eTOC blurb

Xenikoudakis et al. report a partial mitochondrial genome of the extinct giant beaver *Castoroides* and estimate the origin of aquatic behavior in beavers to approximately 20 million years. This time estimate coincides with the extinction of terrestrial beavers and raises the question whether the two events had a common cause.

Ancient DNA reveals twenty million years of aquatic life in beavers

Georgios Xenikoudakis[1], Mayeesha Ahmed[2], Jacob Colt Harris[2], Rachel Wadleigh[2,3], Johanna L. A. Paijmans[1], Stefanie Hartmann[1], Axel Barlow[1], Heather Lerner[2], Michael Hofreiter[1]

[1]Institute for Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Straße 24–25, 14476 Potsdam, Germany

[2]Joseph Moore Museum, Earlham College, 801 National Road West, Richmond IN 47374 US

[3]University of Michigan, 1105 N University Ave, Ann Arbor, MI 48109 USA

Lead Contact: Heather Lerner

*Correspondence: hlerner@gmail.com, geoxenik@gmail.com

With approximately 30 recognised extinct genera, beavers were once a taxon-rich rodent group adapted to both terrestrial and aquatic habitats [1,2]. Today, only two morphologically similar species survive, the Eurasian and the North American beaver [3]. Both are known for their aquatic lifestyle as well as their woodcutting and engineering behaviour that allows them to alter the environment and affect sympatric species [4]. Palaeontological studies suggest that aquatic and woodcutting behaviours are derived traits shared only

between the extinct group of giant beavers and members of the extant beaver lineage [1,3,5]. Here we use 7,686 base-pairs (bp) of mitochondrial DNA from the extinct giant beaver *Castoroides ohioensis* to investigate when these novel behaviours originated in beavers. Our phylogenetic analysis retrieves the anticipated sister relationship of giant beavers to the extant beavers and places the time to their common ancestor during the early Miocene, approximately 20 million years ago. Our results are congruent with inferences from the fossil record [1] in suggesting a single evolutionary transition from terrestrial to aquatic life, although they place this event approximately four million years later in time compared to previous fossil studies [1]. Our study shows that ancient DNA can be used for studying behavioural shifts, even those happening many millions of years ago.

The giant beaver went extinct during the megafaunal extinction at the onset of the Holocene around twelve thousand years ago [5] and was one of the largest rodents that has ever lived, possibly reaching the size of the American black bear [6]. We combined DNA hybridisation capture and iterative mapping methods to retrieve around half of a giant beaver mitochondrial genome (7,686bp, see also Supplemental information, Figure S1). The retrieved sequences indicate advanced DNA fragmentation and nucleotide misincorporations typical for ancient DNA, supporting their authenticity

(Supplemental Information, Figure S2a, S2c). A pairwise comparison with potential sources of contamination suggested negligible incorporation of exogenous sequences during the giant beaver mitochondrial assembly (Supplemental information, Figure S2b). We then compared the assembled giant beaver partial mitochondrial genome against a dataset of rodent, lagomorph (hares and rabbits) and primate mitochondrial genomes to determine their phylogenetic affinities. Both maximum likelihood and Bayesian phylogenetic methods returned congruent results (Figure 1; See also Supplemental Information Figure S2d), strongly supporting a sister group relationship of the giant beaver to the extant beaver clade, in agreement with evidence from the fossil record [1,5] but in contrast with recent palaeoproteomic studies that suggested a closer phylogenetic relationship to *Ictidomys tridecemlineatus* (thirteen-lined ground squirrel) [7]. Using molecular dating, the mean time of divergence between giant and extant beavers is estimated to be 19.7 million yars ago (mya; 95% Credibility Interval: 17.4–22.2 mya) (Figure 1). Given that giant and extant beavers share an aquatic lifestyle [1], the most parsimonious scenario is that this derived trait evolved in their common ancestor. The estimated age of 19.7 mya thus represents a minimum age for the evolution of aquatic life in beavers.

Extant beavers possess unique adaptations to aquatic life. Examples are the

enlarged hind feet and a paddle-like tail used for propulsion in water [1,8] as well as a "combing-claw" used to maintain a non-wettable fur while swimming [8]. A primitive form of this specialised claw has already been identified in early Miocene (23-16 mya) fossils of *Steneofiber* [8], an extinct beaver genus distributed across Eurasia [5,9]. It is interesting that the earliest fossils assigned to *Steneofiber* are considered as progenitors of the common lineage of extant and extinct giant beavers during the Tertiary beaver radiation in Europe [9]. The age of *Steneofiber* fossils overlaps with our molecular estimate for the common ancestor of giant and modern beavers (Figure 1). This implies that the evolution of aquatic traits likely started with *Steneofiber* or its ancestor in Eurasia during the Early Miocene, prior to the diversification of modern and giant beavers around 20 mya, further supporting a single evolutionary change in beaver behaviour towards an aquatic lifestyle. This date coincides with the disappearance of terrestrial beavers, known only from North America, from the fossil record as well as with the appearance of aquatic giant beavers in this region [2]. Around the Oligocene-Miocene boundary around 23 mya ago, the first of a series of glaciations that lasted during the early Miocene (23-16 mya) took place [10]. It is possible that these climatic fluctuations have played a role in species turnover in aquatic and terrestrial beavers, but further study is needed to investigate this hypothesis.

