out, these results have implications for resolving the phylogeny of the machairodonts based on morphological characters. Gene flow between lineages during the Early Miocene radiation of the Felidae was not detected, within machairodonts or between machairodonts and stem Felinae, potentially owing to available data and analytical approaches that were possible. An absence of gene flow between machairodonts and extant cats in the more recent past is not too surprising.

An interesting study illuminating the history of a charismatic group. Unfortunately, perhaps owing to genome coverage, the authors did not examine genomic adaptations as Barnett *et al.* 2020 did for *Homotherium latidens*.

Tables & Figures
It seems unnecessary to avoid any Tables in the main text. For example, having the table of species analysed (Table S10) in the Methods section would benefit clarity, as these data suddenly find mention in the data processing section, as well as being featured in multiple subsequent sections.

The mapping results (Table S1) are summarized in the Results & Discussion section, but the table could feature in the main text as well.

Plain language summary
“assess” rather than “access”?

Introduction
The first sentence reads rather clunky.

Results & Discussion
The start of this section would benefit from more details about (1) baits used for capture (based on data available for the lion; exome capture missing here?), and (2) mapping approach used (i.e. to which species: lion).

Can their approach to detect gene flow be applied to some Felinae (e.g. Panthera) in their dataset to confirm that it does indeed function for more recent events?

Methods

Exome capture used DNA baits? (typo?)

Why was the lion genome used as reference? Why not a genome with higher contiguity? Did this not potentially result in unnecessary loss of data during filtering (e.g. due to scaffold lengths)?

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Conservation genetics (incl felids), evolutionary genetics (incl felids), archival and ancient DNA (incl felids).

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 10 Jun 2021

Michael Westbury

We would like to thank Daniel Förster for taking the time to review our manuscript and for providing useful feedback that we have now incorporated into our manuscript. We have included our responses to each comment below.
Tables & Figures It seems unnecessary to avoid any Tables in the main text. For example, having the table of species analysed (Table S10) in the Methods section would benefit clarity, as these data suddenly find mention in the data processing section, as well as being featured in multiple subsequent sections.
Answer: We have included the table listing the names of the additional species included in this study to the methods section.
The mapping results (Table S1) are summarized in the Results & Discussion section, but the table could feature in the main text as well.
Answer: We have now shifted the mapping results table to the main text.

Plain language summary “assess” rather than “access”?
Answer: Fixed
Introduction The first sentence reads rather clunky.
Answer: We have now removed “a family within Carnivora” to make it smoother to read.
Results & Discussion The start of this section would benefit from more details about (1) baits used for capture (based on data available for the lion; exome capture missing here?), and (2) mapping approach used (i.e. to which species: lion).
Answer: We have now added more information about the baits and mapping approach to this section.
Can their approach to detect gene flow be applied to some Felinae (e.g. Panthera) in their dataset to confirm that it does indeed function for more recent events?
Answer: We have now tested the D3 method on some Felinae species. We find that our method does not detect gene flow between Panthera species. However, we do find significant signs of different levels of gene flow between Panthera (more specifically the snow leopard) and its sister genus, Neofelis (clouded leopard). Although the lack of detection between Panthera is unexpected based on the literature (e.g. Figueiro et al 2017 DOI: 10.1126/sciadv.1700299), this may be due to our strict filtering, artificially making the closely related species look more similar to each other. We note that the results between more distantly related Panthera species are only just not significant (p=0.066). We have included a table with the results here.

Methods Exome capture used DNA baits? (typo?)
Answer: Thank you for finding this. Yes it is a typo and has now been corrected to RNA baits.

Why was the lion genome used as reference? Why not a genome with higher contiguity? Did this not potentially result in unnecessary loss of data during filtering (e.g. due to scaffold lengths)?
Answer: We used the lion as reference genome as the baits were designed with lion sequences and to make results more comparable with our previous study (Barnett et al 2020) where we investigated the Homotherium genome within the context of extant Felinae species which were also all mapped to the same reference genome. Although there are currently other Felinae chromosome level assemblies available, they were not available at the time of the bait design. Despite this, even after filtering of shorter scaffolds, we were still left with 613 scaffolds >1Mb in length equating to 2,107,406,870bp which was more than sufficient for our purposes especially as we either relied on the complete dataset (trees) or 1Mb sliding windows (gene flow analyses).

Competing Interests: No competing interests were disclosed.