Journal Pre-proof

Evidence for Pleistocene gene flow through the ice-free corridor from extinct horses and camels from Natural Trap Cave, Wyoming

Kieren J. Mitchell, Pere Bover, Alexander T. Salis, Caitlin Mudge, Holly Heiniger, Mary Thompson, Bryan Hockett, Laura S. Weyrich, Alan Cooper, Julie A. Meachen

Please cite this article as: Mitchell, K.J., Bover, P., Salis, A.T., Mudge, C., Heiniger, H., Thompson, M., Hockett, B., Weyrich, L.S., Cooper, A., Meachen, J.A., Evidence for Pleistocene gene flow through the ice-free corridor from extinct horses and camels from Natural Trap Cave, Wyoming, *Quaternary International* (2021), doi: https://doi.org/10.1016/j.quaint.2021.11.017.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1   **Evidence for Pleistocene gene flow through the ice-free corridor from extinct horses and**
2   **camels from Natural Trap Cave, Wyoming**
3
4
5   Kieren J. Mitchell[a,b,*], Pere Bover[c], Alexander T. Salis[a,d], Caitlin Mudge[a], Holly Heiniger[a], Mary
6   Thompson[e], Bryan Hockett[f], Laura S. Weyrich[a,g,h], Alan Cooper[i,j], Julie A. Meachen[k]
7
8
9   [a] Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Adelaide, South Australia,
10 Australia
11
12 [b] Otago Palaeogenetics Laboratory, Department of Zoology, University of Otago, Dunedin, New Zealand
13
14 [c] ARAID Foundation, Instituto Universitario de Investigación en Ciencias Ambientales (IUCA) - Aragosaurus Group,
15 Universidad de Zaragoza, Zaragoza, Spain
16
17 [d] Division of Vertebrate Zoology, American Museum of Natural History, New York, New York, USA
18
19 [e] Idaho Museum of Natural History, Idaho State University, Idaho, USA
20
21 [f] US Department of Interior, Bureau of Land Management, Nevada State Office, Reno, Nevada, USA
22
23 [g] Department of Anthropology, Pennsylvania State University, Pennsylvania, USA
24
25 [h] Huck Institutes of the Life Sciences, Pennsylvania State University, Pennsylvania, USA
26
27 [i] South Australian Museum, Adelaide, South Australia, Australia
28
29 [j] BlueSky Genetics, Ashton, South Australia, Australia
30
31 [k] Des Moines University, Department of Anatomy, Des Moines, Iowa, USA
32
33 * Corresponding author.
34
35 E-mail address: kieren.j.mitchell@gmail.com (K.J. Mitchell).

Abstract

Natural Trap Cave (Bighorn Mountains, Wyoming) preserves an abundance of fossil remains from extinct Late Pleistocene fauna and is situated near a past migration route that likely connected populations in Eastern Beringia and the contiguous US—the ice-free corridor between the Cordilleran and Laurentide icesheets. Some palaeontological evidence supports a correspondingly high affinity between fauna recorded in Natural Trap Cave and Eastern Beringia versus elsewhere in the contiguous US, but this hypothesis has not yet been extensively tested using genetic data. In the present study, we analysed 16 horse specimens and one camel specimen from Natural Trap Cave. Of the horse specimens we analysed, we obtained 10 unique and previously unreported mitochondrial haplotypes belonging to two distinct (extinct) genetic clades—two haplotypes corresponded to a caballine horse (*Equus* sp.) and eight corresponded to the stilt-legged horse (*Haringtonhippus francisci*). With only one exception, it appears these newly sequenced individuals all shared a common ancestor more recently with Eastern Beringian individuals than with others from the contiguous US. In addition, mitochondrial data from a specimen assigned to *Camelops* sp. revealed that it shares a closer affinity with specimens from the Yukon Territory than those from Idaho or Nevada, though all appear to belong to a single species ("yesterday's camel"; *Camelops* cf. *hesternus*). Together, these results are consistent with a high level of genetic connectivity between horse and camel populations in the Bighorn Mountains and Eastern Beringia during the Pleistocene.

Keywords

1. Introduction

Throughout the Pleistocene, glacial cycles caused the periodic expansion and contraction of the Cordilleran and Laurentide continental icesheets in the east and west, respectively, of northern North America. During glacial maxima, these two icesheets expanded to the point that they coalesced, likely limiting the dispersal of terrestrial mammals between ice-free areas in Eastern Beringia (Alaska and north-east Canada) and the southern interior of North America (including the modern day contiguous USA). Outside of glacial maxima, an ice-free corridor of varying extent connected these areas and would presumably have permitted faunal dispersal and gene flow. Indeed, fossil data suggest this was the case; for example, morphologically distinct Beringian wolves may have migrated southwards through the ice-free corridor during the Pleistocene prior to the Last Glacial Maximum (Meachen et al., 2016). However, the extent to which the periodic opening and closing of the ice-free corridor influenced the distribution and population structure of

74   Pleistocene fauna, particularly extinct megafauna, has not been extensively tested using genetic
75   data. Ancient DNA is particularly advantageous for studying this phenomenon, because the
76   ancestry and affinities of ancient individuals—including those from extinct species—can be
77   observed directly rather than inferred. However, application of this approach has been limited by
78   relatively poor DNA preservation in ancient specimens from temperate localities in the contiguous
79   US.

80

81   Pleistocene fossils excavated from Natural Trap Cave provide an excellent opportunity to test for
82   past gene flow between animal populations in Eastern Beringia and the North American southern
83   interior (i.e. the contiguous US). Firstly, Natural Trap Cave is situated in the northern Bighorn
84   Mountains, Wyoming, south of the maximum extent of the Pleistocene ice-sheets and very close to
85   the southern terminus of the ice-free corridor (Figure 1). Secondly, fossil remains belonging to a
86   wide range of animal species have been recovered from three distinct periods of deposition—155
87   to 132 thousand years ago, 53 to 17 thousand years ago, and 11 thousand years ago to the
88   present (Lovelace et al., This issue)—spanning multiple glacial cycles. Finally, previous studies
89   have indicated that ancient DNA can successfully be obtained from fossil remains excavated from
90   Natural Trap Cave (e.g. Barnett et al., 2005; Bover et al., 2018; Heintzman et al., 2016; Heintzman
91   et al., 2017; Orlando et al., 2008; Perri et al., 2021; Salis et al., 2020; Salis et al., 2021; Vershinina
92   et al., 2021). Consequently, if dispersal occurred through the ice-free corridor and resulted in gene
93   flow between populations in Eastern Beringia and the contiguous US, this is likely to be reflected in
94   the ancestry of ancient specimens from Natural Trap Cave. Specifically, we might expect to
95   observe a closer affinity between specimens from Natural Trap Cave and Eastern Beringia, as
96   opposed to those from populations further from the southern terminus of the ice-free corridor. We
97   would also expect shared ancestry between specimens from Natural Trap Cave and Eastern
98   Beringia to date to periods when the ice-free corridor would have been traversable—interglacials
99   (e.g. Marine Isotope Stage 3 [MIS 3]).

100

101   Ancient DNA data from North American bison (*Bison* sp.), including specimens from Natural Trap
102   Cave, indicate that bison dispersal through the ice-free corridor likely occurred bi-directionally
103   across multiple glacial cycles (Heintzman et al., 2016). In contrast, the contiguous US appears to
104   have been colonised by brown bears and lions from Eastern Beringia only once, with no evidence
105   for subsequent gene flow (Salis et al., 2020). However, the evidence is less clear-cut either way for
106   other taxa. For example, genetic data have been obtained from the extinct musk-oxen *Bootherium*
107   *bombifrons*—including specimens from Eastern Beringia, Natural Trap Cave, and as far south as
108   Nebraska—but sampling was too sparse to establish a detailed picture of past dispersal and
109   relatedness through time and space (Bover et al., 2018). Similarly, genetic data from caballine
110   (*Equus*) and stilt-legged (*Haringtonhippus*) horses have been obtained from Natural Trap Cave
111   and other southern localities (e.g. Heintzman et al., 2017; Vershinina et al., 2021; Weinstock et al.,

112 2005), but these data are also relatively sparse—complicating conclusions about species-level
113 diversity—and have not specifically been examined in the context of dispersal facilitated by the ice-
114 free corridor.