In addition to an aquatic lifestyle, modern beavers also perform woodcutting

for both food and building material. Studies on isotopes [11] showed that *Castoroides* ate soft aquatic plants, rather than wood. These findings do not directly reject woodcutting for building purposes, although dam and lodge building (engineering) remains an open question for all extinct beaver species due to a lack of fossil evidence. The general similarity of extant beavers reasonably suggests woodcutting behaviour occurred in their common ancestor. In this case our 7my estimate (95% Credibility Interval: 6.7-8.7) (Figure 1) of their divergence serves as a minimum age for the evolution of woodcutting. The earliest evidence for woodcutting based on cut marks are assigned to the extinct beaver genus *Dipoides* and date to the Pliocene (4-5 mya) [12]. Evidence for extant beavers is even more recent, dating to the Pleistocene [3]. Assuming that *Dipoides* belongs to the giant beaver lineage [1,3], the most parsimonious explanation would be that woodcutting emerged prior to the divergence of the extinct giant and modern beavers during the early Miocene (Figure 1). This would suggest that fossil evidence underestimates the age of this behaviour. Alternatively, if *Dipoides* shares a more recent common ancestor with extant beavers, then woodcutting behaviour would date some time prior to the divergence of modern beavers 7 mya (Figure 1). A molecular phylogenetic study on *Dipoides* subfossils could shed light upon these alternative scenarios.

This study provides the first molecular evidence for a single evolutionary

change in the behavioural lifestyle of beavers from terrestrial to aquatic habitats. This change coincided closely with the extinction of terrestrial beavers in North America, suggesting that the two events may share a common, albeit as yet unknown, cause.

Declaration of Interests

The authors declare no competing interests.

Author Contribution

Conceptualization, H.L. and M.H.; Methodology, G.X., J.L.A.P., A.B., H.L and M.H.; Software, G.X., J.L.A.P., A.B., S.H. and H.L; Validation, G.X, M.A., R.W. and H.L.; Formal Analysis, G.X, M.A., R.W and H.L.; Investigation, G.X., M.A., J.C.H., J.L.A.P. and H.L.; Resources, M.A., J.C.H., H.L. and M.H.; Data Curation, G.X. and H.L.; Writing –Original Draft, G.X; Writing – Review & Editing, G.X., M.A, J.C.H., W.R., J.L.A.P, S.H., A.B., H.L. and M.H.; Visualization, G.X., A.B., H.L. and M.H; Supervision, H.L. and M.H.; Project Administration, H.L. and MH; Funding Acquisition, H.L. and M.H.;

References

1. Rybczynski, N.J. (2007). Castorid phylogenetics: Implications for the evolution of swimming and tree-exploitation in beavers. Journal of Mammalian Evolution *14*, 1-35.
2. Hugueney, M., and Escuillié, F. (1996). Fossil Evidence for the Origin of Behavioral Strategies in Early Miocene Castoridae, and Their Role in the Evolution of the Family. Paleobiology *22(4),* 507-513.
3. Rybczynski, N.J. (2008). Woodcutting behavior in beavers (Castoridae, Rodentia): estimating ecological performance in a modern and a fossil taxon. Paleobiology, *34(03)*,

389-402.

4. Hastings, A., Byers, J.E., Crooks, J.A., Cuddington, K., Jones, C.G., Lambrinos, J.G., Talley, T.S., and Wilson, W.G. (2007). Ecosystem engineering in space and time. Ecology Letters *10*, 153–164.
5. Korth, W.W. (2001). Comments on the systematics and classification of the beavers (Rodentia, Castoridae). Journal of Mammalian Evolution *8*, 279-296.
6. Reynolds, P.S. (2002). How big is a giant? The importance of method in estimating body size of extinct mammals. Journal of Mammalogy *83*, 321–332.
7. Cleland, T.P., Schroeter, E.R., Feranec, R.S., and Vashishth, D. (2016). Peptide sequences from the first *Castoroides ohioensis* skull and the utility of old museum collections for palaeoproteomics. Proceedings of the Royal Society B *283*, (1832).
8. Hugueney, M., and Escuillie, F. (1994). K-strategy and adaptive specialization in *Steneofiber* from Montaigu-le-Blin (dept. Allier, France; Lower Miocene, MN 2a, +23 Ma): first evidence of fossil life- history strategies in castorid rodents. Palaeogeography, Palaeoclimatology, Palaeoecology *113*, 217-225.
9. Stefen, C. (2011). A Brief Overview On The Evolution Of European Tertiary Beavers. Baltic Forestry *17(1)*, 148-153.
10. Billups, K., Channell, J.E.T., and Zachos, J. (2002). Late Oligocene to early Miocene geochronology and paleoceanography from the subantarctic South Atlantic. Paleoceanography *17(1)*, 1004.
11. Plint, T., Longstaffe, F.J., and Zazula, G. (2019). Giant beaver palaeoecology inferred from stable isotopes. Scientific Reports *9*, 71-79.
12. Tedford R.H., and Harington C.R. (2003). An Arctic mammal fauna from the Early Pliocene of North America. Nature *425,* 388–390.

