Genomic Adaptations and Evolutionary History of the Extinct Scimitar-Toothed Cat, Homotherium latidens

Barnett, Ross; Westbury, Michael V.; Sandoval-Velasco, Marcela; Vieira, Filipe Garrett; Jeon, Sungwon; Zazula, Grant; Martin, Michael D.; Ho, Simon Y.W.; Mather, Niklas; Gopalkrishnan, Shyam; Ramos-Madrigal, Jazmín; de Manuel, Marc; Zepeda-Mendoza, M. Lisandra; Antunes, Agostinho; Baez, Aldo Carmona; De Cahsan, Binia; Larson, Greger; O’Brien, Stephen J.; Eizirik, Eduardo; Johnson, Warren E.; Koepfli, Klaus-Peter; Wilting, Andreas; Fickel, Jörns; Dalén, Love; Lorenzen, Eline D.; Marques-Bonet, Tomas; Hansen, Anders J.; Zhang, Guojie; Bhak, Jong; Yamaguchi, Nobuyuki; Gilbert, M. Thomas P.

Published in:
Current Biology

DOI:
10.1016/j.cub.2020.09.051

Publication date:
2020

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC-ND

Citation for published version (APA):
Genomic Adaptations and Evolutionary History of the Extinct Scimitar-Toothed Cat, *Homotherium latidens*

**Highlights**

- Nuclear genome and exome analyses of extinct scimitar-toothed cat, *Homotherium latidens*

- *Homotherium* was a highly divergent lineage from all living cat species (~22.5 Ma)

- Genetic adaptations to cursorial and diurnal hunting behaviors

- Relatively high levels of genetic diversity in this individual

**Authors**

Ross Barnett, Michael V. Westbury, Marcela Sandoval-Velasco, ..., Jong Bhak, Nobuyuki Yamaguchi, M. Thomas P. Gilbert

**Correspondence**

m.westbury@sund.ku.dk (M.V.W.), tgilbert@sund.ku.dk (M.T.P.G.)

**In Brief**

Here, Barnett et al. sequence the nuclear genome of *Homotherium latidens* through a combination of shotgun and target-capture approaches. Analyses confirm *Homotherium* to be a highly divergent lineage from all living cat species (~22.5 Ma) and reveal genes under selection putatively related to a cursorial and diurnal hunting behavior.

Barnett et al., 2020, Current Biology 30, 5018–5025

December 21, 2020 © 2020 The Authors. Published by Elsevier Inc.

https://doi.org/10.1016/j.cub.2020.09.051
**Summary**

*Homotherium* was a genus of large-bodied scimitar-toothed cats, morphologically distinct from any extant felid species, that went extinct at the end of the Pleistocene [1–4]. They possessed large, saber-form serrated canine teeth, powerful forelimbs, a sloping back, and an enlarged optic bulb, all of which were key characteristics for predation on Pleistocene megafauna [5]. Previous mitochondrial DNA phylogenies suggested that it was a highly divergent sister lineage to all extant cat species [6–8]. However, mitochondrial phylogenies can be misled by hybridization [9], incomplete lineage sorting (ILS), or sex-biased dispersal patterns [10], which might be especially relevant for *Homotherium* since widespread mito-nuclear discrepancies have been uncovered in modern cats [10]. To examine the evolutionary history of *Homotherium*, we generated a ~7x nuclear genome and a ~38x exome from *H. latidens* using shotgun and target-capture sequencing approaches. Phylogenetic analyses reveal *Homotherium* as highly divergent (~22.5 Ma) from living cat species, with no detectable signs of gene flow. Comparative genomic analyses found signatures of positive selection in several genes, including those involved in vision, cognitive function, and energy consumption, putatively consistent with diurnal activity, well-developed social behavior, and cursorial hunting [5]. Finally, we uncover relatively high levels of genetic diversity, suggesting that *Homotherium* may have been more abundant than the limited fossil record suggests [3, 4, 11–14]. Our findings complement and extend previous inferences from both the fossil record and initial molecular studies, enhancing our understanding of the evolution and ecology of this remarkable lineage.

**Results and Discussion**

We used a combination of shotgun and whole genome and exome target-capture Illumina sequencing to generate the nuclear genome of a single *Homotherium latidens* individual to a depth of ~7x coverage, and its exome to ~38x coverage. The genome was sequenced from a fossil humerus (specimen YG 439.38), which was determined to be older than the limits of radiocarbon dating (>47.5 kya [UCIAMS-142835]), recovered from Pleistocene permafrost sediments near Dawson City, Yukon Territory, Canada.
Warren E. Johnson,16,17,18 Klaus-Peter Koepfli,16 Andreas Wilting,19 Jörens Fickel,19,20 Love Dalén,21,22 Eline D. Lorenzen,1,7 Tomas Marques-Bonet,8,23,24,25 Anders J. Hansen,7,34 Guojie Zhang,26,27,28 Jong Bhak,2,3,29,30 Nobuyuki Yamaguchi,31 and M. Thomas P. Gilbert1,5,7,*