115

116 Both lineages of horse represented at Natural Trap Cave became extinct at the end of the
117 Pleistocene—after the Last Glacial Maximum but prior to the beginning of the Holocene
118 (Heintzman et al., 2017; Lorenzen et al., 2011; Vershinina et al., 2021). As a result, they have no
119 direct modern descendants from which past gene flow can be inferred. Only ancient DNA from
120 temporally and geographically distributed fossil specimens can reveal detailed patterns of
121 population structure and gene flow for these taxa. Similarly, a genus of endemic North American
122 camels ("yesterday's camel"; *Camelops*) became extinct at the end of the Pleistocene (Kooyman et
123 al., 2012; Waters et al., 2015), but it remains unclear exactly how many species this genus
124 comprised (Baskin and Thomas, 2016). Genetic data from three *Camelops* individuals from
125 Eastern Beringia have tentatively been referred to *Camelops hesternus* (Heintzman et al., 2015),
126 but the genetic identity and ancestry of more southern populations currently remains untested.
127 Since all these taxa—*Equus*, *Haringtonhippus*, and *Camelops*—were distributed widely across
128 North American prior to their extinction, including both Eastern Beringia and the contiguous US,
129 they are good models for exploring patterns of past migration through the ice-free corridor.

130

131 In this study, we present new mitochondrial genome sequences obtained from 16 horse specimens
132 and one camel specimen excavated from Late Pleistocene deposits in Natural Trap Cave.
133 Because published genetic data for yesterday's camels are otherwise only available from three
134 Eastern Beringian individuals, we also sequenced mitochondrial genomes from two additional
135 specimens for comparison: one from Spider Cave in Idaho and one from Mineral Hill Cave in
136 Nevada. One additional mitochondrial genome was also obtained from a horse specimen from
137 American Falls Reservoir, Idaho. We used these data to better characterise the phylogenetic
138 affinities of horses and camels from Natural Trap Cave with respect to both those from Eastern
139 Beringia and from populations in the contiguous US further from the southern terminus of the ice-
140 free corridor.

141

142 **2. Material and Methods**

143

144 *2.1 Samples*

145

146 In this study we analysed DNA from a total of 20 fossil specimens, 17 of which were from Natural
147 Trap Cave. Table S1 lists details for all specimens, including provenance, museum accession
148 numbers, and three newly reported radiocarbon ages. All new and previously published
149 radiocarbon ages were calibrated using OxCal v4.4 (Bronk Ramsey, 2016) based on the IntCal20

4

150 calibration curve (Reimer et al., 2020). All pre-PCR genetic research undertaken as part of this

151 study was conducted in the purpose-built ancient DNA clean-room facilities at the University of

152 Adelaide's Australian Centre for Ancient DNA (ACAD).

153

154 *2.2 DNA extraction*

155

156 To reduce contamination, each sample was UV irradiated for 15 min and then the surface layer

157 was abraded using a Dremel tool with a carborundum cutting disc. Each sample was subsequently

158 powdered using a Mikrodismembrator (Sartorius) or fragmented using a BioPulversiser (BioSpec),

159 and 20-200 mg was transferred to a 2 mL screw-cap tube. 1 mL of 0.5 M EDTA was added, and

160 the sample was incubated at room temperature on a rotary mixer for 30-45 min. Samples were

161 then centrifuged, and the EDTA was removed from the screw-cap tube and discarded. An

162 additional 970 μL 0.5 M EDTA and 30 μL 20 mg/mL Proteinase-K were added, and the sample

163 was incubated on a rotary mixer overnight for ~24 hr at 55 °C. The DNA released by these

164 digestion steps was bound and purified using a modified version of a previously published method

165 (Dabney et al., 2013), involving a binding step with a buffer comprising 12.6 mL PB buffer

166 (QIAGEN), 6.5 μL Tween-20, and 390 μL NaOAc 3M with in-solution silicon dioxide, followed by

167 two washes with 80% ethanol. Purified DNA was eluted in 200 μL of TE buffer (10 mM Tris, 1 mM

168 EDTA) with 0.05% Tween-20. Negative (no template) controls were included in each batch of

169 samples to monitor background and cross-contamination.

170

171 *2.3 Library preparation, enrichment, DNA sequencing*

172

173 Illumina DNA sequencing libraries were made from our extracted DNA and negative controls

174 following the protocol of Meyer and Kircher (2010), but using Rohland et al.'s (2015) partial uracil-

175 DNA-glycosylase (UDG) treatment during the end-polishing step and unique 7-mer 5' and 3'

176 barcoded adapters during the ligation step (Rohland and Reich, 2012). We then performed a real-

177 time PCR assay to determine how many cycles of PCR were required to optimise library quantity

178 and complexity (Gamba et al., 2016). Duplicate real-time PCR assays were performed for each

179 library in a final volume of 10 µL, each comprising 1 µL of a 1:5 dilution of library, 1 x Platinum Taq

180 DNA Polymerase High Fidelity buffer (ThermoFisher Scientific), 2 mM MgSO4 (ThermoFisher

181 Scientific), 0.25 mM of each dNTP (ThermoFisher Scientific), 0.4 µM of each primer

182 (IS7_short_amp_P5 and IS8_short_amp_P7; Meyer and Kircher, 2010), 0.004 x ROX (Life Tech),

183 0.2 x SYBR (Life Tech), 0.56 M DMSO (Sigma-Aldrich), and 0.2 U of Platinum Taq DNA

184 Polymerase High Fidelity (ThermoFisher Scientific), in laboratory grade water. Real-time PCRs

185 were performed on a LightCycler 96 (Roche) with the following cycling conditions: 94 °C for 6 min;

186 40 cycles of 94 °C for 30 s, 60 °C for 30 s, 68 °C for 40 s; followed by a high-resolution melt.

187 Results from our rtPCR suggested substantially fewer cycles were required to amplify libraries

188 created from our samples (11-17 cycles) compared with those made from our negative controls

189 (21-28 cycles), suggesting low DNA template quantities in our controls and negligible levels of

190 cross-contamination.

191

192 The libraries were then amplified using conventional PCR. In order to maintain library complexity

193 and minimise PCR bias, each library was amplified in eight separate 25 µL reactions, each

194 comprising 3 µL of undiluted library, 1 x Platinum Taq DNA Polymerase High Fidelity buffer

195 (ThermoFisher Scientific), 2 mM MgSO4 (ThermoFisher Scientific), 0.25 mM of each dNTP

196 (ThermoFisher Scientific), 0.4 µM of each primer (IS7_short_amp_P5 and IS8_short_amp_P7;

197 Meyer and Kircher, 2010), and 0.2 U of Platinum Taq DNA Polymerase High Fidelity

198 (ThermoFisher Scientific), in laboratory grade water. Cycling conditions for the PCR were as

199 follows: 94 °C for 6 min; between 11 and 17 cycles (as determined above using rtPCR) of 94 °C for

200 30 s, 60 °C for 30 s, 68 °C for 40 s; and 68 °C for 10 min. The exception was the library made for

201 UW51516, which was subjected to Recombinase Polymerase Amplification (RPA) for 40 min using

202 a TwistAmp Basic kit (TwistDx Inc.) and following the manufacturer's protocol. Amplified libraries

203 were pooled and purified using 1.8 x volume AxyPrep (Axygen), washed twice with 80% ethanol,

204 and then resuspended in 30 µL of buffer comprising 10 mM Tris, 0.1 mM EDTA, and 0.05%

205 Tween-20.

206

207 All libraries were enriched for placental mammal mitochondrial DNA using hybridisation enrichment

208 with the commercially synthesised RNA probes described by Mitchell et al. (2016). Hybridisation

209 enrichment was performed according to the manufacturer's protocol (Arbor Biosciences: myBaits

210 v3 chemistry) with several modifications: (1) we extended the incubation step to 44 hr (15 hr at 55

211 °C, 16 hr at 50 °C, 17 hr at 55 °C); (2) we used RNA blockers complementary to our truncated

212 library adapters instead of the Blocker #3 provided by the manufacturer; (3) prior to immobilising

213 the RNA baits, we incubated the Dynabeads MyOne Streptavidin C1 (ThermoFisher Scientific) with

214 100 µg of yeast tRNA to saturate bead sites that bind nucleic acids in a non-specific manner

215 (Llamas et al., 2016); and (4) we washed the RNA baits—once bound to the streptavidin beads—

216 three times by incubating for 5 min at 55 °C with 0.1 SSC and 0.1% SDS (discarding the

217 supernatant after each wash).