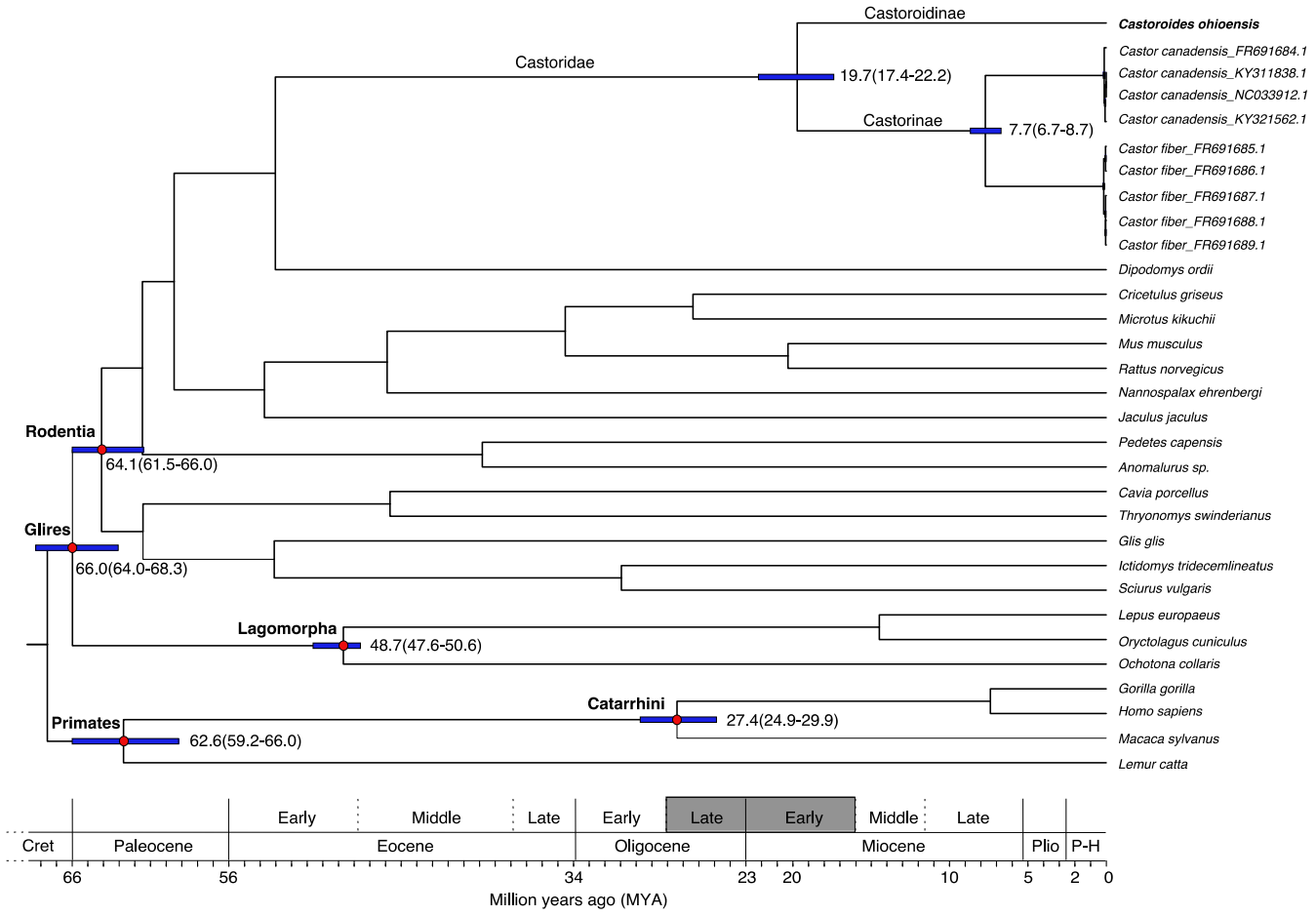

Figure 1. Maximum clade credibility tree. All nodes returned high posterior probabilities (>0.98). Mean estimates for divergences of Castoridinae-Castoroidinae and the European Beaver (*Castor fiber*) - North American beaver (*Castor canadensis*) clade, respectively, are presented in million years ago (mya) next to the corresponding node with node bars depicting the 95% highest posterior density (95% Credibility Interval, shown in parentheses) of the mean estimated age. Nodes marked with a red dot signify constraints based on dated fossils during Bayesian analysis. Scale shown is in million years along with the geological epochs (Cret: Cretaceous, Plio: Pliocene, P-H: Pleistocene-Holocene). The grey bar on the timescale represents the approximate stratigraphic range of *Steneofiber* fossils (adapted from [2]; see also [9] for an overview on beaver taxonomy). The time to the most recent common ancestor of extant and giant beavers is estimated to the early Miocene.