12The Palaeogenomics and Bio-Archaeology Research Network, Research Laboratory for Archaeology and History of Art, University of Oxford, 1 South Parks Road, Oxford OX1 3TG, UK
13Laboratory of Genomic Diversity, Center for Computer Technologies, ITMO University, 49 Kronverkskiy Pr., St. Petersburg 197101, Russia
14Guy Harvey Oceanographic Center, Halmos College of Natural Sciences and Oceanography, Nova Southeastern University, 8000 North Ocean Drive, Ft Lauderdale, FL 33304, USA
15Laboratory of Genomics and Molecular Biology, Escola de Ciências da Saúde e da Vida, PUCRS, Porto Alegre, RS, Brazil
16Center for Species Survival, Smithsonian Conservation Biology Institute, National Zoological Park, 1500 Remount Road, Front Royal, VA 22630, USA
17The Walter Reed Biosystematics Unit, Museum Support Center MRC-534, Smithsonian Institution, 4210 Silver Hill Rd., Suitland, MD 20746-2863, USA
18Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Silver Spring, MD 20910, USA
19Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowsalke-Str. 17, Berlin 10315, Germany
20Institute for Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Strasse 24-25, Potsdam 14476, Germany
21Centre for Palaeogenetics, Svante Arrenhius vag 20C, Stockholm SE-10691, Sweden
22Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Box 50007, Stockholm 10405, Sweden
23CNAG-CRG, Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Baildir Reixac 4, Barcelona 08028, Spain
24Institució Catalana de Recerca i Estudis Avançats, ICREA, Barcelona 08003, Spain
25Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Edifici ICTA-ICP, c/ Columnes s/n, Cerdanyola del Vallès, Barcelona 08193, Spain
26BGI-Shenzhen, Shenzhen 518083, China
27Section for Ecology and Evolution, Department of Biology, Faculty of Science, University of Copenhagen, Universitetsparken 15, Copenhagen, Denmark
28State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China
29Clinomics, Inc., Ulsan 44919, Republic of Korea
30Personal Genomics Institute (PGI), Genome Research Foundation (GRF), Osong 28160, Republic of Korea
31Institute of Tropical Biodiversity and Sustainable Development, Universiti Malaysia Terengganu, Kuala Nerus, Terengganu 21030, Malaysia
32INCT Ecologia, Evolução e Conservacção da Biodiversidade (INCT-EECBio), Goiânia, GO, Brazil
33Instituto Pró-Carnívoros, Atibaia, SP, Brazil
34Instituto de Genética, Universidad de los Andes, Bogotá, Colombia
35These authors contributed equally
36Lead Contact
*Correspondence: m.westbury@sund.ku.dk (M.V.W.), tgilbert@sund.ku.dk (M.T.P.G.)

https://doi.org/10.1016/j.cub.2020.09.051

This nuclear dataset represents a large increase in the number of independent genetic loci available for analysis, each of which has been uniquely passed down over many generations of lineage sorting, drift, selection, and recombination. By investigating such a large number of independent loci, we are able to better understand the Homotherium lineage and make broad evolutionary inferences that are not possible with datasets comprising a more limited number of markers. Therefore, we refer to the genus without mentioning the species name throughout the manuscript if results can be interpreted as representing the lineage as a whole. After filtering for quality and mapping against the lion (Panthera leo) nuclear genome [15] (NCBI accession: JAAVKH000000000), we undertook phylogenomic analyses using a supermatrix of the exonic data, as well as a combined tree constructed from multiple individual gene trees. The inferred topologies are consistent with those of previous studies based on both fossil [2] and mitochondrial datasets [8] in which Homotherium is a sister lineage to all extant cat species (Figures 1 and S1).

We next dated the divergence between Homotherium and the extant cat species using MCMCTree [16] and four fossil calibrations (see STAR Methods and Table S1). We found that Homotherium diverged from extant cat species ~22.5 Ma (95% credibility interval [CI] 17.4–27.8 Ma), close to the Oligocene/Miocene boundary ~23 Ma. This timing coincides with either late Proailurus sp. or early Pseudaelurus sp. as potential ancestors [17]. Our inferred date is similar to the estimate of ~20 Ma (95% CI 18.2–22.0 Ma) that was previously determined based on mitochondrial genomes [8]. This deep divergence confirms that Homotherium belonged to a distinct clade from all living felines, providing support for Machairodontinae as a separate subfamily from Felinae.

Still, the clear-cut phylogenetic relationships reported here might not represent the full evolutionary relationship of Homotherium to living felines. Several recent genomic studies have demonstrated that interspecific hybridization is much more common than previously believed in several lineages within Carnivora, including crown canids [18], hyenas [9], and extant cats [10, 19, 20]. For this reason, we explored our dataset for evidence of topological incongruencies that could arise due to incomplete lineage sorting (ILS) or potentially indicate gene flow. In regard to gene flow, Homotherium is of considerable interest, as it once spanned from southern Africa, across Eurasia and North America, to South America, arguably the largest geographical range of all the saber-toothed cats.
(Machairodontinae), although it is uncertain whether it occurred in all those regions simultaneously [2–4, 14, 21]. In addition to occupying a huge geographical distribution, it also colonized a variety of different habitats, from densely vegetated Java to open steppe-tundra across the Holarctic [13, 22]. Furthermore, fossil evidence suggests that *Homotherium* succeeded in expanding its distribution despite potential competition with other sympatric large cats, including lions (*Panthera leo* and the now-extinct cave lion *P. spelaea*), leopards (*P. pardus*) and the now-extinct dirk-toothed cat *Megantereon cultridens* across Eurasia and Africa, tigers (*P. tigris*) in Southeast Asia, and jaguars (*P. onca*), American lions (*P. atrox*), and other now-extinct saber-toothed cats (*Smilodon* spp.) in the Americas [2, 23].