218

219 Following, post-enrichment purification, all libraries were eluted in 125uL of PCR master mix (1 ×

220 PCR buffer, 2.5 mM MgCl2, 1 mM dNTPs, 0.5 mM primer, 6.25 U AmpliTaq Gold). The master mix

221 from each library was then split into five reactions and subjected to the following thermocycling

222 regime: 94 °C 6 min; 15 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 45 s; and a final

223 extension of 72 °C for 10 min. Forward and reverse primers included full-length indexed Illumina

224 sequencing adapters (see Meyer and Kircher, 2010). PCR products from each library were pooled

225 and purified using 1.1 x volume AxyPrep (Axygen), washed twice with 80% ethanol, and then
226 resuspended in 30 μL of water.

227

228 *2.4 High-throughput sequencing and data processing*

229

230 All libraries were pooled and sequenced together on either an Illumina NextSeq or HiSeq in paired-
231 end sequencing mode. Raw sequencing reads were demultiplexed using "sabre"
232 (http://github.com/najoshi/sabre) according to their unique 7-mer barcode combinations. Using
233 AdapterRemoval v2.1.2 (Schubert et al., 2016) we trimmed residual adapters and low-quality
234 bases (<Phred20 –minquality 4); merged overlapping paired-end reads (minimum overlap = 11 nt);
235 and discarded merged reads <30 bp (–minlength 30). Read quality was visualised using fastQC
236 v0.10.1 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) before and after trimming to
237 make sure the trimming was efficient.

238

239 Mitochondrial consensus sequences were obtained by mapping all merged reads for each library
240 against a previously published reference sequence for their respective species—KT168321 for
241 stilt-legged horses, KT168318 for caballine horses, and KR822421 for yesterday's camels—using
242 BWA v0.7.8 (Li and Durbin, 2009; aln -t 8 -l 1024 -n 0.04 -o 2). Reads with a mapping quality
243 Phred score >30 were selected and retained using the SAMtools v1.4 (Li et al., 2009) view
244 command (-q 30), and duplicate reads were discarded using 'FilterUniqueSAMCons.py' (Kircher,
245 2012). A final 75% majority consensus sequence was then generated for each library and checked
246 by eye in Geneious v9.1.6 (https://www.geneious.com), calling nucleotides for sites with a
247 minimum depth-of-coverage of 3x. Summary statistics for each consensus sequences are provided
248 in Table S1.

249

250 *2.5 Phylogenetic analyses*

251

252 We aligned our new mitochondrial genome sequences with previously published data (Table S1)
253 using the MUSCLE v3.8.425 (Edgar, 2004) algorithm as implemented in Geneious. Three separate
254 alignments were created: one for stilt-legged horses (n=39; 16,655 bp), one for caballine horses
255 (n=34; 16,662 bp), and one for yesterday's camels (n=6; 16,681 bp). Ambiguously aligned columns
256 were removed using Gblocks v0.91b (Castresana, 2000) with default settings, which reduced the
257 length of our caballine horse alignment to 16,317 bp. We inferred maximum likelihood phylogenies
258 based on our stilt-legged and caballine horse alignments using IQ-TREE v1.6.11 (Nguyen et al.,
259 2015), with the best-fitting substitution model (HKY+I) selected using ModelFinder (according to
260 the Bayesian Information Criterion) as implemented in IQ-TREE (Kalyaanamoorthy et al., 2017)
261 and 1000 ultrafast bootstrap replicates to assess topological support (Hoang et al., 2017). We

262 created a median-joining haplotype network (Bandelt et al., 1999) from our yesterday's camel
263 alignment using PopART v1.7 (Leigh and Bryant, 2015).

264

265 We used BEAST v1.8.4 (Drummond and Rambaut, 2007) to co-estimate phylogenies, node ages,
266 and tip ages (for specimens without ages measured using radiocarbon dating) using our stilt-
267 legged and caballine horse alignments. We first evaluated the temporal signal in these two
268 alignments using leave-one-out cross-validation (see Stiller et al., 2014) after pruning our
269 alignment to only the sequences from specimens with finite radiocarbon ages (18 stilt-legged
270 horses and 23 caballine horses; see Table S1). Cross validation involved a series of analyses
271 wherein the age of each sample was sequentially omitted and estimated (applying a uniform prior
272 of 0-150 ka—reflecting a range of plausible deposition ages—instead of specifying the radiocarbon
273 age of the specimen). In each case, we applied the best fitting model estimated previously using
274 IQ-TREE (HKY+I), used a strict molecular clock model, and applied a constant population size
275 coalescent tree prior. A uniform prior of 10^{-11} to 10^{-5} substitutions per site per year was placed on
276 the clock rate. The Markov chain Monte Carlo (MCMC) was run for 2×10^{6} generations sampling
277 trees and parameter values every 2000 generations. Convergence of parameter values and ESSs
278 > 200 were monitored using Tracer v1.7.1 (Rambaut et al., 2018). For all except one caballine
279 horse, the calibrated radiocarbon age fell within the 95% Highest Posterior Density (95% HPD) of
280 the Bayesian estimate (Table S1), suggesting that our data collectively included sufficient temporal
281 information to estimate the age of undated samples. For the one caballine horse sample that failed
282 cross-validation (IMNH 1136/11898), the estimated age (median = 52.5 ka; 95% HPD = 33.3ka –
283 74.7 ka) was substantially older than the calibrated radiocarbon age (median = 17275 cal BP;
284 CAMS_LLNL-175552), possibly due to contamination of the sample with relatively young carbon
285 (e.g. from adhesives or consolidants; Crann and Grant, 2019) that was not removed prior to
286 radiocarbon dating; as a result, we used the median estimated age from BEAST for this one
287 specimen in all downstream analyses.

288

289 We subsequently performed another series of BEAST analyses wherein those sequences from
290 horse specimens without radiocarbon ages or with infinite radiocarbon ages were sequentially and
291 individually added into their respective alignments in order to estimate the age of the specimens
292 (applying a uniform prior of 0-150 ka on the unknown age). Otherwise, these runs used the same
293 priors and MCMC settings as for the cross-validation analyses described above. Two stilt-legged
294 horse sequences were excluded at this point because their position in the tree precluded accurate
295 date estimate (they were an outgroup to all directly dated samples; Figure 2A). Once all other
296 samples were assigned an age (either based on radiocarbon dating or Bayesian date estimation),
297 we conducted a date-randomisation test for each alignment (Ramsden et al., 2008; Stiller et al.,
298 2014). The date randomisation tests involved assigning each sample an age from the set of all
299 sample ages (sampling without replacement), which we did by extracting the sample ages from our

300 BEAST XML files, re-ordering them according to a randomly assigned integer, and then re-
301 assigning the re-ordered ages to the samples as they were ordered in the original BEAST XML file.
302 For both alignments the posterior substitution rate estimate of the original data did not overlap the
303 95% HPDs of the rate estimates from ten such randomised replicates, suggesting that our dataset
304 could be used to reliably estimate evolutionary rate and divergence times (Figure S1). Again, these
305 date randomisation runs used the same priors and MCMC settings as for the cross-validation.

306

307 We then ran two final BEAST analyses for our stilt-legged horse and caballine horse alignments.
308 These analyses were run as above, except we used an Extended Bayesian Skyline coalescent
309 tree prior, posterior medians for the age of sequences without finite radiocarbon ages, and three
310 separate MCMCs. After removing the first 10% of values sampled by each MCMC, we combined
311 the remaining samples using LogCombiner v1.8.4 and created a maximum clade credibility tree
312 using TreeAnnotator v1.8.4. Convergence of parameter values between the three chains and
313 combined effective sample sizes > 200 were assessed using Tracer v1.7.1. We observed that the
314 inclusion of several sequences in each alignment with >20% indeterminate nucleotides (i.e. coded
315 as N) were contributing to topological uncertainty—reducing branch support across the tree—so
316 we excluded these from our final BEAST analyses (see Table S1; Figure 2A, 3A). Consequently,
317 our final alignments for stilt-legged and caballine horses comprised 32 sequences each.

