Title
Genomic Data from Extinct North American Camelops Revise Camel Evolutionary History.

Permalink
https://escholarship.org/uc/item/4zm8b6kJ

Journal
Molecular biology and evolution, 32(9)

ISSN
0737-4038

Authors
Heintzman, Peter D
Zazula, Grant D
Cahill, James A
et al.

Publication Date
2015-09-01

DOI
10.1093/molbev/msv128

Peer reviewed
Genomic Data from Extinct North American *Camelops* Revise Camel Evolutionary History

Peter D. Heintzman,*1 Grant D. Zazula,2 James A. Cahill,1 Alberto V. Reyes,3 Ross D.E. MacPhee,4 and Beth Shapiro1,5

1Department of Ecology & Evolutionary Biology, University of California Santa Cruz
2Yukon Palaeontology Program, Department of Tourism & Culture, Government of Yukon, Whitehorse, YT, Canada
3Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, AB, Canada
4Department of Mammalogy, Division of Vertebrate Zoology, American Museum of Natural History, New York, NY
5UCSC Genomics Institute, University of California Santa Cruz

*Corresponding author: E-mail: peteheintzman@gmail.com.
Associate editor: Claudia Russo

Abstract

Recent advances in paleogenomic technologies have enabled an increasingly detailed understanding of the evolutionary relationships of now-extinct mammalian taxa. However, a number of enigmatic Quaternary species have never been characterized with molecular data, often because available fossils are rare or are found in environments that are not optimal for DNA preservation. Here, we analyze paleogenomic data extracted from bones attributed to the late Pleistocene western camel, *Camelops hesternus*, a species that was distributed across central and western North America until its extinction approximately 13,000 years ago. Despite a modal sequence length of only around 35 base pairs, we reconstructed high-coverage complete mitochondrial genomes and low-coverage partial nuclear genomes for each specimen. We find that *Camelops* is sister to African and Asian bactrian and dromedary camels, to the exclusion of South American camelids (llamas, guanacos, alpacas, and vicuñas). These results contradict previous morphology-based phylogenetic models for *Camelops*, which suggest instead a closer relationship between *Camelops* and the South American camelids. The molecular data imply a Late Miocene divergence of the *Camelops* clade from lineages that separately gave rise to the extant camels of Eurasia. Our results demonstrate the increasing capacity of modern paleogenomic methods to resolve evolutionary relationships among distantly related lineages.

Key words: paleogenomics, *Camelops*, Pleistocene, arctic, phylogeny, morphology, camel evolution.

Introduction

The western camel, *Camelops*, was the largest of the late Pleistocene North American camels, until its extinction some 13,000 calendar years ago (13 ka), just before the end of the Pleistocene Epoch (Kooyman et al. 2012; Waters et al. 2015). Although its geographic range fluctuated, *Camelops* was widely distributed throughout western North America, from the sub-tropics of Honduras to the arctic latitudes of Canada (fig. 1, supplementary table S1, Supplementary Material online; Webb 1965; Webb and Perrigo 1984; Pinsof 1998; Graham and Lundelius 2010). Across this range, *Camelops* lived in a variety of habitats (Webb 1965; Zazula et al. 2011; Waters et al. 2015), where it fed on a mixed herbivorous diet of leaves, shrubs, and grasses (Zazula et al. 2011 and references therein). *Camelops* is one of the few megafaunal taxa that was demonstrably preyed upon by early humans in North America and has figured prominently in discussions of late Pleistocene megafaunal extinction (Kooyman et al. 2012; Waters et al. 2015).