To explore whether gene flow may have occurred between *Homotherium* and living cat species, we investigated conflicting phylogenetic signals in our phylogenomic tree using ASTRAL-II [24] and DiscoVista [25] (Figures 1 and S2). Although our results reproduce some of the notorious cases of genealogical discordance in other felids (e.g., the relationships among lion, jaguar, and leopard [node 6]; or the relative placement of the *Lynx* and *Puma* lineages [node 15]), gene-tree quartet frequencies of alternative topologies for the node separating *Homotherium* from all other cat species were lower than would be expected (less than one-third) if gene flow or ILS had played an important role in this divergence (Figure 1). This implies no gene flow between the lineage that our *Homotherium* sample represents and the extant cat species, but does not exclude the possibility of undetectable gene flow between the ancestor of all extant cat lineages and the *Homotherium* lineage. The most plausible explanation for the inferred lack of gene flow is the deep divergence between *Homotherium* and extant cat species. The oldest signal of gene flow we would be able to detect with the current dataset would be after the divergence of stem Felinae ~14 Ma (Figure 1). This means that any gene flow over an ~8-million-year period between Machairodontinae and Felinae is undetectable. In contrast, the radiation of the main lineages of modern cats occurred within a ~5-million-year period (Figure 1). This rapid radiation may have allowed gene flow to occur between these lineages before they became too genetically divergent from one another to produce viable offspring [10]. Although alternative explanations could apply to the apparent lack of gene flow, these are much more unlikely. One possibility is that *Homotherium* were simply unable to interact with other cat species due to either eco-geographical barriers, competitive exclusion, or low population density. The former seems implausible given their above-described distributions and affiliations to very different habitats. Competitive exclusion also seems unlikely, as *Homotherium* fossils are found at the same sites as *Panthera*, and they are likely contemporaneous [23]. Low population density could be interpreted as more likely, as the fossil record of *Homotherium* is considerably more fragmentary than those of other contemporary large cats, including *Smilodon* and *Panthera* (especially lions), tentatively suggesting that *Homotherium* occurred at lower population densities [3, 4, 11–14]. However, even at low densities, occasional contact with sympatric species would not be precluded.

Another alternative explanation is that behavioral and/or other ecological mechanisms may have prevented hybridization. An example of this can be seen between extant lions and leopards. Though they often occupy the same area, leopards are known to actively avoid lions [26]. A similar behavioral and/or ecological mechanism of avoidance may have occurred between *Homotherium* and other sympatric felids. To further add evidence to our finding of a lack of gene flow, although archaic admixture between living mammalian species and unknown, unsampled lineages has been uncovered (e.g., humans [27], bonobos [28], and dogs [18]), including signatures of introgression into leopards from an extinct Felinae species closely related to *Panthera* [19], no signs of archaic admixture between living cat species and an unknown lineage as divergent as *Homotherium* have been detected in previous studies.

This divergence thus leads to the question of what genetic adaptations occurred as these lineages differentiated from one another. To address the question of what genomic underpinnings made *Homotherium* unique, we performed a comparative analysis of concatenated sequences of 21,034 exons (total length 29,216,712 bp) (Figure S1). A separate correlated-rates relaxed clock was applied to each of the three codon positions. Relative frequencies of the three possible bifurcations (possible arrangements of a quartet on an unrooted tree) are shown for the internal branch containing *Homotherium* (branch 12) and the remaining nodes in Figure S2. Dashed lines show the threshold value of one-third, shown theoretically to be the minimum frequency for a true bipartition. Relevant branch labels have been given based on those in Figure S2. Blue horizontal bars represent 95% credibility intervals of node times. Blue circles indicate internal nodes with fossil-based age constraints. A geological timescale is shown below the tree (Q = Quaternary, Pl = Pleistocene, Ps = Pliocene). Ages of key nodes in the phylogeny can be found in Table S1.
genomics analysis and uncovered signatures of positive selection in several protein-coding regions across its genome (Table S2). We found evidence of positive selection ($5 > d_n/d_s$ ratio > 1) in 230 genes out of the 2,191 assessed 1:1 orthologous loci. Of these 230 genes, 31 were deemed to be highly significant (a free ratio >2) and were further explored for putative functions and phenotypic roles (Table S3). Several highly significant positively selected genes were consistent with a putative diurnal behavior in *Homotherium* [5] (Figure 2). Strong positive selection was detected in genes linked to vision (B3GALNT2, AGBL5, CAPNS2, and SLC1A7) that have known phenotypes involving retinal degeneration, retinitis pigmentosa, hydrolysis of natural lens proteins, and visual processing, e.g., [29, 30]. We also found evidence of positive selection in genes linked with entrainment and the master regulation of circadian clock rhythm (SFPQ and Per1) [31, 32]. Although not in the same genes specifically, previous studies have shown polymorphisms in circadian clock regulatory genes to be related to diurnal preference [33, 34], strengthening the link between circadian genes and diurnal behavior. While speculative, these results do provide some support toward the idea that *Homotherium* hunted during the day, unlike many extant cat species, which are either crepuscular or nocturnal [35]. This hypothesis is further supported by several anatomical features of this species, including an enlarged optic bulb and a large and complex visual cortex [5].

Signals of positive selection were also inferred in genes that were plausibly involved in adaptations helping to increase endurance for a cursorial hunting style in *Homotherium*. These include genes with major influences on the respiratory system/hypoxia (TMEM45A, which has been tied to the deprivation of oxygen supply), the circulatory system (F5 and MMP12), angiogenesis (ECSR—involvement in the formation of new blood vessels), adipogenesis (TAF8—a positive adipogenesis regulator playing a vital role in energy homeostasis), the respiratory/circulatory systems (MMP12, with a role in aneurysm, a result of a weakened blood vessel wall), and mitochondrial respiration (AK3, ISCU, and SURF, with relevance in skeletal muscle and the myopathy phenotype with severe exercise intolerance, e.g., [36] and the biogenesis of the cytochrome c oxidase [COX] complex). Novel adaptations in these genes may have enabled sustained running necessary for hunting in more open habitats and the pursuit of prey until their point of exhaustion. The synergistic interplay of these various functional enhancements could have also been assisted by improvements in bone mineralization (PGD). Improved bone mineralization would have been invaluable to develop and sustain a robust skeletal framework and powerful forelimbs needed for a putatively cursorial hunting style. Furthermore, although not the PGD gene specifically, two other genes playing a role in bone development and repair (DMP1 and PTN) have been found to be under positive selection in a number of...
carnivore genomes (including several felid species) [37], suggesting that robust bones may be key for adaptation to predatory behavior. The signatures of selection in the above-mentioned genes add additional evidence to a cursorial hunting style suggested by post-cranial skeletal morphological data, including less-retractable claws, which are thought to improve traction during long-distance, medium-speed travel, similar to canids or hyenas [23], along with a higher brachial index (radius to humerus ratio) [2, 14, 38, 39].