318

319 In addition to analyses of our horse alignments, we also ran a BEAST analysis for our yesterday's
320 camel alignment. However, because that alignment only contained six sequences—including only
321 three specimens with finite radiocarbon ages (Table S1)—we could not perform the cross-
322 validation or date randomisation tests described above for the stilt-legged and caballine horses.
323 Instead, we constrained the age of the three Eastern Beringian camel specimens with infinite
324 radiocarbon ages using uniform distributions from 50 to 150 ka. We also placed a uniform
325 distribution on the substitution rate of 5.0×10^{-9} to 4.0×10^{-8} substitutions per site per year, which
326 spans a range of values typical for large terrestrial mammals, and we used a constant population
327 size coalescent tree prior. Three separate MCMCs were run for 10^6 generations sampling trees
328 and parameter values every 1000 generations. Otherwise, this analysis was performed using the
329 same settings as for our final analyses of the horse alignments and results were summarised in the
330 same way. Importantly, the posterior ages estimated for camels in this study should only be taken
331 as indicative until they are subject to more rigorous analyses with larger sample sizes that allow for
332 internal validation.

333

334 **3. Results and Discussion**

335

336 *3.1 Stilt-legged horses* (Haringtonhippus)

337

338   Of the 16 horse specimens from Natural Trap Cave that we analysed, 13 yielded mitochondrial
339   haplotypes that showed a close affinity for published sequences from the stilt-legged horse,
340   *Haringtonhippus francisci* (Figure 2A). Eight of these new haplotypes were unique, with the
341   remaining five plausibly representing different specimens from the same individual animals. These
342   eight new haplotypes were all distinct from sequences published by Heintzman et al. (2017), which
343   brings the total number of unique *Haringtonhippus* haplotypes—effectively equivalent to the
344   minimum number of individuals—known from Natural Trap Cave to 14.

345

346   The results of our phylogenetic analyses revealed that 11 of our 13 new *Haringtonhippus*
347   sequences fall within the mitochondrial diversity described by Heintzman et al. (2017), though
348   sequences from Natural Trap Cave do not form a monophyletic clade to the exclusion of
349   sequences from Eastern Beringia or Nevada (Figure 2A). The remaining two sequences—
350   comprising a single unique haplotype—may represent a sister-lineage to all other sequences
351   (Figure 2A). In our view, the genetic distance between these two outgroup sequences and the
352   remaining sequences is unlikely to be of taxonomic significance, although that hypothesis could be
353   tested more rigorously in the future with additional data. Currently, these outgroup specimens have
354   not been directly radiocarbon dated and are also excluded from the results of our Bayesian
355   analysis because the long branch separating them from the remaining samples caused problems
356   with date estimation and convergence of the MCMC (see Section 2.5; Table S1).

357

358   All of the stilt-legged horse sequences from Natural Trap Cave included in our final Bayesian
359   analysis shared a common ancestor more recently with a sequence from Eastern Beringia, and
360   *vice versa*, than with any of the three sequences previously reported from Gypsum Cave in
361   Nevada (Bayesian posterior probability, BPP = 1.0; Figure 2B). This pattern is consistent with
362   ongoing gene flow between Eastern Beringian stilt-legged horse populations and those near to
363   Natural Trap Cave during the Pleistocene, though uncertainty associated with our node age
364   estimates makes the precise timeframe unclear. In addition, high levels of missing data prevented
365   us from confidently determining the affinities of six additional specimens from Natural Trap Cave
366   and one specimen from Mineral Hill Cave, Nevada (see Section 2.5, Figure 2A, Table S1).
367   Consequently, while our results may be suggestive, it is difficult to draw firm conclusions about the
368   connectivity of stilt-legged horse populations in Eastern Beringia and near Natural Trap Cave with
369   those further from the southern terminus of the ice-free corridor (e.g. those in Nevada; Figure 1).

370

371   *3.2 Caballine horses* (Equus)

372

373   Three of our horse specimens from Natural Trap cave yielded caballine horse (*Equus*)
374   mitochondrial haplotypes (Figure 3A)—only two of these were unique, suggesting that they may
375   represent only two different individuals. Vershinina et al. (2021) recently reported caballine horse

376 mitochondrial genome sequences from another four specimens from Natural Trap Cave; however,
377 three of their sequences are identical and could plausibly represent multiple specimens from a
378 single individual animal, especially since radiocarbon ages for all three specimens are practically
379 indistinguishable. Our new data therefore bring the total number of unique caballine horse
380 haplotypes known from Natural Trap Cave to four, which all belong to Vershinina et al.'s (2021)
381 "clade B". Consequently, we only included clade B haplotypes in our downstream analyses. Within
382 North America, clade A haplotypes—specifically A1 and A2 haplotypes—have been reported only
383 from Eastern Beringia and appear to derive from eastward migration across the Bering Land
384 Bridge from Eurasia between 50 and 200 ka (Vershinina et al., 2021).

385

386 As with the stilt-legged horses (Section 3.1, Figure 2), we observed no close affinity between
387 caballine horse sequences from Natural Trap Cave and those obtained from specimens elsewhere
388 in the contiguous US (Figure 3)—represented in our analyses by two sequences from Idaho (one
389 of which we sequenced as part of this study). Instead, one of our new Natural Trap Cave
390 sequences formed a clade with one of Vershinina et al.'s (2021) Natural Trap Cave sequences
391 (BPP = 1.0), which in turn shared a more recent common ancestor with a sequence from Eastern
392 Beringia (BPP = 0.94; Figure 3B), while the remaining two of our new Natural Trap Cave
393 sequences were excluded from our final analysis due to high levels of missing data (see Section
394 2.5, Table S1, Figure 3A). In contrast, Vershinina et al.'s (2021) remaining three sequences from
395 Natural Trap Cave represent a relatively distinct lineage, which last shared a common ancestor
396 with other sequences >100 ka; as for similarly distinct stilt-legged horse lineages from Natural Trap
397 Cave (Section 3.1), this distinct caballine horse lineage may indicate persistent local
398 phylogeographic structure in addition to gene flow with populations in Eastern Beringia.

399

400 Unlike the stilt-legged horse samples from Nevada, which were all closely related (Figure 2), the
401 caballine horse sequences from Idaho were the respective sister lineages to two distinct clades
402 otherwise comprising samples from Eastern Beringia and Natural Trap Cave or Alberta (BPP =
403 0.94 & 1.0, respectively; Figure 3B). The majority of node age estimates within these clades—
404 including the common ancestors of our new Natural Trap Cave sequence and its nearest Eastern
405 Beringian relative—fall within Marine Isotope Stage 3 (MIS 3; 29-57 ka; Figure 3B), when an ice-
406 free corridor was likely present. This is consistent with the occurrence of gene flow between horse
407 populations in Eastern Beringia and those near to Natural Trap Cave via the ice-free corridor prior
408 to coalescence of the ice-sheets during the Last Glacial Maximum. However, age estimates for the
409 common ancestors between these clades and their respective nearest relatives from Idaho (95%
410 HPDs = 65.1-84.9 ka & 50.5-65.9 ka, respectively) substantially overlap with MIS 4 (57-71 ka),
411 when the ice-free corridor may have been inaccessible or less traversable. Together, these
412 observations suggest that populations in the contiguous US, particularly those near to Natural Trap
413 Cave, may have been the source of *Equus* clade B diversity observed in Eastern Beringia during

414 MIS 3, with the majority of gene flow occurring from south to north. This hypothesis is further

415 supported by the apparent absence of *Equus* clade A1 and A2 haplotypes from the contiguous

416 US—otherwise found only in Eastern Beringia—despite increasing the number of samples that

417 were examined by Vershinina et al. (2021).

418

419 *3.3 Yesterday's camel* (Camelops)

420

421 Our Bayesian phylogenetic analysis of yesterday's camel (*Camelops* sp.) included sequences from

422 six specimens: three from Eastern Beringia published by Heintzman et al. (2015), one from Natural

423 Trap Cave, one from Spider Cave in Idaho, and one from Mineral Hill Cave in Nevada (Figure 4A).

424 Our results strongly supported reciprocal monophyly of a clade comprising the sequences from

425 Idaho and Nevada (BPP = 1.0) and a clade comprising the Eastern Beringian sequences and the

426 sequence from Natural Trap Cave (BPP = 1.0). The common ancestor of these two clades

427 occurred between 213 and 836 ka (95% HPD; median = 405 ka; Figure 4B), suggesting that they

428 all likely belong to a single species (*Camelops* cf. *hesternus*; Heintzman et al., 2015). However,

429 our node age estimates for *Camelops* are relatively imprecise and should be treated with caution

430 because they were not estimated using as informative or objective priors compared to our

431 analyses of caballine and stilt-legged horses (see Section 2.5).