*Camelops* was among the last surviving of the North American camels. The family Camelidae originated in North America during the Middle Eocene, approximately 46–42 Ma (Honey et al. 1998). Camelidae was taxonomically diverse, comprising as many as 13 now-extinct genera during the Miocene (Honey et al. 1998). Paleontological and mitochondrial DNA evidence together suggest that two of its constituent major tribes diverged in North America during the late Early Miocene, about 17.5–16 Ma (supplementary fig. S1, Supplementary Material online; Honey et al. 1998; Hassanin et al. 2012). Today, one tribe, the Camelini, is represented by the Afroasian dromedary (*Camelus dromedarius*) and wild and domestic bactrian camels (*C. ferus, C. bactrianus*, respectively), the presumed common ancestor of which first dispersed across the then-exposed Bering Isthmus into Eurasia during the Late Miocene, 7.5–6.5 Ma (Pickford et al. 1995; Rybczynski et al. 2013). The other major tribe, the South American Aucheniini, consists today of llamas (*Lama glama*), guanacos, (*L. guanicoe*), and vicuñas/alpacas (*Vicugna vicugna*) (Grubb 2005). The ancestor of these South American camelids probably dispersed into South America across the Isthmus of Panama during the Great American Biotic Interchange, perhaps just before the Pliocene/Pleistocene boundary around 2.7 Ma (Bell et al. 2004).

Satisfactory resolution of fossil camelid taxonomy and phylogeny has been impeded by rampant morphological
parallelism, with virtually every “diagnostic” character being encountered in more than one proposed lineage (Harrison 1979; see supplementary information and supplementary fig. S1, Supplementary Material online, for an overview). Based on morphology, *Camelops* is widely recognized as a highly derived member of the Aucheniini, probably most closely related to Late Miocene *Alforjas* (~10 to 5 Ma; Honey et al. 1998; see also Harrison 1979; Breyer 1983; Voorhies and Corner 1986; Webb and Meachen 2004, but see Scherer 2013).

It has become increasingly routine to use molecular data to test morphology-based hypotheses concerning evolutionary relationships. In particular, ancient DNA (aDNA) and fossil proteins offer the possibility of testing phylogenetic hypotheses involving both living and extinct species (e.g., Orlando et al. 2003; Rohland et al. 2007; Campos et al. 2010; Llamas et al. 2015; Welker et al. 2015). Here, we evaluate three *Camelops* fossils, all lower limb bones recovered from a late Pleistocene, presumably Sangamonian Interglacial (~125 to 75 ka), context in subarctic Yukon, Canada. Fossils from interglacial deposits tend to be rare at higher latitudes, at least partly because the generally acidic nature of interglacial soils is not conducive to long-term preservation of bones or biomolecules (Lindahl 1993; Muhs et al. 2001). Consequently, there have been few successful attempts to recover aDNA from fossils of subarctic and arctic interglacial faunas (but see Rohland et al. 2007) despite this interval being well within the current temporal envelope for successful aDNA characterization (Orlando et al. 2013). Using a combination of morphological and paleogenomic analyses, we demonstrate that 1) these fossils can be assigned to *Camelops* with confidence on the basis of morphological and mensurational comparisons and that 2) *Camelops* is phylogenetically closer to *Camelus* (Camelini) than *Lama/Vicugna* (Aucheniini). Divergence dating indicates that the lineages of *Camelops* and *Camelus* separated during the Middle to Late Miocene.

### Systematic Paleontology

**Classification**

Class: Mammalia; Order: Artiodactyla; Family: Camelidae; Subfamily: Camelinae; Genus: *Camelops* Leidy 1854.

**Fossil Locality and Age**

Three specimens of *Camelops* (YG 29.199, 328.21, and 328.23) were recovered from an active placer gold mine along Hunker Creek (64.019167 N, 139.158056 W), a waterway located near Dawson City, Yukon, Canada (fig. 1). Pleistocene vertebrate fossils are commonly uncovered at these localities by hydraulic stripping of frozen sediments, during mining operations designed to gain access underlying gold-bearing gravel (Froese et al. 2009). Mining in this manner does not proceed stratigraphically, with the result that there is usually only limited information for placing individual fossils in a precise chronological context. However, the fossil assemblage that yielded these *Camelops* fossils also includes steppe-bison (*Bison priscus*), considered to be chronologically restricted to the late Pleistocene.
Phylogenetic Position of Camelops within Camelinea

Analysis of the full mitochondrial and low coverage nuclear genomes (figs. 2 and 3, Supplementary figs. S3 and S4 and tables S2 and S3, Supplementary Material online) of the three Yukon Camelops specimens establishes unequivocally that they are more closely related to African and Asian Camelus (Camelini) than to South American Lama and Vicugna (Auchenini). This contrasts with the conventional interpretation of placing Camelops within the Auchenini (e.g. Harrison 1979, 1985; Voorhies and Corner 1986; Wheeler 1995; Honey et al. 1998; Scherer 2013) and suggests that a wider systematic revision of fossil camels is required.

In the mitogenomic phylogeny, the Camelops–Camelus clade is recovered with strong statistical support regardless of analytical parameters (fig. 2 and supplementary figs. S3 and S4, Supplementary Material online), with 100% posterior probability and 98–100% maximum likelihood bootstrap support in the presence of an outgroup and 94% posterior probability support in analyses lacking an outgroup (supplementary figs. S3 and S4 and table S2, Supplementary Material online). We estimated nuclear phylogenies using pairwise transversion distances to minimize the influence of damaged sites in the low-coverage paleogenomic dataset. The branching order of the resulting phylogeny was concordant with that of the mitochondrial phylogeny, regardless of whether the wild bactrian camel or alpaca was used as a reference genome for mapping the Camelops reads (fig. 3; supplementary table S3, Supplementary Material online). More Camelops reads aligned to the wild bactrian camel genome than to the alpaca genome per megabase (table 1), again suggesting that Camelops and Camelus genomes are more closely related to each other than either is to Lama Vicugna. We also note that the branches leading to Camelops were shorter when the Camelops reads were mapped to the alpaca genome than when mapped to the wild bactrian camel genome (fig. 3b). This is likely due to biased recovery of increasingly conserved regions of the genome as evolutionary distance increases between the reference genome and Camelops (Prüfer et al. 2010).

The mitochondrial and nuclear genome analyses are in slight disagreement with respect to resolved relationships...
among the three Camelops specimens, with either YG 328.23 or YG 328.21 standing as sister to the remaining specimens, respectively (fig. 2 and supplementary figs. S3 and S4 and table S3, Supplementary Material online). This observation is consistent with a lack of lineage sorting, due to the three specimens probably having originated from the same population.

**Dating the Divergence between Camelops and Camelus**

Molecular clock-based analyses of the mitogenomic data from the Yukon fossils indicate that Camelops diverged from the lineage of Old World Camelus between 17.5 and 7 Ma (fig. 4; supplementary table S2, Supplementary Material online). This range spans the late Early Miocene through the middle Late Miocene (Gradstein et al. 2012). Crucially, analyses assuming a wide range of model parameters resulted in broadly congruent results, with overlapping 95% credibility intervals. We note that the analyses that assumed a birth–death (BD) serially sampled prior resulted in noticeably younger estimates than models assuming alternate speciation priors and that the choice of outgroup also had an observable effect on the divergence estimate when calibration set one was enforced (fig. 4; supplementary table S2, Supplementary Material online). Our divergence estimates between living camel clades (supplementary table S2, Supplementary Material online) are consistent with previous ages derived from analyses of whole genomes (Wu et al. 2014) and mitochondrial sequences (Ji et al. 2009; but see supplementary information, Supplementary Material online) from extant...
species and are in broad agreement with the fossil record (Wheeler 1995).

To further test the Camelops–Camelus divergence estimate, we performed an additional analysis using the low coverage nuclear genome data. Assuming a genome-wide strict molecular clock, we inferred that these two lineages diverged between around 11 and 10 Ma (supplementary table S3, Supplementary Material online). This result is consistent with that derived from the mitogenomic data (fig. 4; supplementary table S2, Supplementary Material online).

Reconstructing the Evolutionary History of Camelidae

Both molecular and fossil record data suggest that crown group Camelidae arose in the late Early Miocene, around 17.5–16 Ma (supplementary fig. S1, Supplementary Material online; Honey et al. 1998; Hassanin et al. 2012; Wu et al. 2014). Our data support a subsequent divergence between the extant Old World camel (Camelus) lineage and the extinct Camelops clade prior to the late Late Miocene. This proposed divergence timing is consistent with hypotheses based on the fossil record that propose close affinities between Late Miocene Paracamelus and living Camelus (Titov 2008) and between Miocene Alforjas and Camelops (supplementary fig. S1, Supplementary Material online; Harrison 1979). Paracamelus, which is the presumed ancestor of living Old World camels, expanded from North America into Eurasia by approximately 7.5 to 6.5 Ma (Pickford et al. 1995). Eurasian Paracamelus later became separated from contemporaneous populations in arctic North America, probably as a result of the flooding of the Bering Isthmus approximately 5.5 Ma (Gladenkov et al. 2002). Paracamelus finally became extinct in the North American Arctic and Subarctic by the middle Pleistocene, roughly 1 Ma (Rybczynski et al. 2013). Stratigraphic and radiocarbon evidence suggest that Camelops and Paracamelus did not coexist in arctic and subarctic North America (Zazula et al. 2011; Rybczynski et al. 2013).

The fossil record suggests that Camelops first appeared during the middle Pliocene (~4.0 to 3.2 Ma) in southern North America (Thompson and White 2004) and achieved its maximum range during the late Pleistocene, with a reconstructed distribution from Alaska to Central America (fig. 1; supplementary table S1, Supplementary Material online). Fossils of C. hesternus in the northern part of its distribution in eastern Beringia (unglaciated Alaska and Yukon; fig. 1i) are very rare (Harington 1997), which may indicate either that populations were continuously present but always small or that Camelops dispersed into the subarctic only during brief interglacial intervals (Zazula et al. 2011). From a paleoecological standpoint, being able to distinguish between these alternatives is crucial, as the former implies a much greater capacity to withstand major climate change than does the latter.

Although Camelops was extinct throughout North America by roughly 13 ka (Waters et al. 2015), our nonfinite radiocarbon dates from Hunker Creek suggest that populations in eastern Beringia may have been locally extinct tens of millennia earlier, in parallel with the extinction chronology of American mastodon Mammut americanum (Zazula et al. 2014). This pattern may also apply to other rare fossil taxa found in eastern Beringia, such as Jefferson’s ground sloth (Megalonyx jeffersonii) and the giant beaver (Castoroides...
2008). We then combined the resulting alignments to produce the mitochondrial genome. We merged, adapter trimmed, and mapped each read to four extant camelid reference mitogenomes using an iterative assembler (Green et al. 2010). We restricted our analyses to transversions so as to prevent cytosine deamination from biasing analyses (Dabney, Meyer, et al. 2013).

Mitogenomic Reconstruction and Phylogenetic Analysis
We combined the sequenced reads from all samples and sequencing runs and used these to reconstruct the Camelops mitochondrial genome. We merged, adapter trimmed, and mapped each read to four extant camelid reference mitogenomes using an iterative assembler (Green et al. 2008). We then combined the resulting alignments to produce a draft of the Camelops mitogenome. Independent remapping of reads from each sample to the draft mitogenome, as well as visual inspection, gene annotation, and assessment of DNA fragment length distributions and deamination patterns (supplementary fig. S2, Supplementary Material online), were all consistent with what is expected from authentic aDNA sequences.

We aligned the three Yukon Camelops mitogenomes to seven extant camelid and ten other artiodactyl mitochondrial genomes collected from GenBank (supplementary table S4, Supplementary Material online). We partitioned the alignments into four partitions: control region, coding regions, rRNAs, and tRNAs. We then used PartitionFinder (Lanfear et al. 2012) and jModelTest (Darriba et al. 2012) to identify the most suitable model of molecular evolution for each partition (supplementary table S5, Supplementary Material online). Assuming these models, we then ran Bayesian and maximum likelihood phylogenetic analyses in MrBayes (Ronquist et al. 2012) and RAxML (Stamatakis 2014), respectively, to test for consistent camelid topologies under different combinations of three conditions: 1) including or excluding an outgroup taxon; 2) including or excluding the control region partition; and 3) using both a rate heterogeneity parameter (gamma) and a proportion of invariant sites or using only the rate heterogeneity parameter.

Materials and Methods
This section provides an overview of the methods of this study; full details can be found in the supplementary information, Supplementary Material online.

Radiocarbon Dating
To generate radiocarbon dates for the fossil Camelops specimens, we extracted approximately 150 mg of collagen from each bone using a combination of standard Longin (1971) methods and ultrafiltration (Beaumont et al. 2010). The prepared collagen was radiocarbon dated at the KECK Accelerator Mass Spectrometry (AMS) Laboratory (University of California, UC, Irvine).

DNA Extraction, Library Preparation, and Sequencing
We followed standard protocols for aDNA research, outlined in Cooper and Poinar (2000). In a purpose-built aDNA facility, we extracted DNA from 100 to 120 mg of bone powder following Dabney, Knapp, et al. (2013) and constructed DNA libraries following Meyer and Kircher (2010). We sequenced the libraries using the Illumina HiSeq-2500 platform at UCSF.

Phylogenetic Inference from Nuclear Genomes
To assess the phylogenetic relationships among Yukon Camelops and extant species in the Camelinina and Auchenini based on low coverage nuclear genomes, we calculated mean pairwise transversion differences between all pairs of Camelops specimens and the wild bactrian camel (Camelini) and alpaca (Auchenini) reference genomes. To test for mapping bias that might lead to underestimation of the divergence between Camelops and the reference genome, we mapped all Camelops shotgun reads to both the wild bactrian camel and alpaca reference genomes. To calculate the divergence between the reference genome sequences, we created artificial datasets in which short fragments mimicking Illumina sequencing reads were sampled from each of the two reference genomes. These artificial “reads” were then remapped to the alternate reference genome. In inferring pairwise distances between genomes, we restricted our analyses to transversions so as to prevent cytosine deamination from biasing analyses (Dabney, Meyer, et al. 2013).
the ratio of transversions between wild bactrian camel-Camelops and alpaca-Camelops (supplementary table S3, Supplementary Material online).

**Supplementary Material**

Supplementary information, figures S1–S4, tables S1–S5, and dataset S1 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

**Acknowledgments**

We thank John Alton and other placer gold miners of the Klondike region for collecting fossils and providing access to study sites, as well as Duane Froese, Elizabeth Hall, and Susan Hewitson for their help in collecting, identifying, and archiving the remains studied here. John Southon and the KECK radiocarbon laboratory provided radiocarbon dates, Julie Meachen provided access to unpublished *Palaeolama* measurement data, Joaquin Arroyo-Cabrales provided data on Mexican localities of Camelops, Robert Calef assisted with the annotation of the mitochondrial genomes, and Lily Shiue assisted with the experimental work. We thank three anonymous reviewers for comments that improved the manuscript. This work was supported by grants from the Packard Foundation and the Gordon and Betty Moore Foundation. Fossil specimens YG 29.199, 328.21, and 328.23 are curated as part of the Government of Yukon Palaeontology Program's fossil collections in Whitehorse, Yukon. The three Camelops mitogenomes have been deposited in GenBank (accession numbers: KR822420–KR822422). Whole-genome shotgun nucleotide data have been deposited in the NCBI Short Read Archive (BioProject accession number: PRJNA284284).

**References**