Finally, although social behavior is very complex and therefore difficult to directly link to certain genetic traits, we found evidence of positive selection in genes that putatively play roles in cognition/behavior (SCTR, which is linked to impaired synaptic plasticity and social behavior [Mouse Genome Informatics—MGI database—http://www.informatics.jax.org]) and the nervous system (NTF3—influencing the nerve growth factor). We speculate that these genes may have played a role in coordinated social interactions that others have suggested were required for Homotherium to successfully hunt large prey species [13, 14], but additional data would be needed to support this notion. Although our results do not provide definitive links between genes under positive selection and their ecological consequences, we provide the starting ground for functional studies into the relationship between the genetics and ecology of the extinct hypercarnivore Homotherium. We find a number of positively selected genes whose known functions complement the paleontological data and represent the putative genetic underpinnings for several unique adaptations of Homotherium that appear to be closely linked to cursorial group-hunting strategies and diurnal behavior.

As mentioned above, the fossil record for Homotherium is considerably more fragmentary than those of other contemporary large cats, including Smilodon and Panthera (especially lions), tentatively suggesting that Homotherium occurred at lower population densities [3, 4, 11–14]. This leads to the question of whether the fragmentary fossil record of Homotherium truly reflects low population densities or is a stochastic result based on the likelihood of preservation due to their habitat or behavior. To examine the relative abundance of Homotherium latidens compared with extant cat species, we compared the genetic diversity (which is correlated with the effective population size), in the form of heterozygosity, from our Homotherium latidens genome to that of single individual representatives from 15 extant cat species included in this study. We used two different approaches (autosome-wide and exome-wide) to estimate heterozygosity. To ensure comparability between species when calculating heterozygosity, we only considered transversions to remove biases that could occur due to the much higher ancient DNA (aDNA) damage levels expected in our Homotherium individual relative to the extant species. For the autosome-wide heterozygosity estimates, to further ensure comparability between results, we downsampled all individuals to a comparable read depth to the Homotherium latidens genome before estimating heterozygosity. As we performed exome capture on our Homotherium latidens individual and therefore had relatively high coverage, we did not perform any downsampling with this dataset. We caution that our estimations are based on estimates from single representatives of each species, and thus may not hold true species-wide. However, it is striking that our Homotherium latidens genome exhibits medium to high levels of genetic diversity, higher than that inferred for the other big cat species. This inference held true regardless of whether autosome-wide (Figure 3) or exome-wide heterozygosity (Figure S3) was compared. Given the relationship between genome-wide heterozygosity, effective population size, and census size, this finding suggests that...
Homotherium latidens was relatively abundant in contrast to previously suggested low population densities based on low levels of mitochondrial genetic diversity compared with other species of large cats [5]. However, as this previous suggestion was purely based on the single-locus mitochondrial genome from three individuals, the inference on diversity was limited.

Our finding of a putative high abundance in Homotherium latidens is strengthened by Homotherium having had the largest geographical range of all Machairodontinae, ranging from southern Africa, across Eurasia and North America, to South America [2–4, 14, 21]. Although genetic diversity often reflects population size, the demography and life history traits of the population/species can also play a role. One notable factor that may have resulted in higher levels of genetic diversity in Homotherium occupying the (sub)Arctic would be long-distance migration between structured populations. In general, species (including cursorial predators) dwelling in areas with dramatic seasonal variability or low primary productivity (e.g., the (sub)Arctic travel greater distances than those in more stable and productive regions [40]. Therefore, it may have been possible that Homotherium, as a cursorial predator living in the less-productive (sub)Arctic region, moved across an enormous range, leading to increased gene flow between distant populations. Regardless, the huge geographical distribution of Homotherium, and the apparent ability of the genus to colonize different habitats, suggests it was a very successful taxon and fits our inference that it may have been relatively abundant.

The apparent success of Homotherium leads to the question of why the lineage was unable to survive to the present. Although it cannot be known for certain, some of the exact adaptations/specializations that led to Homotherium’s success could also have led to its downfall. Toward the end of the Late Pleistocene, a decrease in large prey availability may have caused more direct competition with other cat species that were likely more effective at capturing the remaining smaller prey species. The specific adaptations Homotherium had acquired would have suddenly become less advantageous, leaving the lineage stuck on a path to extinction.

Conclusions
The genome sequence of the scimitar-toothed cat, Homotherium latidens, improves our understanding of its evolutionary relationships with other extant cat species and into the genetic underpinnings of its unique adaptations. Our results demonstrate that Homotherium was highly divergent (~22.5 Ma) from all living cat species and did not undergo any detectable gene flow with extant felid species after their initial radiation ~14 Ma. This placement supports the recognition of Machairodontinae as a distinct subfamily within the Felidae. Furthermore, we found evidence of positive selection in several genes involved in vision, cognitive function, and energy consumption, potentially consistent with the diurnal and hunting/social behavior of this extinct lineage. Finally, we uncovered relatively high levels of genetic diversity in our Homotherium latidens individual, suggesting that it was not only a successful lineage, but also rather abundant relative to extant cat species. Our study shows how the fossil record and paleogenomics can be used synergistically to better understand the evolution and ecology of extinct species that lack extant close relatives for comparison.

STAR METHODS
Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Fossil information
  - Radiocarbon dating
- **METHOD DETAILS**
  - Ancient DNA extraction, library preparation, and sequencing
  - Genome capture
  - Modern DNA extraction, library preparation, and sequencing
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Data processing pipeline
  - Extracting coding and protein sequences
  - Phylogenetic analyses
  - Phylogenomic dating
  - Tests of positive selection
  - Genetic diversity

SUPPLEMENTAL INFORMATION
Supplemental Information can be found online at https://doi.org/10.1016/j.cub.2020.09.051.