432

433 The camel sample from Natural Trap Cave was most closely related to one of the Eastern

434 Beringian samples—YG 328.23—to the exclusion of the remaining two (Figure 4). Our results

435 suggest that the common ancestor of the Natural Trap Cave specimen and its nearest relative

436 occurred between 65.3 and 195 ka (95% HPD; median = 122 ka; Figure 4B). As for the horses

437 (Sections 3.1 & 3.2), this pattern is consistent with greater population connectivity and gene flow

438 between Eastern Beringian populations and those near to Natural Trap Cave versus populations in

439 the contiguous US further from the southern terminus of the ice-free corridor. However, because

440 our dataset includes very few individuals—as for a previous study the extinct musk-oxen

441 *Bootherium bombifrons* (Bover et al., 2018)—this result may be a sampling artefact and needs to

442 be confirmed in the future with more comprehensive sampling.

443

444 *3.4 Synthesis*

445

446 Contrary to previous work that suggested as many as four distinct horse species were represented

447 among the Pleistocene fossils from Natural Trap Cave (e.g. Eisenmann et al., 2008), genetic data

448 to date—including our new sequences—only provide strong evidence for two species: one species

449 of stilt-legged horse (*Haringtonhippus francisci*) and one species of caballine horse (*Equus* sp.).

450 This conclusion remains true even if the wider mitochondrial diversity described by Vershinina et

451 al. (2021) is interpreted as corresponding to several distinct species (e.g. *Equus ferus*, *E. lambei*,

452 *E. scotti*, *E. occidentalis*), because sequences from Natural Trap Cave all belong to a relatively
453 restricted subset of overall caballine horse diversity ("clade B"; Figure 3). Additional sampling may
454 yet reveal genetic evidence for additional lineages at Natural Trap Cave, but if so they must occur
455 only at very low abundance, having not been detected among the 18 unique horse haplotypes thus
456 far obtained. With respect to abundances, we also note that—assuming horse samples have been
457 randomly chosen for genetic analysis—genetic data are consistent with a roughly three-fold higher
458 abundance of stilt-legged horses versus caballine horses in the Natural Trap Cave assemblage.

459

460 Overall, our results from stilt-legged horses, caballine horses, and yesterday's camels are all
461 consistent with a higher level of connectivity between populations in Eastern Beringia and those
462 near Natural Trap Cave during the Pleistocene when compared with populations further from the
463 southern terminus of the ice-free corridor (e.g. those in Idaho or Nevada; Figure 1). However, the
464 strength of this conclusion is limited by sparse sampling from localities in the contiguous US other
465 than Natural Trap Cave and the imprecision of our node age estimates, specifically for stilt-legged
466 horses and yesterday's camels. Greater sampling intensity in future studies may overcome these
467 limitations. Hall's Cave in Texas, where short fragments of DNA from both *Camelops* and
468 *Haringtonhippus* have been detected in bulk bone samples (Seersholm et al., 2020), may be a
469 promising site for expanding the geographical breadth of datasets for these species. The inclusion
470 of nuclear DNA—if it can reliably be obtained from specimens from the contiguous US—would also
471 help to reveal evidence for finer-scale gene flow that is not captured in the mitochondrial phylogeny
472 of these taxa. Nevertheless, our data reveal intriguing patterns, specifically the lack of strong
473 evidence for southward versus northward dispersal of caballine horses through the ice-free
474 corridor during MIS3. This contrasts with data from bison, brown bears, and lions, which suggest
475 dispersal of these taxa through the ice-free corridor occurred primarily from Eastern Beringia into
476 the contiguous US (Heintzman et al., 2016; Salis et al., 2020), emphasising that patterns of
477 megafaunal dispersal during the Pleistocene are species specific. It may therefore be illuminating
478 to expand sampling of ancient DNA from the contiguous US—including Natural Trap Cave—to
479 include taxa like grey wolves and bighorn sheep, which may also have traversed the ice-free
480 corridor (e.g. Meachen et al., 2016).

481

482 **Author contributions**

483

484 Conceptualisation: KJM, AC, JAM; Investigation: KJM, CM, PB, ATS, HH; Formal analysis: KJM;
485 Data Curation: KJM, CM, PB, ATS, HH, JAM, MT, BH; Resources: KJM, AC, LSW, JAM, MT, BH;
486 Visualisation and Writing (Original Draft): KJM; Writing (Review & Editing): all authors; Funding
487 acquisition: AC, JAM.

488

489 **Data availability**

13

490

491 Mitochondrial consensus sequences produced as part of this study are available on GenBank

492 (TBA-TBA). Consensus sequences, demultiplexed sequencing reads, and phylogenetic analysis

493 files—including BEAST XMLs—are available through figshare (DOI: TBA).

494

495 **Declaration of competing interests**

496

497 The authors declare that they have no known competing financial interests or personal

498 relationships that could have appeared to influence the work reported in this paper.

499

500 **Acknowledgements**

501

512

513 **References**

514

515 Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific

516   phylogenies. Mol Biol Evol 16(1):37-48.

517 Barnett R, Barnes I, Phillips MJ, Martin LD, Harington CR, Leonard JA, Cooper A. 2005. Evolution

518   of the extinct sabretooths and the American cheetah-like cat. Curr Biol 15(15):R589-R590.

519 Baskin J, Thomas R. 2016. A review of *Camelops* (Mammalia, Artiodactyla, Camelidae), a giant

520   llama from the Middle and Late Pleistocene (Irvingtonian and Rancholabrean) of North

521   America. Hist Biol 28(1-2):120-127.

522 Bover P, Llamas B, Thomson VA, Pons J, Cooper A, Mitchell KJ. 2018. Molecular resolution to a

523   morphological controversy: The case of North American fossil muskoxen *Bootherium* and

524   *Symbos*. Mol Phylogenet Evol 129:70-76.

525 Bronk Ramsey C. 2016. Bayesian analysis of radiocarbon dates. Radiocarbon 51(1):337-360.

526 Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in

527   phylogenetic analysis. Mol Biol Evol 17(4):540-552.

528 Crann CA, Grant T. 2019. Radiocarbon age of consolidants and adhesives used in archaeological
529   conservation. Journal of Archaeological Science: Reports 24:1059-1063.

530 Dabney J, Knapp M, Glocke I, Gansauge M-T, Weihmann A, Nickel B, Valdiosera C, García N,
531   Pääbo S, Arsuaga J-L, Meyer M. 2013. Complete mitochondrial genome sequence of a
532   Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proceedings of
533   the National Academy of Sciences 110(39):15758-15763.

534 Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC
535   Evol Biol 7(214).

536 Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
537   Nucleic Acids Res 32(5):1792-1797.

538 Eisenmann V, Howe J, Pichardo M. 2008. Old world hemiones and new world slender species
539   (Mammalia, Equidae). Palaeovertebrata 36:159-233.

540 Gamba C, Hanghøj K, Gaunitz C, Alfarhan AH, Alquraishi SA, Al-Rasheid KAS, Bradley DG,
541   Orlando L. 2016. Comparing the performance of three ancient DNA extraction methods for
542   high-throughput sequencing. Molecular Ecology Resources 16(2):459-469.

543 Heintzman PD, Froese D, Ives JW, Soares AER, Zazula GD, Letts B, Andrews TD, Driver JC, Hall
544   E, Hare PG, Jass CN, MacKay G, Southon JR, Stiller M, Woywitka R, Suchard MA,
545   Shapiro B. 2016. Bison phylogeography constrains dispersal and viability of the Ice Free
546   Corridor in western Canada. Proc Natl Acad Sci U S A 113(29):8057-8063.

547 Heintzman PD, Zazula GD, Cahill JA, Reyes AV, MacPhee RDE, Shapiro B. 2015. Genomic Data
548   from Extinct North American *Camelops* Revise Camel Evolutionary History. Mol Biol Evol
549   32(9):2433-2440.

550 Heintzman PD, Zazula GD, MacPhee RDE, Scott E, Cahill JA, McHorse BK, Kapp JD, Stiller M,
551   Wooller MJ, Orlando L, Southon J, Froese DG, Shapiro B. 2017. A new genus of horse
552   from Pleistocene North America. eLife 6:e29944.

553 Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2017. UFBoot2: Improving the
554   Ultrafast Bootstrap Approximation. Mol Biol Evol 35(2):518-522.

555 Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast
556   model selection for accurate phylogenetic estimates. Nature methods 14(6):587-589.

557 Kircher M. 2012. Analysis of High-Throughput Ancient DNA Sequencing Data. Ancient DNA:
558   Methods and Protocols. p 197-228.