Supplemental Information: Ancient DNA reveals twenty million years of aquatic life in beavers

Georgios Xenikoudakis, Mayeesha Ahmed, Jacob Colt Harris, Rachel Wadleigh, Johanna L. A. Paijmans, Stefanie Hartmann, Axel Barlow, Heather Lerner, Michael Hofreiter

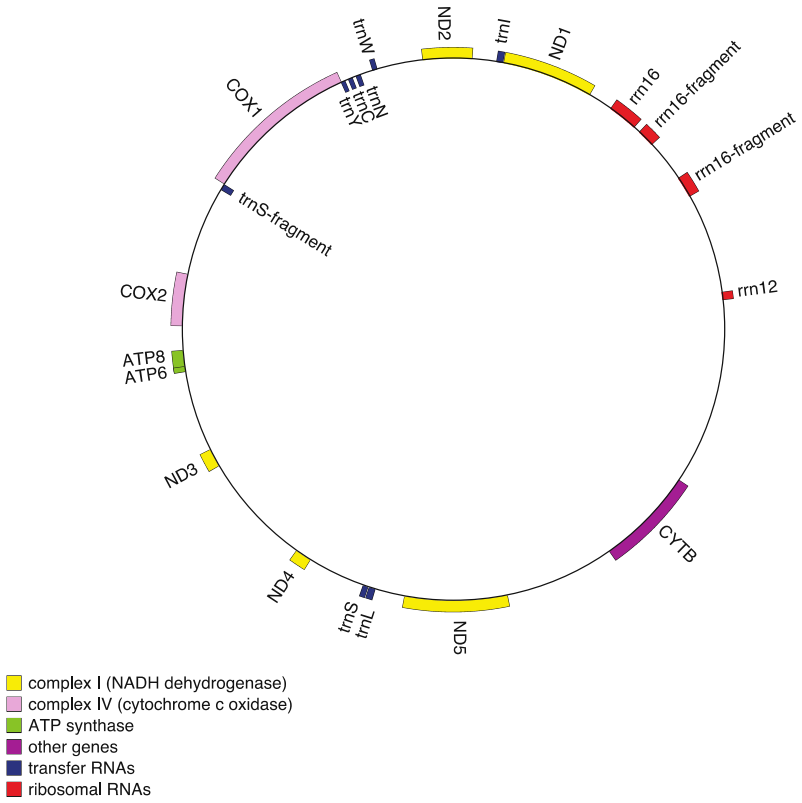

Figure S1. Gene map of the *Castoroides ohioensis* mitochondrial genome showing the parts of the genome that could be recovered. The map was generated with the software GeSeq [S27] using blat search tool settings and the North American beaver *Castor canadensis* (NC_033912.1) as reference for the annotation of the giant beaver mitogenome. Positions colored determine parts of the reference mitochondrial genome that are covered. Genes are positioned in the outer and inner side of the circle based on

clockwise (inner) and counterclockwise (outer) transcription direction on the L-strand and H-strand, respectively.