ACKNOWLEDGMENTS
This project 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 and ERC Consolidator Award 681396—Extinction Genomics to M.T.P.G. Portions of this manuscript were prepared while W.E.J. held a National Research Council Research Associateship Award at the Walter Reed Army Institute of Research (WRAIR). The material has been reviewed by WRAIR and there is no objection to its presentation and/or publication. The opinions and assertions contained herein are the private views of the authors, and are not to be construed as official or as reflecting true views of the Department of the Army or the Department of Defense. 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. Thanks to the field crew that recovered bone YG 439.38: Beth Shapiro, Matthias Stiller, Duane Froese, and Tyler Kuhn. We are grateful to the placer gold mining community and the Tr’ondëk Hwëch’in for their continued support and partnership with our research in the Klondike goldfields region. We thank Rheaon Slade for his assistance on the layout of Figure 2. Lastly, we thank Jean-Marie Rouillard (Arbor Biosciences) for help with balt design.

AUTHOR CONTRIBUTIONS
The authors declare no competing interests.

REFERENCES


# STAR★METHODS

## KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Biological sample</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homotherium latidens</td>
<td>This paper</td>
<td>Yukon Government Palaeontology Program</td>
</tr>
<tr>
<td>Lion (Panthera leo)</td>
<td>This paper</td>
<td>Copenhagen Zoo</td>
</tr>
<tr>
<td>Sand cat (Felis margarita)</td>
<td>This paper</td>
<td>Leibniz Institute for Zoo and Wildlife Research</td>
</tr>
<tr>
<td>Fishing cat (Prionailurus viverrinus)</td>
<td>This paper</td>
<td>Leibniz Institute for Zoo and Wildlife Research</td>
</tr>
<tr>
<td>Leopard cat (P. bengalensis)</td>
<td>This paper</td>
<td>Leibniz Institute for Zoo and Wildlife Research</td>
</tr>
<tr>
<td>Caracal (Caracal caracal)</td>
<td>This paper</td>
<td>Copenhagen Zoo</td>
</tr>
<tr>
<td><strong>Chemicals, Peptides, and Recombinant Proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AccuPrim Pfx DNA Polymerase</td>
<td>Invitrogen</td>
<td>Cat# 12344024</td>
</tr>
<tr>
<td>Phusion High-Fidelity DNA Polymerase</td>
<td>New England Biolabs</td>
<td>Cat# M0530S</td>
</tr>
<tr>
<td>KAPA HiFi HotStart polymerase</td>
<td>Roche</td>
<td>Cat# KK2801 07959052001</td>
</tr>
<tr>
<td><strong>Critical Commercial Assays</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEBNext DNA Library Prep Master Mix Set</td>
<td>New England Biolabs</td>
<td>Cat# E6070</td>
</tr>
<tr>
<td>MinElute PCR Purification Kit</td>
<td>QIAGEN</td>
<td>Cat# 28006</td>
</tr>
<tr>
<td>QIAquick column system</td>
<td>QIAGEN</td>
<td>Cat# 28104</td>
</tr>
<tr>
<td>myBaits target enrichment kit</td>
<td>Arbor Biosciences</td>
<td>NA</td>
</tr>
<tr>
<td>Kingfisher blood DNA extraction kit</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 98010196</td>
</tr>
<tr>
<td>PCR-free Truseq Illumina library kit</td>
<td>Illumina</td>
<td>Cat# 20015962</td>
</tr>
<tr>
<td><strong>Deposited data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Felis nigripes</td>
<td>99 Lives Cat Genome Sequencing Initiative</td>
<td>NCBI SRA accession code: SRR2511865</td>
</tr>
<tr>
<td>Felis catus</td>
<td>99 Lives Cat Genome Sequencing Initiative</td>
<td>NCBI SRA accession code: SRR2224864</td>
</tr>
<tr>
<td>Felis chaus</td>
<td>[20]</td>
<td>NCBI SRA accession code: SRR2062187</td>
</tr>
<tr>
<td>Felis margarita</td>
<td>This study</td>
<td>NCBI Bioproject accession code: PRJNA649575</td>
</tr>
<tr>
<td>Felis margarita</td>
<td>[20]</td>
<td>NCBI SRA accession code: SRR2062538</td>
</tr>
<tr>
<td>Prionailurus viverrinus</td>
<td>This study</td>
<td>NCBI Bioproject accession code: PRJNA649563</td>
</tr>
<tr>
<td>Prionailurus bengalensis</td>
<td>NA</td>
<td>NCBI SRA accession code: SRR2062628</td>
</tr>
<tr>
<td>Prionailurus bengalensis</td>
<td>This study</td>
<td>NCBI Bioproject accession code: PRJNA649572</td>
</tr>
<tr>
<td>Lynx pardinus</td>
<td>[41]</td>
<td>European nucleotide archive accession code: ERA562804</td>
</tr>
<tr>
<td>Acinonyx jubatus</td>
<td>[42]</td>
<td>NCBI SRA accession code: SRS1123638</td>
</tr>
<tr>
<td>Caracal caracal</td>
<td>This study</td>
<td>NCBI SRA sample accession code: SAMN15096300</td>
</tr>
<tr>
<td>Panthera tigris altaica</td>
<td>[43]</td>
<td>NCBI SRA accession code: SRR836311</td>
</tr>
<tr>
<td>Panthera tigris tigris</td>
<td>[43]</td>
<td>NCBI SRA accession code: SRR836354</td>
</tr>
<tr>
<td>Panthera uncia</td>
<td>[43]</td>
<td>NCBI SRA accession code: SRR836372</td>
</tr>
<tr>
<td>Panthera onca</td>
<td>[19]</td>
<td>NCBI SRA sample accession code: SAMN05907657</td>
</tr>
<tr>
<td>Panthera pardus</td>
<td>[37]</td>
<td>NCBI SRA accession code: SRR3041424</td>
</tr>
<tr>
<td>Panthera leo</td>
<td>[43]</td>
<td>NCBI SRA accession code: SRR836361</td>
</tr>
<tr>
<td>Panthera leo</td>
<td>[43]</td>
<td>NCBI SRA accession code: SRR836370</td>
</tr>
<tr>
<td>Hyaena hyaena</td>
<td>[44]</td>
<td>NCBI Bioproject accession code: PRJNA390068</td>
</tr>
<tr>
<td>Crocuta crocuta</td>
<td>[9]</td>
<td>NCBI Bioproject accession code: PRJNA554753</td>
</tr>
<tr>
<td>Homotherium latidens</td>
<td>This study</td>
<td>NCBI Bioproject accession code: PRJNA649760</td>
</tr>
</tbody>
</table>