559 Kooyman B, Hills LV, Tolman S, McNeil P. 2012. Late Pleistocene western camel (*Camelops
560   hesternus*) hunting in southwestern Canada. Am Antiq 77(1):115-124.

561 Leigh JW, Bryant D. 2015. popart: full-feature software for haplotype network construction.
562   Methods in Ecology and Evolution 6(9):1110-1116.

563 Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform.
564   Bioinformatics 25(14):1754-1760.

15

565 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R,
566   Subgroup GPDP. 2009. The Sequence Alignment/Map (SAM) format and SAMtools.
567   Bioinformatics 25(16):2078-2079.

568 Llamas B, Fehren-Schmitz L, Valverde G, Soubrier J, Mallick S, Rohland N, Nordenfelt S,
569   Valdiosera C, Richards SM, Rohrlach A, Romero MIB, Espinoza IF, Cagigao ET, Jiménez
570   LW, Makowski K, Reyna ISL, Lory JM, Torrez JAB, Rivera MA, Burger RL, Ceruti MC,
571   Reinhard J, Wells RS, Politis G, Santoro CM, Standen VG, Smith C, Reich D, Ho SYW,
572   Cooper A, Haak W. 2016. Ancient mitochondrial DNA provides high-resolution time scale of
573   the peopling of the Americas. Science Advances 2(4).

574 Lorenzen ED, Nogues-Bravo D, Orlando L, Weinstock J, Binladen J, Marske KA, Ugan A,
575   Borregaard MK, Gilbert MT, Nielsen R, Ho SY, Goebel T, Graf KE, Byers D, Stenderup JT,
576   Rasmussen M, Campos PF, Leonard JA, Koepfli KP, Froese D, Zazula G, Stafford TW, Jr.,
577   Aaris-Sorensen K, Batra P, Haywood AM, Singarayer JS, Valdes PJ, Boeskorov G, Burns
578   JA, Davydov SP, Haile J, Jenkins DL, Kosintsev P, Kuznetsova T, Lai X, Martin LD,
579   McDonald HG, Mol D, Meldgaard M, Munch K, Stephan E, Sablin M, Sommer RS, Sipko T,
580   Scott E, Suchard MA, Tikhonov A, Willerslev R, Wayne RK, Cooper A, Hofreiter M, Sher A,
581   Shapiro B, Rahbek C, Willerslev E. 2011. Species-specific responses of Late Quaternary
582   megafauna to climate and humans. Nature 479(7373):359-364.

583 Lovelace DM, Redman CM, Minckley TA, Schubert BW, Mahan S, Wood JR, McGuire JL, Laden J,
584   Bitterman K, Heiniger H, Fenderson L, Cooper A, Mitchell KJ, Meachen JA. This issue. An
585   age-depth model and revised stratigraphy of vertebrate-bearing units in Natural Trap Cave,
586   Wyoming. Quat Int.

587 Meachen JA, Brannick AL, Fry TJ. 2016. Extinct Beringian wolf morphotype found in the
588   continental U.S. has implications for wolf migration and evolution. Ecology and Evolution
589   6(10):3430-3438.

590 Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target
591   capture and sequencing. Cold Spring Harbor Protocols 2010(6):1-10.

592 Mitchell KJ, Bray SC, Bover P, Soibelzon L, Schubert BW, Prevosti F, Prieto A, Martin F, Austin JJ,
593   Cooper A. 2016. Ancient mitochondrial DNA reveals convergent evolution of giant short-
594   faced bears (Tremarctinae) in North and South America. Biol Lett 12(4).

595 Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic
596   algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32(1):268-274.

597 Orlando L, Male D, Alberdi MT, Prado JL, Prieto A, Cooper A, Hanni C. 2008. Ancient DNA
598   clarifies the evolutionary history of American late Pleistocene equids. J Mol Evol 66(5):533-
599   538.

600 Perri AR, Mitchell KJ, Mouton A, Álvarez-Carretero S, Hulme-Beaman A, Haile J, Jamieson A,
601   Meachen J, Lin AT, Schubert BW, Ameen C, Antipina EE, Bover P, Brace S, Carmagnini A,
602   Carøe C, Samaniego Castruita JA, Chatters JC, Dobney K, dos Reis M, Evin A, Gaubert P,

603 Gopalakrishnan S, Gower G, Heiniger H, Helgen KM, Kapp J, Kosintsev PA, Linderholm A,
604 Ozga AT, Presslee S, Salis AT, Saremi NF, Shew C, Skerry K, Taranenko DE, Thompson
605 M, Sablin MV, Kuzmin YV, Collins MJ, Sinding M-HS, Gilbert MTP, Stone AC, Shapiro B,
606 Van Valkenburgh B, Wayne RK, Larson G, Cooper A, Frantz LAF. 2021. Dire wolves were
607 the last of an ancient New World canid lineage. Nature 591(7848):87-91.

608 Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior Summarization in
609 Bayesian Phylogenetics Using Tracer 1.7. Syst Biol 67(5):901-904.

610 Ramsden C, Melo FL, Figueiredo LM, Holmes EC, Zanotto PMA, the VC. 2008. High Rates of
611 Molecular Evolution in Hantaviruses. Mol Biol Evol 25(7):1488-1492.

612 Reimer PJ, Austin WEN, Bard E, Bayliss A, Blackwell PG, Bronk Ramsey C, Butzin M, Cheng H,
613 Edwards RL, Friedrich M, Grootes PM, Guilderson TP, Hajdas I, Heaton TJ, Hogg AG,
614 Hughen KA, Kromer B, Manning SW, Muscheler R, Palmer JG, Pearson C, van der Plicht
615 J, Reimer RW, Richards DA, Scott EM, Southon JR, Turney CSM, Wacker L, Adolphi F,
616 Büntgen U, Capano M, Fahrni SM, Fogtmann-Schulz A, Friedrich R, Köhler P, Kudsk S,
617 Miyake F, Olsen J, Reinig F, Sakamoto M, Sookdeo A, Talamo S. 2020. The IntCal20
618 Northern Hemisphere radiocarbon age calibration curve (0–55 cal kBP). Radiocarbon
619 62(4):725-757.

620 Rohland N, Harney E, Mallick S, Nordenfelt S, Reich D. 2015. Partial uracil-DNA-glycosylase
621 treatment for screening of ancient DNA. Philos Trans R Soc Lond B Biol Sci
622 370(1660):20130624.

623 Rohland N, Reich D. 2012. Cost-effective, high-throughput DNA sequencing libraries for
624 multiplexed target capture. Genome Res 22(5):939-946.

625 Salis AT, Bray SCE, Lee MSY, Heiniger H, Barnett R, Burns JA, Doronichev V, Fedje D,
626 Golovanova L, Harington CR, Hockett B, Kosintsev P, Lai X, Mackie Q, Vasiliev S,
627 Weinstock J, Yamaguchi N, Meachen J, Cooper A, Mitchell KJ. 2020. Lions and brown
628 bears colonized North America in multiple synchronous waves of dispersal across the
629 Bering Land Bridge. Mol Ecol. Accepted Author Manuscript.
630 https://doi.org/10.1111/mec.16267

631 Salis AT, Gower G, Schubert BW, Soibelzon LH, Heiniger H, Prieto A, Prevosti FJ, Meachen J,
632 Cooper A, Mitchell KJ. 2021. Ancient genomes reveal hybridisation between extinct short-
633 faced bears and the extant spectacled bear (*Tremarctos ornatus*).
634 bioRxiv:2021.2002.2005.429853.

635 Schubert M, Lindgreen S, Orlando L. 2016. AdapterRemoval v2: rapid adapter trimming,
636 identification, and read merging. BMC Research Notes 9:88.

637 Seersholm FV, Werndly DJ, Grealy A, Johnson T, Keenan Early EM, Lundelius EL, Winsborough
638 B, Farr GE, Toomey R, Hansen AJ, Shapiro B, Waters MR, McDonald G, Linderholm A,
639 Stafford TW, Bunce M. 2020. Rapid range shifts and megafaunal extinctions associated
640 with late Pleistocene climate change. Nature Communications 11(1):2770.

641   Stiller M, Molak M, Prost S, Rabeder G, Baryshnikov G, Rosendahl W, Münzel S, Bocherens H,
642   Grandal-d'Anglade A, Hilpert B, Germonpré M, Stasyk O, Pinhasi R, Tintori A, Rohland N,
643   Mohandesan E, Ho SYW, Hofreiter M, Knapp M. 2014. Mitochondrial DNA diversity and
644   evolution of the Pleistocene cave bear complex. Quat Int 339-340:224-231.