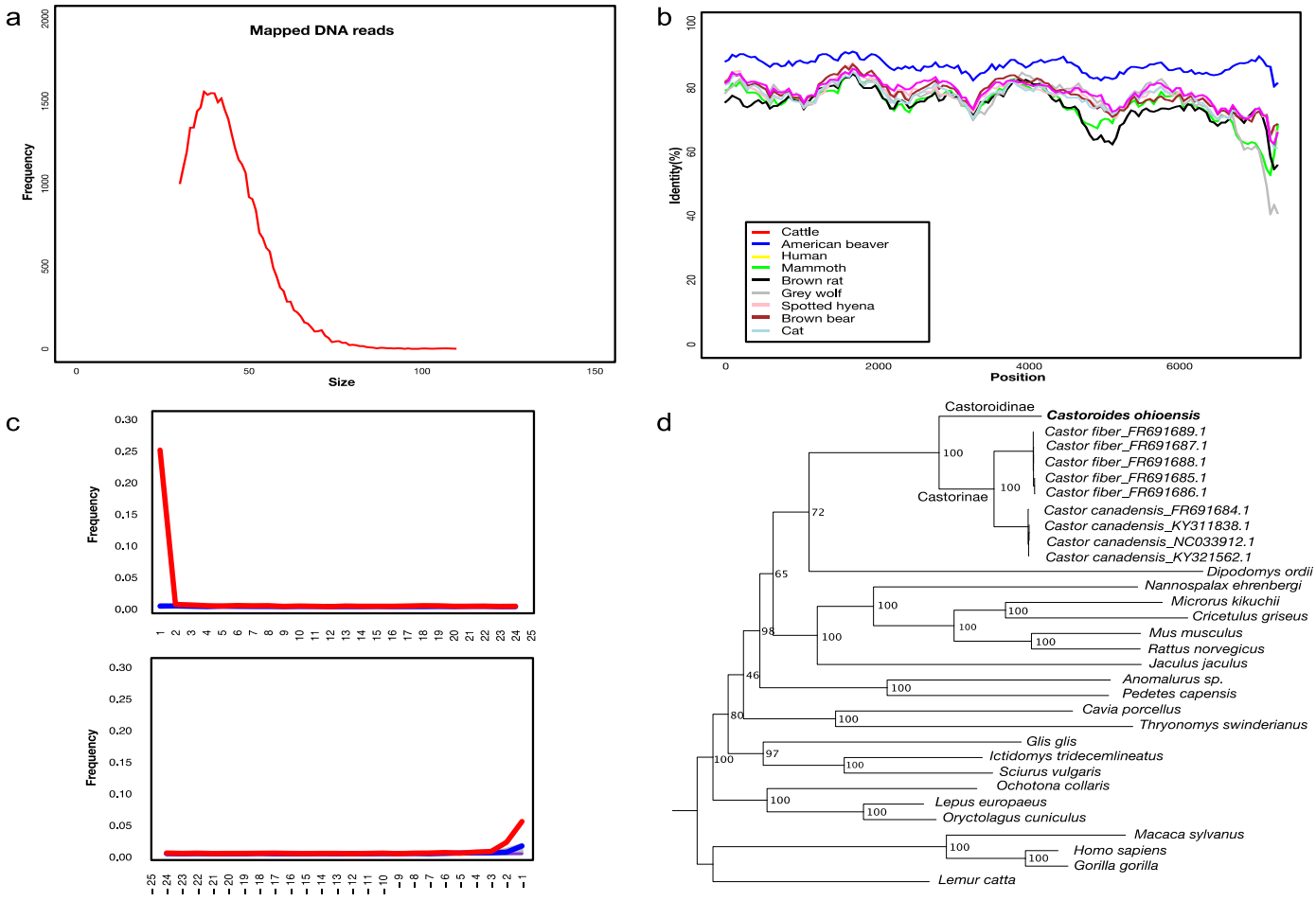

Figure S2. (a) Read length distribution of reads mapping to the final consensus sequence. The DNA fragment size distribution is in line with expectations for ancient DNA. (b) Pairwise comparison of the giant beaver (*Castoroides ohioensis*) final consensus sequence against potential contamination sources. Results are depicted as percentage (%) of identity (y-axis) across 500 bp overlapping sliding windows. Each window progresses by 50 bp and each line represents a single pairwise comparison. The giant beaver shows consistently higher pairwise DNA similarity to modern beaver than to all other species compared, in line with phylogenetic expectations. (c) Frequency (y-axis) of C to T substitutions (red) and G to A (blue) per position (x-axis) across mapped reads from the 5′ (upper) and the 3′ end (lower). The increased frequency of C to T transitions at the two

ends of the reads is indicative of DNA damage. (d) Maximum likelihood phylogeny using 100 bootstrap replicates to infer the phylogenetic position of *Castoroides ohioensis*. Nodes are annotated with bootstrap percentages.

Supplemental Experimental Procedures

Sampling, DNA extraction and library preparation

Five giant beaver teeth (incisors VP-45 and VP-13, and molars VP-73, VP-74 and VP-149) housed at Earlham College's Joseph Moore Museum collection were processed during this study. Growth in rodent teeth occurs by the secretion of mineralised dentin in odontoblasts that forms a coating around the inner lining of the pulp cavity [S1]. Therefore, we sampled by drilling incisors and molars at the dental pulp interior retrieving ~50mg of material. Sampling was performed in a dedicated sampling area for ancient DNA at the Joseph Moore Museum. All subsequent DNA extractions, library preparations and sequencing work were carried out at the University of Potsdam in a dedicated ancient DNA laboratory.

DNA extractions were performed for all samples following a published protocol [S2] with exception of the centrifugation speed setting during the binding stage, which was reduced (1500 RPM) [S2]. Subsequently, extracts were enzymatically treated with uracil DNA glycosylase and endonuclease VIII to excise uracils before being converted into single-stranded DNA libraries [S3]. Prior to indexing amplification, we determined the optimal number of indexing PCR cycles through qPCR (Thermo Scientific Pikoreal 96) as described previously [S4]. After library indexing, quantification was performed by measuring DNA fragment size on a Tapestation 2200 (Agilent Technologies) using the D1000 ScreenTape System and DNA concentration on a Qubit Fluorometer (ThermoFisher Scientific). During the extraction and library building stages, negative blanks were also included. All libraries were used for DNA enrichment prior to sequencing.

Hybridisation capture

Castor canadensis (North American beaver) tissue samples were retrieved from the Joseph Moore Museum tissue collection for generating capture baits (JMMNH G-130). A Qiagen DNeasy Blood and Tissue Kit was used to extract DNA. We amplified the *C. canadensis* mitochondrial DNA in two products of ~6,000 and ~11,500 bp using primers and protocol described in [S5]. Gel excision and purification was used to remove smaller products from non-specific amplification using a NucleoSpin Gel and PCR Clean-up kit (Macherey-22) according to the manufacturer's protocol. The purified PCR products were subsequently sheared to a length of ~150bp with a Covaris S220 System.

We applied a modified version of the protocol proposed for ancient DNA capture in [S6]. Briefly, streptavidin-coated Dynabeads were used to isolate the hybrid target-bait molecule complexes. We performed two successive capture experiments, the first using 24 hours hybridisation time, and the second 48 hours. After each capture, optimal amplification cycles for each product were estimated using qPCR prior to product amplification. After amplification, products were purified using a Qiagen MinElute PCR kit according to the manufacturer's instructions. Purified libraries were quantified as

described above for DNA fragment size and concentration on a Tapestation 2200 (Agilent Technologies) and Qubit Fluorometer (ThermoFisher Scientific), respectively. Quantification analysis showed that only three (samples VP-45, VP-73, VP-74) out of five sample libraries had successfully been enriched, e.g. showing traces of DNA on the Tapestation. These three libraries were subsequently used for sequencing.

Sequencing and computational analysis

We sequenced approximately 2 million paired-end reads (2x75 bp) for each of the three captured library samples along with a negative blank on an Illumina Miseq Platform. Library adapters were trimmed from paired-end sequencing reads with Skewer software (ver.0.2.2) [S7] using a minimum post-trimming read length of 30bp (-l 30). Trimmed paired reads were merged with Flash (ver. 1.2.11) [S8] using default parameters.

Preliminary computational analysis involved mapping reads to the modern North American beaver mitochondrial genome, which showed no detectable contamination in our negative blank. During this preliminary mapping, only one (VP-73) of the three libraries returned enough uniquely mapping reads (VP-73: 22326 reads; VP-45: 44 reads; VP-74: 153 reads) against the North American beaver mitochondrial genome and thus it was the only one used in the phylogenetic analysis.

Given the lack of prior genetic information and in order to avoid any reference genome related biases [S9], we used a collection of different mitochondrial reference genomes against which merged reads were mapped. The selection was based on the proposed rodent phylogenies inferred from DNA, protein or morphological data [S5,S10-S16]. The following mitochondrial sequences were downloaded from GenBank and used as references: North American Beaver (*C. canadensis*, FR691684.1), European Beaver (*Castor fiber*, NC_028625.1), African squirrel (*Anomalurus sp*. AM159537.1), Guinea Pig (*Cavia porcellus*, AJ222767.1), South African springhare (*Pedetes capensis*, HE983623.1), Brown Rat (*Rattus norvegicus*, KP244683.1) and Thirteen-lined ground squirrel (*Ictidomys tridecemlineatus*, KP698974.1).

In order to retrieve the maximally possible information for the giant beaver mitochondrial genome, we followed an iterative mapping approach. Using a custom perl script (available at: https://github.com/geoxenik/GiantB/blob/master/GB_iterative.pl), we performed 100 iterations against each of the references to ensure read mapping saturation, i.e. no additional reads mapping to the reference. For the first alignment, bwa (v.0.7.13) [S17] was run with mismatch parameter set to default (-n 0.04) to achieve higher read mapping confidence. During the following 100 iterations, a more relaxed parameter setting (-n 0.01) was allowed for extension of the aligned regions. All bwa runs were performed with disabled seeding length (-l 999) and reads of low mapping quality (-q 25) were removed with samtools (v.1.3) [S18]. PCR duplicates were discarded using the MarkDupsByStartEnd.jar script (available at: https:// github.com/dariober/Java-cafe/tree/master/MarkDupsByStartEnd). After each iteration, a consensus sequence was

called based on the mapped reads using ANGSD (ver.0.9.13) [S19]. Regions across the reference with low depth of coverage (<3X) and regions with no aligned reads were called as missing (N). Each consensus was used as a reference to re-map the entire set of merged reads during the next iteration. The consensus sequences called after the last iteration of each of the seven iterative mapping trials, each corresponding to the different references used for mapping, were aligned with MAFFT (v7.273) [S20]. From this alignment, we called a final consensus requiring 100% identity, ambiguous sites or non-covered sites were called as Ns and the new final consensus was used it as input for further phylogenetic analysis.

Consensus validation

We calculated the size distribution of mapped reads to check if it was consistent with the expected size for ancient DNA molecules (Figure S1a). We used mapDamage (v.2.0.8) [S21] to check for the characteristic damage patterns at the two ends of mapped reads created by increased level of cytosine to uracil substitutions during DNA degradation [S22] (Figure S1b). Finally, in order to identify ambiguous regions across the final consensus that would indicate incorporation of exogenous reads (contamination) during the iterative mapping approach, we performed a pairwise comparison of the consensus sequence with mitochondrial genomes of common contaminants as well as species that have been processed in our laboratory. The following species were used for the comparisons: cattle (*Bos taurus*, V00654.1), North American beaver (*Castor canadensis*, FR691684.1), human (*Homo sapiens*, J01415.2), grey wolf (*Canis lupus*, NC_009686.1), woolly mammoth (*Mammuthus primigenius*, JF912200.1), spotted hyena (*Crocuta crocuta*, NC_020670.1), brown bear (*Ursus arctos*, AP012559.1) domestic cat (*Felis catus*, NC_001700.1) and brown rat (*Rattus norvegicus*, NC_001665.2). Comparisons were performed by calculating the percentage of pairwise identity in 500bp overlapping sliding windows, with a 50 bp step size (Figure S1c).