(Continued on next page)
RESOURCE AVAILABILITY

Lead Contact
Further information and requests for reagents and data may be directed to and will be fulfilled by the Lead Contact, Michael V Westbury (m.westbury@sund.ku.dk).

Materials Availability
This study did not generate new unique reagents.

Data and Code Availability
NCBI accession codes for all newly generated raw fastq reads can be found in the key resources table.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Fossil information
Specimen YG 439.38 was sampled from the collections of the Yukon Government Palaeontology Program housed in Whitehorse, Yukon Territory, Canada, and was recovered from the placer gold mining site of Dominion Creek, near Dawson City, Yukon Territory. Initially identified as a large feline humerus by GZ, it was later assigned based on morphology to Homotherium, the scimitar-toothed cat. Here, we refer to our fossil material as H. latidens in accordance with the mitochondrial genomic data which suggested that all late Pleistocene Holarctic Homotherium belonged to a single species [8].

Radiocarbon dating
A section of bone was delivered to Stafford Research LLC for radiocarbon dating. Briefly, samples of crushed bone were decalcified and washed, treated with 0.05 N NaOH overnight to remove humics, soaked in 0.1 N HCl, gelatinised at 60°C at pH 2, and ultrafiltered at 30 kDa. Resultant collagen was then graphitised and run on an accelerator mass spectrometer according to standard protocol [61, 62].

Continued

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oligonucleotides</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illumina compatible adapters</td>
<td>Illumina</td>
<td>NA</td>
</tr>
<tr>
<td>Hybridization capture probes</td>
<td>Arbor Biosciences</td>
<td>NA</td>
</tr>
<tr>
<td>Software and Algorithms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PALEOMIX v1.2.5</td>
<td>[45]</td>
<td><a href="https://github.com/MikkelSchubert/paleomix">https://github.com/MikkelSchubert/paleomix</a></td>
</tr>
<tr>
<td>AdapterRemoval v2.0.0</td>
<td>[46]</td>
<td><a href="https://github.com/MikkelSchubert/adapterremoval">https://github.com/MikkelSchubert/adapterremoval</a></td>
</tr>
<tr>
<td>BWA v0.7.5a</td>
<td>[47]</td>
<td><a href="http://bio-bwa.sourceforge.net/">http://bio-bwa.sourceforge.net/</a></td>
</tr>
<tr>
<td>Picard v2.18.0</td>
<td>NA</td>
<td><a href="https://broadinstitute.github.io/picard">https://broadinstitute.github.io/picard</a></td>
</tr>
<tr>
<td>GATK v3.8.0</td>
<td>[48]</td>
<td><a href="https://gatk.broadinstitute.org/hc/en-us">https://gatk.broadinstitute.org/hc/en-us</a></td>
</tr>
<tr>
<td>mapDamage v2.0.5</td>
<td>[49, 50]</td>
<td><a href="https://ginojac.github.io/mapDamage/">https://ginojac.github.io/mapDamage/</a></td>
</tr>
<tr>
<td>vcf-tools v0.1.14</td>
<td>[51]</td>
<td><a href="https://vcftools.github.io/">https://vcftools.github.io/</a></td>
</tr>
<tr>
<td>bedtools v2.29.0</td>
<td>[52]</td>
<td><a href="https://github.com/arg5x/bedtools2">https://github.com/arg5x/bedtools2</a></td>
</tr>
<tr>
<td>EMBOSs v6.6.0</td>
<td>[53]</td>
<td><a href="https://www.ebi.ac.uk/Tools/emboss/">https://www.ebi.ac.uk/Tools/emboss/</a></td>
</tr>
<tr>
<td>OrthoMCL v2.0.9</td>
<td>[54]</td>
<td><a href="https://orthomcl.org/">https://orthomcl.org/</a></td>
</tr>
<tr>
<td>RAxML v8.2.11</td>
<td>[55]</td>
<td><a href="https://sco.h-its.org/exelixis/software.html">https://sco.h-its.org/exelixis/software.html</a></td>
</tr>
<tr>
<td>Tracer</td>
<td>[57]</td>
<td><a href="https://github.com/beast-dev/tracer/releases/latest">https://github.com/beast-dev/tracer/releases/latest</a></td>
</tr>
<tr>
<td>PRANK</td>
<td>[58]</td>
<td><a href="https://www.ebi.ac.uk/research">https://www.ebi.ac.uk/research</a> goldman/software/prank</td>
</tr>
<tr>
<td>SAMtools v1.6</td>
<td>[59]</td>
<td><a href="http://samtools.sourceforge.net/">http://samtools.sourceforge.net/</a></td>
</tr>
<tr>
<td>ANGSD v0.921</td>
<td>[60]</td>
<td><a href="https://github.com/ANGSD/angsd">https://github.com/ANGSD/angsd</a></td>
</tr>
</tbody>
</table>
METHOD DETAILS

Ancient DNA extraction, library preparation, and sequencing

Samples of cortical bone were taken from a long bone element (approx. 1 cm³) using a Dremel drill and reduced to powder in a Mikrodismembrator. DNA extraction was performed as described in Orlando et al. [63] in a dedicated ancient DNA laboratory in parallel with negative extraction controls. DNA extracts and negative controls were built into genomic libraries using the NEB E6070 kit and a slightly modified version of the protocol as used by Vilstrup et al. [64]. Briefly, the extract (30 μl) was end-repaired and passed through a MinElute column. The collected flow-through was adaptor-ligated and passed through a QiaQuick column. Adaptor fill-in reaction was performed on the flowthrough, before a final incubation at 37°C (30 min) followed by inactivation overnight at –20°C.

We amplified the resulting library in a 50 μl volume reaction, using 25 μl of library for 12 cycles under the following reaction conditions. Final concentrations of reagents in the master mix were: 1.25 U AccuPrime Pfx DNA Polymerase (Invitrogen), 1x AccuPrime Pfx reaction mix (Invitrogen), 0.4 mg/mL BSA, 120 nM primer inPE, 120 nM of a multiplexing indexing primer containing a unique 6-nucleotide index code (Illumina). 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 72°C extension for 30 s and a final extension step at 72°C for 7 min. Amplified libraries were checked for the presence of DNA on a 2% Agarose gel before purification using the QIAquick column system (QIAGEN) and quantification on an Agilent 2100 BioAnalyzer. Quantified libraries were communally pooled in equimolar ratios and sequenced as 100 bp single-end reads on three lanes of an Illumina HiSeq2000 platform at the Danish National High-Throughput Sequencing Centre and 100 bp paired-end reads on two lanes of Illumina Hiseq2000 at BGI Copenhagen.

Genome capture

Shotgun-sequenced libraries were assessed for endogenous content and the best libraries were taken forward for two sets of capture experiments. The first method used biotinylated RNA probes transcribed from fresh DNA extract derived from modern lion tissue in an attempt to capture the whole genome (whole genome capture/WGC). The second method used previously published lion genomic data [43] to identify exon coding regions and create biotinylated RNA baits that covered these regions (exome capture). Both types of baits were generated by Arbor Biosciences (Ann Arbor, MI, USA). Capture experiments using both sets of baits were carried out using the myBaits target enrichment kit and following the instructions described in manual V3 (https://arborbiosci.com/wp-content/uploads/2017/10/MYbaits-manual-v3.pdf) to enrich the ancient libraries for endogenous DNA. After capture and cleanup, enriched libraries were re-amplified for further sequencing using either Phusion polymerase or KAPA HiFi Hot-Start polymerase with primers IS5_reamp.P5 and IS6_reamp.P7 over 14 cycles [65]. Following quantification using a TapeStation instrument, amplified captured libraries were sequenced on an Illumina HiSeq2000 platform. In total we sequenced 303,621,590 single-end reads and 524,795,913 paired-end reads.

Modern DNA extraction, library preparation, and sequencing

We extracted DNA from the sand cat (Felis margarita), fishing cat (Prionailurus viverrinus), leopard cat (P. bengalensis), and caracal (Caracal caracal) on a KingFisher Duo Prime robot using the Kingfisher blood DNA extraction kit (Thermo Fisher Scientific) according to the manufacturer’s instructions. The extracts were built into PCR-free Truseq Illumina sequencing libraries and sequenced on an Illumina HiSeqX using 2 x 150 bp paired-end sequencing at the National Genomics Infrastructure (NGI) in Stockholm. The sand cat, fishing cat, leopard cat, and caracal had 1,706,953,684, 992,351,075, 1,264,528,157, 1,596,359,637 raw paired-end reads sequenced respectively.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data processing pipeline

Post-sequencing read processing of Homotherium was performed using the PALEOMIX v1.2.5 pipeline [45]. Removal of adaptor contamination and adaptor dimers, as well as trimming of low-quality bases (BaseQ < 5 or Ns) was accomplished with AdapterRemoval v2.0.0 [46], while also removing reads shorter than 30 bp or with over 50 bp missing data. All retained reads were mapped against the publicly available African lion reference genome [15] (NCBI Bioproject: PRJNA615082, NCBI Accession: JAAVKH000000000) with BWA-MEM v0.7.5a and default parameters [47]. We used Picard v2.18.0 (http://broadinstitute.github.io/picard/) to identify and filter PCR duplicates by the 5-prime end mapping coordinate. GATK v3.8.0 [48] was used to perform an insertion/deletion (indel) realignment step to adjust for increased error rates at the ends 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. DNA damage patterns were assessed, and base quality scores were recalibrated using mapDamage v2.0.5 [49, 50]. Genotypes were called with HaplotypeCaller from GATK v3.8. Data processing of the extant species followed the same pipeline with slight modifications: no minimum read length cut-off or missing data cut-off was applied during the adaptor trimming step, and bases were not recalibrated using mapDamage. These steps were removed as the data from the extant species was not highly fragmented and damaged as is expected for ancient DNA. For the sand cat, fishing cat, leopard cat, and caracal 1,661,576,497, 966,303,222, 1,229,116,793, 1,563,010,218 unique reads successfully mapped to the reference nuclear genome respectively. 195,372,434 Homotherium latidens unique reads successfully mapped to the nuclear genome and 14,900,381 successfully mapped to the exome.
Extracting coding and protein sequences

To obtain the coding sequence (CDS) regions for each species, we ignored all indels and calculated a consensus genome sequence using vcf-consensus from vcf-tools v0.1.14 [51]. Subsequently, since vcf-consensus assumes the reference sequence in the case of missing data, we masked all regions in which depth was lower than 5x using maskFastaFromBed from bedtools v2.23.0 [52]. Finally, we used the gene annotations (v2) from the reference African lion genome to extract the coding sequences from the masked genomic sequence, and converted them to protein using transeq from EMBOSs v6.6.0 [53]. Afterward, orthologs were identified using the OrthoMCL v2.0.9 pipeline [54] with a BLAST e-value cut-off of $10^{-5}$ and mcl inflation of 1.5, only retaining 1-1 orthologs for downstream analyses.

Phylogenetic analyses

Custom python scripts were used to construct a multiple sequence alignment of the genotypes for the CDS regions. To achieve this, diploid genotypes were extracted from the aforementioned VCF file for every position in the African lion reference genome with a CDS annotation (v2 annotation). Genotypes passing filters (minimum genotype quality of 20, minimum read depth of 4) were used to populate the multiple sequence alignments with IUPAC ambiguity codes, while genotypes not passing the filters were replaced by the character ‘N’. Indel variants were ignored (replaced by ‘N’) such that the lengths of CDS regions were the same for every taxon. Genotypes from antisense CDS regions were reverse-complemented such that the multiple sequence alignment for each gene consisted of a concatenation of sense genotypes from all CDS regions.

For phylogenetic reconstruction, all of the aforementioned CDS alignments were concatenated into a supermatrix of 29,216,712 bp. Of these, 370,311 (1.3% of total) positions were completely undetermined. RAxML v8.2.11 [55] was used to perform a maximum-likelihood phylogenetic analysis, with 100 bootstrap replicates (Figure S1). The dataset was partitioned such that a separate GTR+G model was assigned to each of the three codon positions.

In addition to the phylogenetic tree estimated using the concatenated exon sequences, a separate phylogenetic tree was constructed for 13,164 genes, using the concatenated exon sequences for each gene. Similar to the supermatrix tree, a separate GTR+G model was assigned for these maximum-likelihood gene trees estimated using RAxML, we inferred the species tree under the multispecies coalescent using ASTRAL-II [24]. Subsequently, we visualized the discordance between the 13,164 gene trees and the species tree. At each node, the local posterior probabilities were computed with ASTRAL-II [24], while the relative frequencies of the three possible bipartitions (possible arrangements of a quartet on an unrooted tree) induced by each internal branch of the estimated species tree were computed and visualized using DiscoVista v1.0 [25] (Figure S2).

Phylogenomic dating

We inferred the evolutionary timescale using a Bayesian phylogenomic dating analysis of the concatenated exon sequences from 20 species (Table S1). The sequence data were analyzed using MCMCTree, part of the PAML package [16]. Rate variation among lineages was modeled using a separate correlated-rates relaxed clock for each of the three codon positions. A uniform substitution rate GTR+G model was assigned to each branch. The branch-site model was additionally used to further investigate potential positive selection. We specified Homotherium latidens as the foreground species and Panthera leo, Panthera pardus, Panthera onca, Panthera tigris tigris, Panthera uncia, Neofelis nebulosa, Acinonyx jubatus, Lynx pardinus, Felis catus, and Crocuta crocuta as background species and Hyaena hyaena as outgroup for the tests. We assessed the statistical significance of the likelihood ratio with a 10% false discovery rate criterion (Table S2). Genes with a free-ratio > 2 were considered as highly significant for signatures of positive selection (Table S3).
Genetic diversity

To estimate the genetic diversity of our Homotherium latidens genome, we used two independent approaches to calculate heterozygosity. These included the exome-wide heterozygosity, and autosome-wide heterozygosity. For comparative purposes, we calculated the same metrics from single representatives from each of the 15 cat species included in our study.

To calculate the exome-wide heterozygosity, we first extracted only the coding regions from the mapped bam files using SAMtools v1.6 [59]. We ran the resultant exome only bam files through ANGSD v0.921 [60] for each individual independently, sampling allele frequencies (-doSaf 1), taking genotype likelihoods into account using the GATK algorithm (-GL 2 and -domajorminor 1), and specifying the following parameters: only include sites with at least 5x read depth (-mininddepth 5), only include reads with a mapping quality > 30 (-minmapq 30), only consider bases with a quality > 30 (-minq 30), only consider reads mapping to one location (-unique-only 1), adjust quality scores around indels (-baq 2). Moreover, we only considered transversions (-notrans 1) for all species as aDNA damage is known to cause C-T transitions (G-A on the reverse strand) which could artificially inflate the diversity estimates in the Homotherium relative to the less damaged extant species. From this we calculated a folded SFS to recover the number of sites that are either homozygous or heterozygous in each individual using the realSFS command as part of the ANGSD package, with a tolerance of 1e-8. To calculate the variance of these results we included the -nSites parameter specifying 200kb windows of covered bases.

Following this, we calculated the autosome-wide patterns of heterozygosity of the same individuals, excluding coding sequences. Coding regions were removed from the mapped bam files using bedtools v2.26.0 intersect [52]. To make results comparable between individuals, and due to the relatively low coverage of our Homotherium latidens individual, we first calculated the average read depth across the genome excluding coding regions of all individuals and downsampling all individuals to an average read depth of ~7.75x using SAMtools. To remove scaffolds putatively linked to sex chromosomes, we calculated the average read depth of the downsampled bam files for each scaffold in two individuals known to be male (African lion, and cheetah). Scaffolds that had an average read depth < 4.5x were deemed as candidates for originating from either the X or Y chromosome so were removed from further analyses. This resulted in the cumulative removal of ~50Mb from the downstream analyses. We used the same ANGSD protocol as for the exome but reduced the minimum read depth from 5x to 3x due to the much lower average read depth of the whole genome data and restricted the analysis to scaffolds > 1MB. Furthermore, as opposed to the 200kb windows of covered bases, to obtain the variance of heterozygosity across the autosomes, we calculated the heterozygosity for each scaffold independently.