645   Vershinina AO, Heintzman PD, Froese DG, Zazula G, Cassatt-Johnstone M, Dalén L, Der
646   Sarkissian C, Dunn SG, Ermini L, Gamba C, Groves P, Kapp JD, Mann DH, Seguin-
647   Orlando A, Southon J, Stiller M, Wooller MJ, Baryshnikov G, Gimranov D, Scott E, Hall E,
648   Hewitson S, Kirillova I, Kosintsev P, Shidlovsky F, Tong H-W, Tiunov MP, Vartanyan S,
649   Orlando L, Corbett-Detig R, MacPhee RD, Shapiro B. 2021. Ancient horse genomes reveal
650   the timing and extent of dispersals across the Bering Land Bridge. Mol Ecol.

651   Waters MR, Stafford TW, Kooyman B, Hills LV. 2015. Late Pleistocene horse and camel hunting at
652   the southern margin of the ice-free corridor: Reassessing the age of Wally's Beach,
653   Canada. Proc Natl Acad Sci U S A 112(14):4263.

654   Weinstock J, Willerslev E, Sher A, Tong W, Ho SYW, Rubenstein D, Storer J, Burns J, Martin L,
655   Bravi C, Prieto A, Froese D, Scott E, Xulong L, Cooper A. 2005. Evolution, Systematics,
656   and Phylogeography of Pleistocene Horses in the New World: A Molecular Perspective.
657   PLoS Biol 3(8):e241.

658

659   **Figure captions**

660

661   **Figure 1 (colour):** Map of study areas relative to the location of the Cordilleran and Laurentide ice

662   sheets and the ice-free corridor connecting Eastern Beringia (dark gray) to locations in the

663   contiguous US (adapted from Meachen et al., 2016). Depiction approximates extent during

664   Pleistocene glacial minima; during glacial maxima the ice sheets would likely have coalesced and

665   no ice-free corridor would have been present. Natural Trap Cave (blue circle) in Wyoming (blue) is

666   closer to the southern terminus of the ice-free corridor compared to study sites in Idaho (orange;

667   light and dark orange circles) or Nevada (red; light and dark red circles).

668

669   **Figure 2 (colour):** A) Maximum likelihood phylogeny of stilt-legged horse (*Haringtonhippus*

670   *francisci*) mitochondrial genome sequences. Ultrafast bootstrap support from IQ-TREE is displayed

671   for nodes with 95% support or higher. Tips are labelled with a shorthand reference number (see

672   Table S1) and specimen ID; new sequences obtained as part of this study are marked with an

673   asterisk. Coloured circles indicate geographical provenance of samples. Branch lengths are

674   proportional to number of substitutions; scale is in number of substitutions per site. Sequences

675   labelled in light grey were excluded from our final Bayesian analysis due to either a high number of

676   indeterminate nucleotides (given in brackets following the specimen ID) or inability to confidently

677   estimate the age of the specimens (see Section 2.5 and Table S1). B) Time-calibrated Bayesian

678   phylogeny of stilt-legged horse mitochondrial genome sequences. Coloured circles indicate
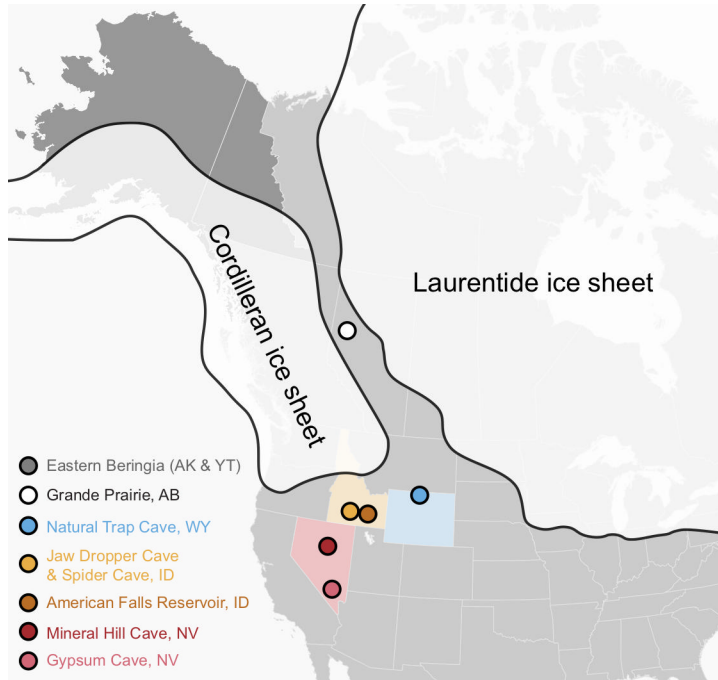
679 geographical distribution of samples. Samples are labelled with a shorthand reference (see Table
680 S1) and new sequences are marked with an asterisk. Shaded vertical bars demarcate Marine
681 Isotope Stages 1 through 7 (even numbered MISs are colder glacials while odd numbered MISs
682 are warmer interglacials). Branch lengths are proportional to time (scaled in thousands of years
683 before present). Tip and node heights are plotted as median values. Horizontal node bars reflect
684 95% Highest Posterior Densities (95% HPDs). Labels reflect Bayesian posterior probability (only
685 displayed for branches with a value of at least 0.90).

686

687 **Figure 3 (colour):** A) Maximum likelihood phylogeny of caballine horse mitochondrial genome
688 sequences corresponding to Vershinina et al.'s (2021) *Equus* sp. "clade B". Ultrafast bootstrap
689 support from IQ-TREE is displayed for nodes with 95% support or higher. Tips are labelled with a
690 shorthand reference number (see Table S1) and specimen ID; new sequences obtained as part of
691 this study are marked with an asterisk. Coloured circles indicate geographical provenance of
692 samples. Branch lengths are proportional to number of substitutions; scale is in number of
693 substitutions per site. Sequences labelled in light grey were excluded from our final Bayesian
694 analysis because they comprised a high number of indeterminate nucleotides (given in brackets
695 following the specimen ID; see Section 2.5 and Table S1). B) Time-calibrated Bayesian phylogeny
696 of caballine horse mitochondrial genome sequences. Coloured circles indicate geographical
697 distribution of samples. New sequences are marked with an asterisk. Shaded bars demarcate
698 Marine Isotope Stages 1 through 6 (even numbered MISs are colder glacials while odd numbered
699 MISs are warmer interglacials). Branch lengths are proportional to time (scaled in thousands of
700 years before present). Tip and node heights are plotted as median values. Node bars reflect 95%
701 Highest Posterior Densities (95% HPDs). Labels reflect Bayesian posterior probability (only
702 displayed for branches with a value of at least 0.90).
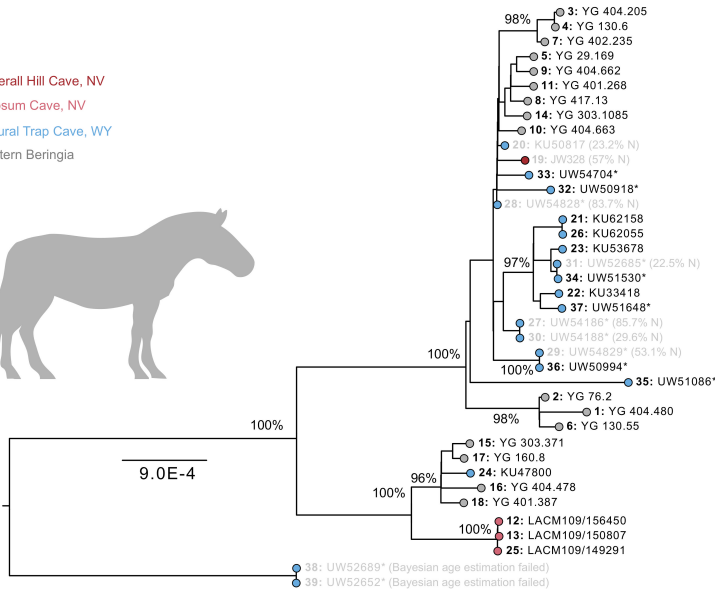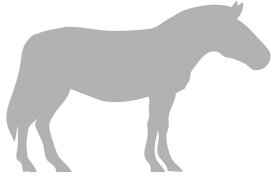
703

704 **Figure 4 (colour):** A) Median-joining haplotype network for yesterday's camel (*Camelops* sp.)
705 mitochondrial genome sequences from Natural Trap Cave (blue circle), Spider Cave (orange
706 circle), Mineral Hill Cave (red circle), and Eastern Beringia (grey circle). Tips are labelled in with a
707 shorthand reference number (see Table S1) and new sequences are marked with an asterisk.
708 Network edges are labelled with the number of substitutions separating haplotypes. B) Time-
709 calibrated Bayesian phylogeny of yesterday's camel mitochondrial genome sequences. Branch
710 lengths are proportional to time (in thousands of years before present); tip and node heights are
711 median values; node bars reflect 95% Highest Posterior Densities (95% HPDs). Labels reflect
712 Bayesian posterior probability.