Phylogenetic analysis

For the phylogenetic analysis, we used the following mammalian mitochondrial genomes: *Anomalurus* sp. (NC_009056.1), *Pedetes capensis* (HE983623.1), *Rattus norvegicus* (NC_001665.2), *Thryonomys swinderianus* (NC_002658.1), *Glis glis* (NC_001892.1), *Castor fiber* (FR691685.1), *C. fiber* (FR691686.1), *C. fiber* (FR691687.1), *C. fiber* (FR691688.1), *C. fiber* (FR691689.1), *C. canadensis* (FR691684.1), *C. canadensis* (NC_033912.1), *C. canadensis* (KY311838.1), *C. canadensis* (KY321562.1), *Ictidomys tridecemlineatus* (KP698974.1), *Dipodomys ordii* (KT327176.1), *Gorilla gorilla* (NC_001645.1), *Cavia porcellus* (NC_000884.1), *Cricetulus griseus* (NC_007936.1), *Lepus europaeus* (AJ421471.1), *Mus musculus* (NC_006915.1), *Homo sapiens* (J01415.2), *Jaculus jaculus* (NC_005314.1), *Lemur catta* (AJ421451.1), *Macaca sylvanus* (AJ309865.1), *Nannospalax ehrenbergi* (NC_005315.1), *Ochotona collaris* (AF348080.1), *Oryctolagus cuniculus* (AJ001588.1), *Sciurus vulgaris*

(NC_002369.1), *Microtus kikuchii* (NC_003041.1). We aligned these sequences with the giant beaver mitochondrial genome using MAFFT (v. 7.245) [S20] and setting a slow iterative refinement while keeping all other parameters to default.

Prior to phylogenetic analysis, we manually determined regions of tRNAs, rRNAs and coding genes across our alignment based on the mitochondrial sequence annotation. Repetitive regions and the D-loop region were removed due to ambiguous sequence alignment. Subsequently, we ran PartitionFinder2 [S23] to identify an appropriate set of data partitions and substitution models using a greedy search algorithm and the BIC criterion for model selection. Our analysis resulted in seven partitions (Partition 1, GTR+I+G; Partition 2: GTR+I+G; Partition 3: GTR+I+G; Partition 4, HKY+I+G; Partition 5, K80+G; Partition 6, GTR+I+G; Partition 7, GTR+I+G;) which were used for maximum likelihood and Bayesian phylogenetic analyses. The maximum likelihood phylogeny was estimated using RAxML-NG v. 0.8.1 [S24] using the primate sequences as outgroup for tree rooting. We ran 100 bootstrap replicates and visualised the results using FigTree v.1.4.3 (available at: http:// tree.bio.ed.ac.uk/software/figtree/). Subsequently, a fossil calibrated analysis was performed using BEAST v1.8.2 [S25]. We used a uniform prior distribution for the most recent common ancestor of Rodentia (56-66mya), Catarrhini (24.4-34mya), Primates (56-66mya), Lagomorpha (47.6-66mya) and Glires (56-164.6mya) [S26]. We applied a Birth-Death speciation tree model and a lognormal relaxed clock model for each genome partition, allowing for a mean rate of 0.05 substitutions per million years under uniform prior distribution (lower and upper values set to 0 and 0.2 respectively). We initiated our analysis with a random starting tree and ran a MCMC chain for 50 million steps sampling every 5000 steps. During preliminary runs, we encountered problems with parameter convergence for Partitions 1 and 2, suggesting overparameterisation. Thus, for those partitions, we switched to a simpler substitution model (HKY+I+G) than the one suggested (GTR+I+G) by the PartitionFinder 2 software**.** Parameter convergence and sufficient posterior sampling (Effective Sample Size above 200) were verified in Tracer v.1.6 (available at: http://tree.bio.ed.ac.uk/software/tracer/). The Maximum Clade Credibility tree was selected from the posterior sample with tree nodes scaled to their mean age using TreeAnnotator ver. 2.4.7 [S25], after removing the first 10 % of trees as burn in, and visualised in FigTree v.1.4.3. For confirmation, results of the time calibrated analysis were replicated by an independent run using the same settings in BEAST to check for consistent results among runs.

Supplemental References

[S1] Madden, R.H. (2014). Hypsodonty in mammals: Evolution, geomorphology and the role of earth surface processes. Cambridge University Press, Cambridge, UK.

[S2] Dabney, J., Knapp, M., Glocke, I., Gansauge, M.T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.L., and Meyer, M.(2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA

fragments. Proceedings of the National Academy of Sciences of the United States of America *110*, 15758-15763.

[S3] Gansauge, M.T., and Meyer M. (2013). Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. Nature Protocols *8*, 737-748.

[S4] Basler, N., Xenikoudakis, G., Westbury, M.V., Song, L., Sheng, G., and Barlow, A. (2017) Reduction of the contaminant fraction of DNA obtained from an ancient giant panda bone. BMC Research Notes *10,* 754.

[S5] Horn, S., Durka, W., Wolf, R., Ermala, A., Stubbe, A., Stubbe, M, and Hofreiter, M. (2011). Mitochondrial genomes reveal slow rates of molecular evolution and the timing of speciation in beavers (*Castor*), one of the largest rodent species. PLoS ONE *6*(1), e14622.

[S6] Horn, S. (2012) Target Enrichment via DNA Hybridization Capture. In: Shapiro B., Hofreiter M. (eds) Ancient DNA. Methods in Molecular Biology (Methods and Protocols), vol 840. Humana Press.

[S7] Jiang, H., Lei, R., Ding, S.W., and Zhu, S. (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics *15,* 182.

[S8] Magoc, T., and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics *27*(21), 2957–2963.

[S9] Westbury, M., Baleka, S., Barlow, A., Hartmann, S., Paijmans, J.L.A., Kramarz, A., Forasiepi, A.M., Bond, M., Gelfo, J.N., Reguero, M.A., Mendoza, P.L., Taglioretti, M., Scaglia, F., Rinderknecht, A., Jones, W., Mena, F., Billet, G., de Muizon, C., Aguilar, J.L., MacPhee, R.D.E., and Hofreiter, M. (2017). A mitogenomic timetree for Darwin's enigmatic South American mammal Macrauchenia patachonica. Nature Communications *8*, 15951.

[S10] Rybczynski, N.J. (2007). Castorid phylogenetics: Implications for the evolution of swimming and tree-exploitation in beavers. Journal of Mammalian Evolution *14*, 1-35.

[S11] Cleland, T.P., Schroeter, E.R., Feranec, R.S., and Vashishth, D. (2016). Peptide sequences from the first *Castoroides ohioensis* skull and the utility of old museum collections for palaeoproteomics. Proceedings of the Royal Society B *283*, (1832).

[S12] Fabre, P.H., Jønsson, K.A.J., and Douzery, E.J.P. (2013). Jumping and gliding rodents: Mitogenomic affinities of Pedetidae and Anomaluridae deduced from an RNA-Seq approach. Gene *531*,388–397.

[S13] Fabre, P.H., Hautier, L., Dimitrov, D., and Douzery, E.J.P. (2012). A glimpse on the pattern of rodent diversification: a phylogenetic approach. BMC Evolutionary Biology *12*, 88.

[S14] Ryu, S.H., Kwak, M.J., and Hwang, U.W. (2013). Complete mitochondrial genome of the Eurasian flying squirrel *Pteromys volans* (Sciuromorpha, Sciuridae) and revision of rodent phylogeny. Molecular biology reports *40*(2), 1917-1926.

[S15] Voloch, C., Vilela, J.F., Loss Oliveira, L., and Schrago, C.G. (2013). Phylogeny and chronology of the major lineages of New World hystricognath rodents: insights on the biogeography of the Eocene/Oligocene arrival of mammals in South America. BMC Research Notes *6*, 160.

[S16] Heritage, S., Fernández, D., Sallam, H.M., Cronin, D.T., Echube, M.E.E., and Seiffert, E.R. (2016). Ancient phylogenetic divergence of the enigmatic African rodent Zenkerella and the origin of anomalurid gliding. PeerJ *4*, e2320.

[S17] Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. Bioinformatics *25*, 1754-1756.

[S18] Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan J., Homer N., Marth G., Abecasis G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics *25*, 2078-2079.

[S19] Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics *15*, 356.

[S20] Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution *30*, 772–780.

[S21] Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P., and Orlando, L. (2013) mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics 29* (13): 1682–1684.

[S22] Dabney, J., Meyer, M., and Paabo, S. (2013) Ancient DNA damage. Cold Spring Harbor Perspectives in Biology *5*.

[S23] Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T.,and Calcott, B. (2016) PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Molecular Biology and Evolution. *34*, (3) 772–773,

[S24] Alexey M Kozlov, Diego Darriba, Tomáš Flouri, Benoit Morel, and Alexandros, Stamatakis, RAxML-NG: A fast, scalable, and user-friendly tool for maximum likelihood phylogenetic inference, Bioinformatics, btz305.

[S25] Drummond, A.J., Suchard, M.A., Xie, D. and Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology And Evolution *29*, 1969-1973.

[S26] Benton, M.J., Donoghue, P.C.J., Asher, R.J., Friedman, M., Near, T.J., and Vinther, J. (2015) Constraints on the timescale of animal evolutionary history. Palaeontologia Electronica 18.1.1FC, 1–107.

[S27] Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E.S., Fischer, A., Bock, R., and Greiner, S. GeSeq – versatile and accurate annotation of organelle genomes (2017). Nucleic Acids Research *45*, W6-W11.