Cordilleran ice sheet
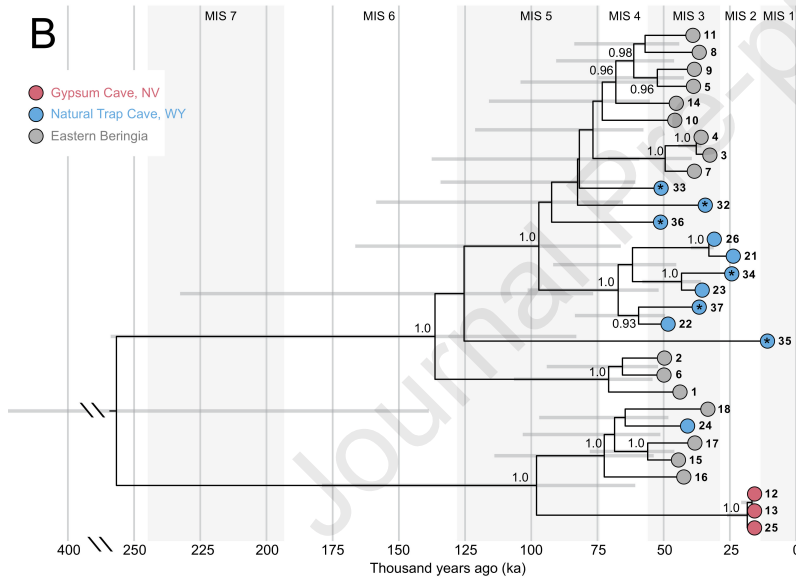
Laurentide ice sheet

- ⬤ Eastern Beringia (AK & YT)
- ⬤ Grande Prairie, AB
- ⬤ Natural Trap Cave, WY
- ⬤ Jaw Dropper Cave & Spider Cave, ID
- ⬤ American Falls Reservoir, ID
- ⬤ Mineral Hill Cave, NV
- ⬤ Gypsum Cave, NV

A

American Falls Reservoir, ID
Jaw Dropper Cave, ID
Natural Trap Cave, WY
Eastern Beringia
Alberta

98% **46:** UAM:ES:25586
95% **68:** YG 328.54
      66: YG 188.42
      45: UAM:ES:7066
      40: UAM:ES:7279
      41: YG 303.428
      42: UAM:ES:22207
      43: YG 126.55
      44: UAM:ES:27507
      69: YG 133.16
      67: YT03/280
      70: UW54187* (78.8% N)
      65: YG 303.325
      47: YG 126.33
100% **73:** IMNH 50001/17237*
      71: UW54706* (26.3% N)
      72: UW54705*
      48: KU42626
      49: QZ-10-393
100% **52:** KU47519
      51: KU47538
      50: KU43413
      53: YG 355.73
      64: P98.21.1
      55: YG 328.94
      54: YG 391.63
      56: YG 401.260
97% **57:** QZ-10-75
      58: YG 303.363
      59: IMNH 1136/11898
      61: YG 328.174
      60: YG 150.52
      62: YG 270.1
100% **63:** UAM:ES:25745

100%

7.0E-4

100%

B

American Falls Reservoir, ID
Jaw Dropper Cave, ID
Natural Trap Cave, WY
Eastern Beringia
Alberta

MIS 6   MIS 5   MIS 4   MIS 3   MIS 2   MIS 1

1.0 **46**
     68
     66
     45
1.0 **40**
     41
     43
     42
1.0 **65**
     47
1.0 **44**
     69
     67
1.0 **72**
0.94 **48**
0.99 **49**
     73
1.0 **52**
     50
     51
     53
     55
     54
1.0 **57**
     56
     58
     59
     64
0.99 **61**
0.96 **60**
1.0 **62**
     63

200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
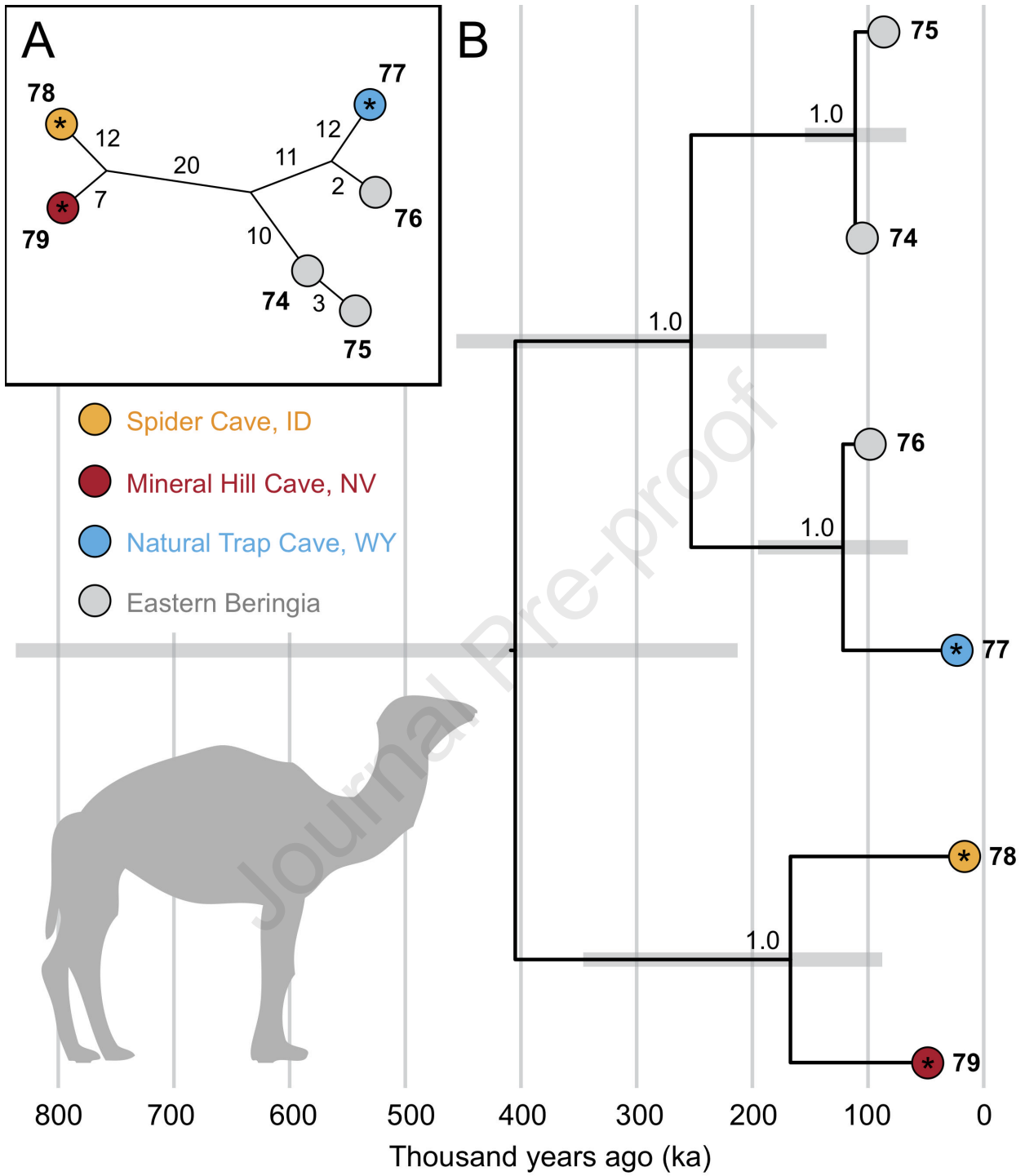Thousand years ago (ka)

Declaration of interests

☐ **The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.**

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: