Palaeogenomics reveal a hybrid origin of the world's largest *Camelus* species

Junxia Yuan[1,2]*, Michael V. Westbury[3]*, Shungang Chen[1,4], Jiaming Hu[5], Fengli Zhang[6], Siren Wang[6], Zhen Zhang[7], Linying Wang[1], Bo Xiao[5], Xindong Hou[8], Fuqiang Li[9], Xulong Lai[2,5], Wenhui Liu[10]*, Guilian Sheng[2,8]*

[1]Faculty of Materials Science and Chemistry, China University of Geosciences, Wuhan, People's Republic of China.

[2]State Key Laboratory of Biogeology and Environmental Geology, China University of Geosciences, Wuhan, People's Republic of China.

[3]Section for Evolutionary Genomics, the GLOBE Institute, University of Copenhagen, Copenhagen, Denmark.

[4]Group of Alpine Paleoecology and Human Adaptation, State Key Laboratory of Tibetan Plateau Earth System, Resources and Environment, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, People's Republic of China.

[5]School of Earth Sciences, China University of Geosciences, Wuhan, People's Republic of China.

[6]Daqing Museum, Daqing, Heilongjiang, People's Republic of China.

[7]Zhaoyuan Museum, Daqing, Heilongjiang, People's Republic of China.

[8]School of Environmental Studies, China University of Geosciences, Wuhan, People's Republic of China.

[9]Yifu Museum, China University of Geosciences, Wuhan, People's Republic of China.

[10]Institute of Environmental Archaeology, National Museum of China, Beijing, People's Republic of China.

*For correspondence: yuanjx@cug.edu.cn (JY); m.westbury@sund.ku.dk (MVW); wenhuiliu89@126.com (WL); glsheng@cug.edu.cn (GS)

Abstract:

The extinct *Camelus knoblochi* is known as the largest camel in genus *Camelus*, but its relationship to modern *Camelus* species remains unclear. In this study, we reported the first mitochondrial and nuclear analyses of seven Late Pleistocene *C. knoblochi* samples from Northeastern China. We found that they are inseparable to wild Bactrian camel on the matrilineal side, but belong to a distinct cluster on the biparental nuclear side. Further admixture proportion analyses suggested hybrid ancestry between both the ancestors of the modern wild and domesticated Bactrian camels, with ~65% contribution from the former and ~35% from the later. By calculating the coalescence time for three *Camelus* species above, we estimated the hybridization event occurred between approximately 0.8 and 0.33 Ma. We also used Bayesian skyline to reconstruct the maternal demographic trajectories for different *Camelus* to better compare their evolutionary histories. Our results provide molecular insights into *C. knoblochi* and fill in a vital piece in understanding the genus *Camelus*.

Key words: camel, *Camelus knoblochi*, Pleistocene, palaeogenome, hybridization

Introduction

The genus *Camelus* contains three extant species, i.e., the one-humped Arabian camel (*C. dromedarius*), and the two-humped wild Bactrian (*C. ferus*) and domesticated Bactrian (*C. bactrianus*) camels. They are all well-adapted to arid climates, with the Arabian camel being distributed in low-altitude torrid regions across northern Africa and southwestern Asia, and the two Bactrian species inhabiting temperate dry environments from China to middle Asia (Xue et al., 2015; Mohandesan et al., 2017; Tibary and El Allali, 2020). Among extant members of *Camelus*, the wild Bactrian camel is the only feral species, while the Arabian and domesticated Bactrian camels, which have played important roles in the progress of human civilization, retain no natural populations (Qi, 2004; Almathen et al., 2016; Orlando, 2016). The genus *Camelus* also contains several fossil species, e.g., *C. knoblochi*, *C. grattatdi*, *C. sivalensis* and *C. thomasi*, among which *C. knoblochi* had the largest body size (Titov, 2008).

Bactrian camels once ranged from the great bend of the Yellow River westward to central Kazakhstan (Tulgat and Schaller, 1992), but their present wild distribution has greatly reduced in recent years (Tulgat and Schaller, 1992). Currently, the wild Bactrian camel is restricted to four isolated regions, three in China and one in Mongolia (Silbermayr et al., 2009). Although the two Bactrian species are physically isolated now, their genetic relationship has been debated for many years. There were two main hypotheses: the wild Bactrian camel was the ancestor of the domesticated, or it was

3

67   feral descendant from the later (Peters and van den Driesch, 1997). Moreover, the wild

68   Bactrian camel was once even considered as a subspecies of the domesticated one,

69   namely *C. bactrianus ferus*. In recent years, mitochondrial and nuclear evidence

70   suggested their clear genetic differentiation without any direct ancestral relationships

71   (Ji et al., 2009; Silbermayr et al., 2009; Zhang et al., 2015; Ming et al., 2020), leaving

72   a suspension for identifying their possibly respective or shared ancestors.

73

74   Based on fossil evidence, *C. knoblochi* ranged from the Sea of Azov Region to

75   Western Transbaikalia during 0.45-0.12 million years ago (Ma). It subsequently shifted

76   its habitat eastward at the beginning of the Late Pleistocene (approximately 0.12 Ma)

77   and was restricted from the Urals to Northeastern China during the Late Pleistocene

78   (Titov, 2008). It is generally thought that *C. knoblochi* became extinct at the end of

79   Pleistocene and was ultimately replaced by Bactrian camels owing to climatic

80   aridization and plant community transitions (Titov, 2008). Morphologically, *C.*

81   *knoblochi* differs from Bactrian camels in its larger body size and broader infraorbital

82   shelf (Geraads et al., 2019), with most other skeletal characteristics being similar.

83   Geraads et al. (2019, 2020) suggested that *C. knoblochi* was more closely related to

84   Arabian camels than to Bactrian camels. In contrast, another study proposed *C.*

85   *knoblochi* was similar to, and potentially ancestral to Bactrian camels (Rowan et al.,

86   2018). Therefore, the relationship between *C. knoblochi* and the extant *Camelus* is not

87   yet well understood.

88

89   Modern and ancient DNA (aDNA) studies have deeply advanced our understanding

90   on the phylogeny and domestication of camelid species (Cui et al., 2007; Ji et al., 2009;

91   Wu et al., 2014; Heintzman et al., 2015; Almathen et al., 2016; Mohandesan et al., 2017;

92   Chen et al., 2019; Lado et al., 2020; Ming et al., 2020; Díaz-Maroto et al., 2021).

93   However, to our knowledge, no molecular analysis has been carried out on *C. knoblochi*.

94   In this study, we retrieved palaeogenomic data from seven Late Pleistocene *C.*

95   *knoblochi* remains. We explored the phylogenetic relationship between this extinct

96   species and extant Old World camels, inferred their respective maternal demographic

97   histories, and estimated the divergence times and genetic diversities of each camelid

98   lineages. The new findings presented here provide us further insight into the genetic

99   components and evolutionary history of genus *Camelus*.

100

Results

Mitochondrial and nuclear genome mapping

103   Seven *C. knoblochi* remains excavated from Northeastern China are included in this

104   study, dating to between around 50 and 25.5 thousand years ago (ka) (Figure 1A;

105   Appendix 1-table 1). Using aDNA methodology and High-throughput sequencing, we

106   successfully obtained near-complete mitochondrial genomes from all individuals with

107   sequencing depths ranging from 6.5 to 17.1 folds. Higher coverages were generated

108   when mapping against the mitogenome of the wild Bactrian camel than to that of the

109   domesticated Bactrian camel (Appendix 1-table 2). The DNA fragments are short

110   (average length: 48-66 bp) and show evident damage characteristics (Appendix 2-

111   figures 1-3), verifying the obtained sequences are not from modern contamination.

112   Endogenous nuclear DNA content ranges between 0.067-2.983% and allowed the

113   recovery of partial nuclear genomes for each sample (0.003-0.069× coverage). Detailed

114   data processing information is listed in Appendix 1-table 2.

115

116   **Sex determination**

117   We found the ratios of X chromosome coverage to autosome coverage to range from

118   0.52 to 0.96. While no individuals displayed the perfect ratios of 0.5 for male or one for

119   female, all individuals were only slightly deviated from these ratios. Therefore, we

120   could clearly see that our *C. knoblochi* dataset contained three males and four females

121   (Appendix 1-table 2).

122

123   **Analysis of maternal phylogeny**

124   A mitogenome BEAST phylogenetic analysis revealed that camelid individuals were

125   divided into two clades (Figure 1B), i.e., Lamini clade and *Camelops-Camelus* clade,

126   consistent with results suggested by Heintzman et al. (2015). In the *Camelops-Camelus*

127   clade, *Camelops* specimens from North America formed a separate subclade while the

128   Old World *Camelus* individuals formed another. The latter branch was further divided

129   into three lineages, with the two species of Bactrian camels being more closely related

130   to each other than to the Arabian camel, also consistent with previous studies

131   (Heintzman et al., 2015; Chen et al., 2019; Ming et al., 2020). Our mitogenome BEAST

132   tree revealed that the extinct *C. knoblochi* was nested within the diversity of extant wild

6

133   Bactrian camel from Mongolia with a high posterior clade probability of 1 (Figure 1B).

134   The entire topology is further robustly support by additional maximum likelihood (ML)

135   tree using RAxML software (Appendix 2-figure 4).

136

137   Using the fossil calibrated node from Camelini vs. Lamini and tip calibrations, we

138   estimated the substitution rate around $6.55×10^{-9}$ substitutions/site/year (95% CI: 5.26-

139   $7.92×10^{-9}$), which is not contradictory to previous calculations, e.g., $8.0×10^{-9}$ in Chen

140   et al. (2019), and $5.0×10^{-9}$ in Fitak et al. (2016). The split between Old World *Camelus*

141   and New World *Camelops* occurred during the Miocene (15.08 Ma, 95% CI: 16.97-

142   13.64 Ma). The divergence time between Arabian and Bactrian camels is approximately

143   2.44 Ma (95% CI: 3.38-1.74 Ma). The coalescent time of all Bactrian species is dated

144   around 0.80 Ma (95% CI: 1.00-0.51 Ma), while the time to the most recent common

145   ancestor (TMRCA) of our samples plus the wild Bactrian camels is 0.33 Ma (95% CI:
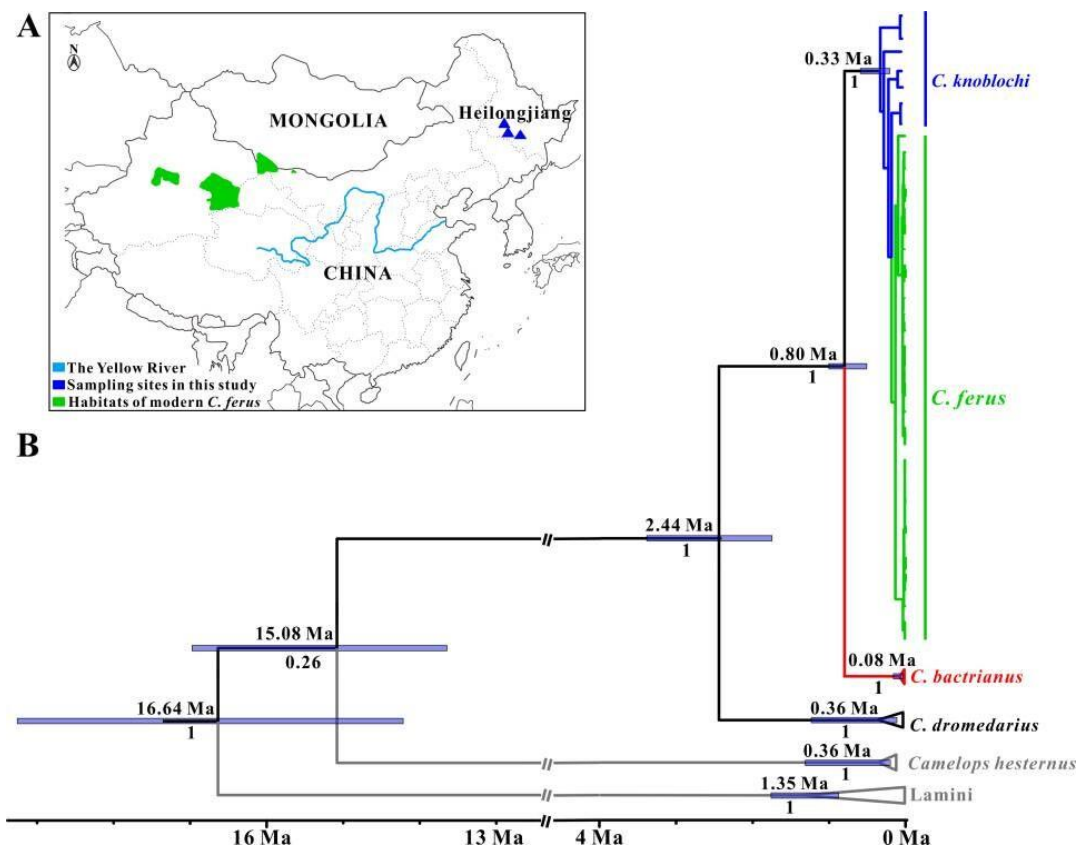
146   0.59-0.21 Ma).

147

Figure 1. (A) Sampling locations of *C. knoblochi* analyzed in this study. (B) Maximum clade credibility tree of family Camelidae based on near-complete mitogenomes computed using BEAST. Node heights are centered on the median posterior age estimates with blue bars showing 95% credibility intervals of the divergence times. Numbers below nodes represent the posterior values. The median radiocarbon ages based on the ^{14}C dating results or the median age from BEAST estimation were used as tip dates of *C. knoblochi* specimens, shown in Appendix 1-table 1. The ages of the remaining samples included in this analysis are listed in Appendix 1-table 3.

Nuclear population structure

Our principal component analysis (PCA) using several representatives per *Camelus* species (Figure 2A) showed Arabian camel to separate first on the PC1 axis (58.9% of

8

161   the variation in the data explained by PC1), and the remaining three species form three

162   distinct clusters along the PC2 axis (7.3% of the variation). We found a near identical

163   result when simulating aDNA damage and down sampling two individuals per extant

164   species, suggesting that this result is biologically driven rather than by biases caused by

165   the low coverage and damaged nature of our *C. knoblochi* data (Appendix 2-figure 5).

166   The PCA computed using the dataset excluding Arabian camel also showed strong

167   separation between all three species along the PC1 axis. On the PC2 axis, substructure

168   appears within domesticated Bactrian camel. As PCs 2, 3, and 4 were explained by

169   similar levels of variance (6.1%, 5.0%, and 4.4% respectively), we also investigated

170   variation along these axes. Similar to PC2, PCs 3 and 4 also arise from variation within

171   domesticated Bactrian camel (Appendix 2-figure 6).

172

173   Our admixture proportion analyses using all *Camelus* species converged on a K=3

174   but not on K=4, suggesting that there are three and not four populations in our dataset.

175   This analysis showed the extant species forming their own populations with no

176   migration between them, and *C. knoblochi* to contain ~35% domesticated Bactrian

177   camel and ~65% wild Bactrian camel ancestry (Figure 2B). We found a similar result

178   when using the aDNA simulated dataset (Appendix 2-figure 7). However, with the

179   aDNA simulated dataset, K=4 did converge. With K=4, we start to uncover some

180   substructure within the domesticated Bactrian camel while *C. knoblochi* still appears as

181   a mixture between the ancestors of domesticated and wild Bactrian camels (Appendix

182   2-figure 7). Due to this convergence, we also looked into the K=4 NGSadmix output

9

183   with the best likelihood for the original dataset. Although it did not converge, we see a

184   similar pattern of substructure within the domesticated Bactrian camel arising as with

185   the aDNA simulated dataset (Appendix 2-figure 8).
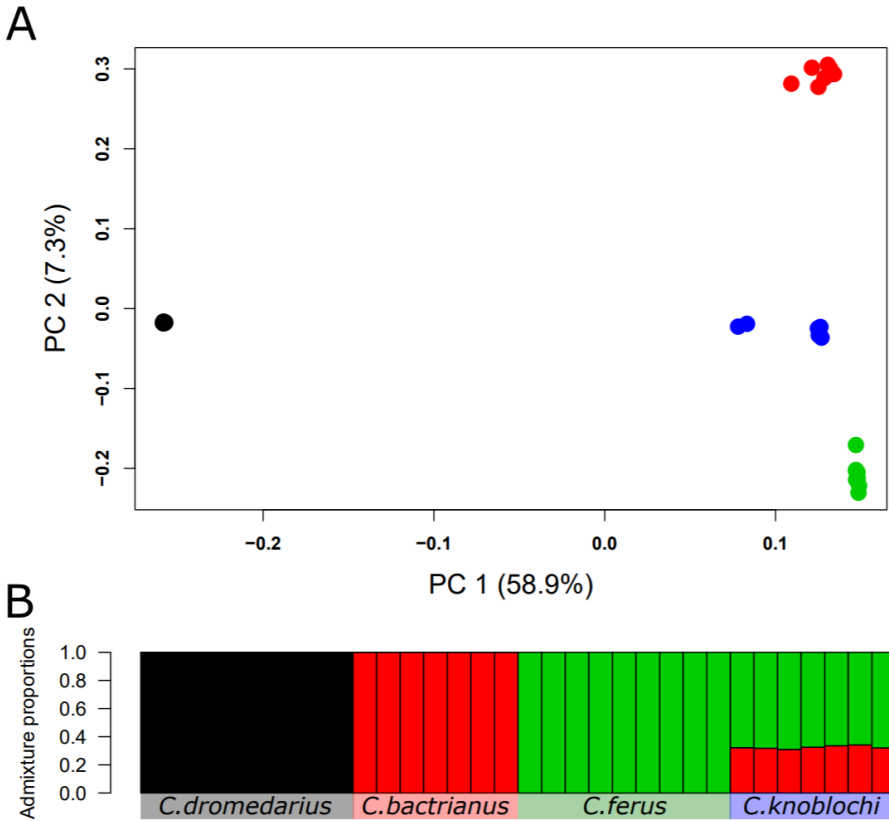
186

187

188   **Figure 2.** Nuclear population structure comparisons of four *Camelus* species. (A)

189   Principal component analysis (PCA) computed using genotype likelihoods and

190   representatives from all *Camelus* species analyzed in this study. Colours correspond to

191   each species: black - *C. dromedarius*, red - *C. bactrianus*, green - *C. ferus*, and blue -

192   *C. knoblochi.* (B) Admixture proportion analysis computed using genotype likelihoods

193   and an *a priori* population number (K) of three.

194

195   **Population dynamics**

10

196   Bayesian skyline analysis of wild Bactrian camel mitogenomes indicated a steep

197   decline in female effective population size in the last several hundred years, while the

198   ancestor of domesticated Bactrian camels showed population expansion during the past

199   80 thousand years. *C. knoblochi* kept a relatively steady maternal population size during

200   the Late Pleistocene, only observing a slight increase of 95% credibility intervals

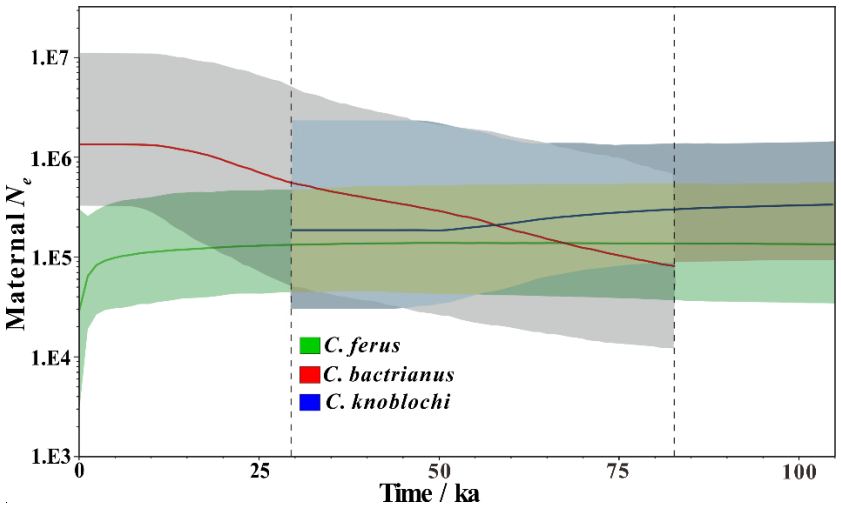201   between 55 and 50 ka (Figure 3).

202

203

Figure 3. Bayesian skyline plot of the two Bactrian camels and *C. knoblochi* individuals

based on near-complete mitochondrial genomes. Green, red and blue lines indicate

median female N_e of *C. ferus*, *C. bactrianus* and *C. knoblochi* changes over time.

Correspondingly, the grey, blue and green shaded areas indicate their 95% credibility

intervals.

Mitochondrial DNA Diversity

211   We also calculated genetic diversity of *Camelus* lineages (Table 1). The lowest

212   genetic diversity is observed in modern wild Bactrian camel, while the Late Pleistocene

11

213 *C. knoblochi* shows higher mitochondrial diversity compared to extant Old World

214 camels.

215

216 **Table 1.** Genetic diversity of *Camelus* lineages based on mitogenomes.

Taxon	*C. ferus*	*C. bactrianus*	*C. dromedarius*	*C. knoblochi*
Genetic diversity	0.001162	0.001298	0.001725	0.003002

217

Discussion

Phylogenetic position and origin for *Camelus knoblochi*

220 All fossils in this study are diagnostic and have been carefully evaluated in

221 morphology (Appendix 2-figures 9-13). They have generally larger body sizes than

222 extant *Camelus* materials. We confirm that size differences are not from the gender

223 biases as our molecular sex determination suggested that both female and male

224 individuals were included. Therefore, our samples are clearly attributed to *C. knoblochi*.

225

226 We reconstructed phylogenetic trees of camelids based on near-complete

227 mitochondrial genomes. Both Bayesian and ML approaches showed that our seven *C.*

228 *knoblochi* do not form a separate lineage but to cluster with the extant wild Bactrian

229 camel (Figure 1B; Appendix 2-figure 4). Such intermixed pattern is also observed in

230 other mammals, for example, polar vs. brown bear (Miller et al., 2012) and Tibetan wild

231 ass vs. Mongolian wild ass (Vilstrup et al., 2013). These phenomenon could be

232 explained by past hybridization events or incomplete lineage sorting, which might lead

233 to the discrepancies between molecular and morphological identification (Taron et al.,

234 2021). Therefore, it is imperative to also consider nuclear data as the mitochondrial

235 marker may not reflect the full picture of the species' evolutionary history (Miller et al.,

236 2012; Sheng et al., 2019).

237

238   Contrary to mitochondrial results, the nuclear analyses clearly suggest that *C.*

239 *knoblochi* is a unique group, which is closer to wild Bactrian camel than to domesticated

240 Bactrian camel (Figures 2A & 1B). Using simulated aDNA dataset, we verified this

241 result while considering potential biases linked to the low genomic coverage or DNA

242 damage (Appendix 2-figures 5-8). Further admixture proportion analyses revealed that

243 *C. knoblochi* shares ancestry from both wild Bactrian camels and the ancestor of

244 modern domesticated Bactrian camels. This extinct camel is mostly a hybrid between

245 the ancestors of two modern Bactrian camel species. The fact that our samples contain

246 unilateral matrilineage (Figure 1B) suggests *C. knoblochi* arose from mating of wild

247 Bactrian camel females and the male ancestors of modern domesticated Bactrian camel.

248 However, we cannot know if this is always the case as we covered limited geographical

249 sampling sites. Previous studies have indicated that hybridization is common in extant

250 *Camelus* species, for instance in modern domesticated Bactrian camel × wild Bactrian

251 camel (Felkel et al., 2019; Ming et al., 2020) and domesticated Bactrian camel ×

252 Arabian camel (Burger, 2016; Ming et al., 2020). Our analyses detect an ancient mixed

253 ancestry, which means hybridization may also be common to Pleistocene *Camelus*.

254 However, it is noteworthy that the earliest known remains of the two Bactrian camels

13

255   were limited in the Late Pleistocene, while the earliest *C. knoblochi* fossil recorders are

256   traced back to Middle Pleistocene (Titov, 2008). Further investigations on fossil

257   materials from the three *Camelus* species above are needed to fully understand their

258   evolutionary histories.

259

260   **Divergence time for *Camelus* species**

261   The family Camelidae originated in North America approximately 45-40 Ma

262   (Stanley et al., 1994). This was followed by the origin of the genus *Paracamelus*, most

263   likely ancestral to all *Camelus*, which dispersed from North America into Eurasia and

264   Africa via the Bering Strait around 6.3-5.8 Ma (van der Made et al., 2002). For all Old

265   World camels, *Camelus grattardi* is thought to be the earliest representative based on

266   remains from the Plio-Pleistocene strata (2.9-2.2 Ma) of Ethiopia (Geraads et al., 2019).

267   Using root-and-tip dating calibrations in BEAST (Figure 1B), we estimated the

268   subsequent split between Arabian and Bactrian camels around 2.44 Ma, much younger

269   than previous molecular dating (8-4.1 Ma) (Cui et al., 2007; Wu et al., 2014;

270   Heintzmann et al., 2015). However, Geraads et al. (2020) argued that the above

271   mentioned problematic calibration points probably led to overestimation and newly

272   suggested 2-1 Ma for that node using fossil data. As the fossil record tends to provide

273   minimal age estimates (Geraad et al., 2020), we confirm our estimation to be reasonable.

274

275   The divergence between the ancestors of domesticated and wild Bactrian camels

276   occurred around 0.8 Ma, indicating the temporal upper limit for their hybridization. The

277 earliest *C. knoblochi* fossils are dated back to the second half of the Middle Pleistocene

278 (0.45-0.12 Ma) (Titov, 2008) and the coalescence for all *C. knoblochi* here was traced

279 to approximately 0.33 Ma (Figure 1B). Hence, the now extinct *C. knoblochi* likely arose

280 due to a hybridization event occurred between 0.8 and 0.33 Ma (Figure 4). However, it

281 is still unclear where hybridization occurred and how *C. knoblochi* became a fixed

282 unique lineage. Genetic investigation into more Pleistocene camels should be essential

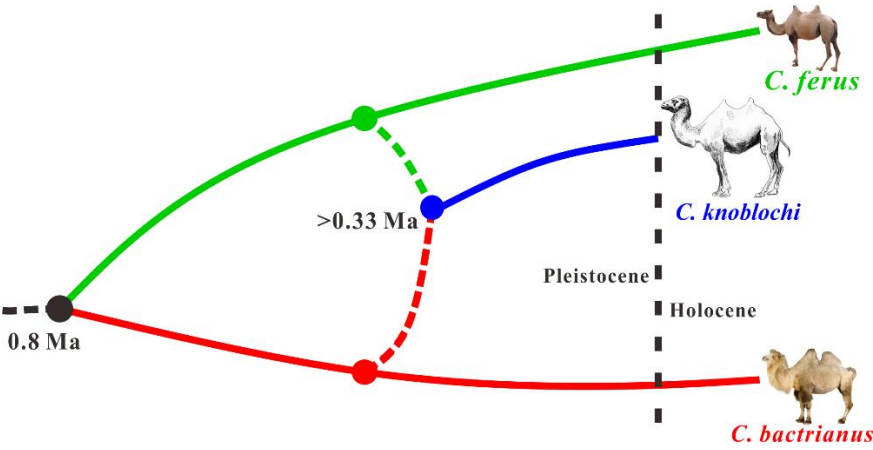283 to answer these questions.

284

285

286 **Figure 4.** Hypothesis for evolutionary outline of *C. knoblochi* and the two Bactrian

287 camels based on BEAST, PCA and admixture proportion analyses.

288

289 **Genetic diversities and demographic histories for *Camelus***

290 We investigated genetic diversity of different lineages within *Camelus* (Table 1).

291 Similar to previous studies (Mohandesan et al., 2017; Ming et al., 2020), wild Bactrian

292 camels display the lowest diversity value (0.001162). This is not unexpected as wild

293 Bactrian camels are suffering from a strong loss of genetic diversity in recent years,

294 which may be linked to geographic isolation and small population size. They are

15

295   currently limited to four isolated geographic regions (Silbermayr et al., 2009) and have

296   only ~1,000 individuals in China and Mongolia (Xue et al., 2015; Burger et al., 2019).

297   In addition, our BSP analysis revealed that the maternal population size of wild Bactrian

298   camels remained relatively stable since 100 ka, except a sharp decline in the last few

299   hundred years (Figure 3), possibly owing to human activities. However, the maternal

300   population size of domesticated Bactrian camels kept a continued growth during the

301   past 80 ka, reflecting that the ancestors of the wild and domesticated Bactrian camels

302   experienced different pressures and evolutionary histories during the Late Pleistocene.

303

304   The available Late Pleistocene *C. knoblochi* individuals exhibit higher genetic

305   diversity compared to extant *Camelus* (Table 1). This may be due to its hybrid origin or

306   might indicate a relatively larger population size during the Late Pleistocene. Fossil

307   records suggested that *C. knoblochi* was widely distributed in temperate and cold

308   steppes of Central and East Asia during the Late Pleistocene. Bayesian skyline

309   reconstruction showed the maternal effective population size of *C. knoblochi* lineage

310   remained generally stable during the Late Pleistocene, while later maintaining a large

311   population (Figure 3). The observed expansion could correspond to the early substage

312   of Guxiangtun periglacial stage (70-53 ka). The climate was very cold in Northeast

313   China during this period and the average annual temperature might be 10 °C lower than

314   that of today (Sun et al., 1985). After this cold stage, there is a slow population increase

315   of *C. knoblochi*. The demographic history of *C. knoblochi* may be further explored by

316   involving more individuals in future studies.

317

Methods

Sampling

A total of seven *Camelus* specimens were collected from three regions, i.e., Qinggang

County, Zhaodong County, and Haerbin City, in Heilongjiang Province, Northeastern

China (Figure 1A). Four of them (CADG387, CADG388, CADG389 and CADG596)

are stored in Zhaoyuan Museum, two samples (CADG529 and CADG533) in Daqing

Museum and one sample (CADG665) in Yifu Museum of China University of

Geosciences. All specimens are assigned to *C. knoblochi* based on morphological

characters. Detailed information of these specimens is shown in Appendix 1-table 1 and

Appendix 2-figures 9-13.

328

DNA extraction, library preparation and sequencing

DNA extraction and library preparation were all conducted in dedicated clean room

facilities at China University of Geoscience (Wuhan). Approximately 200 mg of each

specimen was ground into fine powder. DNA extraction was carried out as described

previously (Hu et al., 2021). DNA was eluted twice by adding EB buffer to produce 80

µL of final DNA extract for each sample. DNA extracts were converted into double-

stranded Illumina sequencing libraries following the protocol from Meyer and Kircher

(2010) with several modifications (Hu et al., 2021). Finally, indexed libraries were

quantified, pooled, and sequenced using an Illumina HiSeqX10 platform in a 2 × 150

bp paired-end run. Additionally, extraction and library blanks were included throughout

17

339   all steps to monitor potential contamination.

340

Mitochondrial data processing

342   Raw reads were trimmed of adapters using Cutadapt v1.4.2 (Martin, 2011), requiring

343   a minimum overlap of 1 bp, and discarding reads shorter than 30 bp. The trimmed reads

344   were then mapped to modified mitogenome reference sequences of wild Bactrian camel

345   (Genbank No. EF212038) and domesticated Bactrian camel (Genbank No. EF507799),

346   independently, using the "aln" algorithm in Burrows-Wheeler aligner (BWA) v0.6.2 (Li

347   and Durbin, 2010) with default parameters. Due to the circular state of the

348   mitochondrial genome, the modified references had 20 bp added to the ends from the

349   opposite edge. Reads with a MapQuality score less than 30 and designated as PCR

350   duplicates were removed using SAMtools v0.1.9 (Li et al., 2009). The final

351   mitochondrial consensus sequences were generated in Geneious

352   (https://www.geneious.com/), using a minimum coverage of 2 and a base agreement

353   greater than 75%.

354

Mitochondrial phylogenetic analysis

356   To investigate the phylogenetic position of *C. knoblochi*, a Bayesian phylogenetic

357   analysis was performed in BEAST v1.8.2 (Drummond et al., 2012) using the near-

358   complete mitochondrial genomes (Appendix 1-table 3). The dataset contains 7 *C.*

359   *knoblochi*, 15 *C. bactrianus*, 32 *C. ferus*, 10 *C. dromedarius*, 2 *Lama glama*, 3 *Lama*

360   *pacos*, 5 *Lama guanicoe*, 2 *Vicugna Vicugna,* 3 *Camelops* cf. *hesternus*, and one *Sus*

361   *scrofa* as the outgroup. A multiple sequence alignment was generated by using MAFFT

362   v7.471 (Katoh et al., 2002), and the problematic section of the control region was

363   removed. The final alignment of the combined dataset contains 16,680 nucleotide

364   positions. PartitionFinder v2.1.1 (Lanfear et al., 2016) was used to identify the optimal

365   partitioning scheme, and our dataset was divided into eight distinct partitions. A root-

366   and-tip-dating approach was used for the calibrated analysis. Firstly, based on the first

367   appearance of Lamini, the divergence time between Camelini and Lamini was set at

368   17.5 Ma with a standard deviation of 1.52 Ma. Secondly, tip dates were considered zero

369   for the modern individuals and the median radiocarbon or stratigraphic ages for the

370   ancient specimens (Appendix 1-tables 1, 3). As to the ancient specimens with no

371   stratigraphic information and also beyond the limits of radiocarbon dating, the median

372   age from BEAST estimation was used as the sample date (Appendix 1-table 1). The

373   BEAST analysis was conducted under a relaxed uncorrelated lognormal molecular

374   clock. A coalescent Bayesian skyline tree model and the GTR+G substitution model

375   were selected. Markov chain Monte Carlo chains were performed with 100,000,000

376   generations and sampled every 10,000 states. The results were analyzed using

377   TRACER v1.7 (Rambaut et al., 2018) to check the convergence of each parameter and

378   sufficient sampling (ESS>200). The first 25,000,000 of trees were discarded as burn-

379   in, the Maximum Clade Credibility tree was summarized using the TreeAnnotator

380   v1.5.4 (Drummond et al., 2012), which was visualized using FigTree v1.4.4

381   (http://tree.bio.ed.ac.uk/software/figtree). Moreover, Bayesian skyline analyses were

382   conducted to infer female effective population size changes over time in the wild

19

383 Bactrian camels, the ancestor of the modern domesticated Bactrian camels, and *C.*

384 *knoblochi* lineages using BEAST v1.8.2 (Drummond et al., 2012).

385

386   Genetic diversity of *C. knoblochi* and extant *Camelus* individuals (Appendix 1-table

387 3) was calculated using MEGA 7 software (Kumar et al., 2016), all positions containing

388 gaps or missing data were excluded and the final dataset had a length of 12,285 bp.

389

390 **Nuclear data processing**

391   For the ancient *C. knoblochi* samples, we followed the same mapping protocol as for

392 the mitochondrial genomes, except we used the wild Bactrian camel nuclear genome

393 (GCF_009834535.1) instead of the mitochondrial genome as reference. As a

394 comparative dataset, we downloaded the raw sequencing reads from nine Arabian

395 camel, seven domesticated Bactrian camel, and nine wild Bactrian camel (Fitak et al.,

396 2020). We merged overlapping paired end reads using FLASH v1.2.11 (Magoč and

397 Salzberg, 2011) utilizing default parameters. We mapped both merged and unmerged

398 reads to the available wild Bactrian camel reference nuclear genome using BWA

399 v0.7.15 (Li and Durbin, 2010) utilizing the mem algorithm and otherwise default

400 parameters. We parsed the alignment files and removed duplicates and reads of

401 mapping quality score <30 using SAMtools v1.6 (Li et al., 2009). Mapping statistics

402 can be found in Appendix 1- tables 2, 4.

403

404 **Sex determination**

405 We determined the sex of our *C. knoblochi* individuals by comparing the average

406 coverage of reads mapping to the X chromosome, to that of a chromosome of similar

407 length (chromosome 1). We did this using SAMtools depth and either specifying the X

408 chromosome or chromosome 1 using the parameter -b, and to include sites with no

409 coverage (-a). This method works as males should have half the amount of coverage on

410 the X chromosome compared to the autosomes due to only having a single copy of the

411 X chromosome but two copies of each autosome, whereas females should have

412 relatively similar coverages for both the X chromosome and autosomes.

413

414 **Nuclear population structure**

415 We evaluated the nuclear relationships between the four *Camelus* species included

416 in this study by performing a PCA, as well as estimating admixture proportions. We

417 performed these analyses twice, once using all four species, and once excluding Arabian

418 camel. Due to the very low coverage manner of our *C. knoblochi* individuals, we

419 performed these analyses using genotype likelihoods (GL) computed with ANGSD

420 v0.921 (Korneliussen et al., 2014). We computed GL for all individuals specifying the

421 parameters: minimum mapping and base qualities of 30 (-minmapQ 30 -minQ 30),

422 calculate genotype likelihoods using the GATK algorithm (-GL 2), output a beagle GL

423 file (-doGlf 2), calculate major and minor alleles based on GL (-doMajorMinor 1),

424 remove transitions (-rmtrans 1), only include SNPs with a p-values <1e-6 (-SNP_pval

425 1e-6), only consider autosomal chromosomes (-rf), a minimum minor allele frequency

426 of 0.05 (-minmaf 0.05), skip triallelic sites (-skiptriallelic 1), only consider reads

21

427   mapping to one region uniquely (-uniqueonly 1), only consider sites if at least 26

428   individuals have coverage (-minind 26). When calculating GL but excluding Arabian

429   camel we used the same parameters but only considered sites if at least 17 individuals

430   had coverage. To construct a covariance matrix for PCA we used PCAngsd v0.98

431   (Meisner and Albrechtsen, 2018). To estimate ancestry components we used NGSadmix

432   v32 (Skotte et al., 2013) specifying the number of *a priori* populations (K) as 3 or 4. To

433   evaluate the reliability of our NGSadmix results, we repeated the NGSadmix analysis

434   up to 100 times, printing the likelihood each time. If the best likelihood was obtained

435   twice, the results were considered to have converged and were therefore reliable. If this

436   was not obtained within 100 replicates, then the reliability of the results could not be

437   determined.

438

439   **Palaeogenomic data reliability assessment**

440   Due to the large disparity between data quality and coverage of the modern and

441   ancient datasets, we assessed whether our results could be data driven as opposed to

442   biological. To test this we randomly selected two individuals from each of the extant

443   species, reduced their raw read lengths down to 40 bp, and added aDNA damage

444   patterns using TAPAS v1.2 (Taron et al., 2018) and the mapdamage output from our *C.*

445   *knoblochi* individual with the highest coverage (CADG533). We mapped the artificially

446   damaged reads to the wild Bactrian camel reference genome as was done with the

447   empirical ancient data and subsampled the resultant bam files to ~0.07x with SAMtools.

448   We reran the PCA and admixture proportion analyses with the aDNA simulated data for

22

449   these individuals instead of the original high quality data. To ensure comparability, we

450   used the same sites as we had recovered with the complete dataset.

451

452   **References**

453   Almathen F, Charruau P, Mohandesan E, Mwacharo JM, Orozco-terWengel P, Pitt D,

454   Abdussamad AM, Uerpmann M, Uerpmann H, De Cupere B, Magee P, Alnaqeeb

455   MA, Salim B, Raziq A, Dessie T, Abdelhadi OM, Banabazi MH, Al-Eknah M,

456   Walzer C, Faye B, Hofreiter M, Peters J, Hanotte O, Burger PA. 2016. Ancient and

457   modern DNA reveal dynamics of domestication and cross-continental dispersal of

458   the dromedary. *PNAS* **113**:6707-6712. DOI:

459   https://doi.org/10.1073/pnas.1519508113, PMID: 27162355

460   Burger PA. 2016. The history of Old World camelids in the light of molecular genetics.

461   *Tropical Animal Health and Production* **48**:905-913. DOI:

462   https://doi.org/10.1007/s11250-016-1032-7, PMID: 27048619

463   Burger PA, Ciani E, Faye B. 2019. Old World camels in a modern world-a balancing

464   act between conservation and genetic improvement. *Animal Genetics* **50**:598-612.

465   DOI: https://doi.org/10.1111/age.12858, PMID: 31532019

466   Chen Q. 2001. Paleolithic site unearthed at Dakanzi, Dandai, Zhenlai, Jilin. *Northern*

467   *Cultural Relics* **2**:1-7.

468   Chen S, Li J, Zhang F, Xiao B, Hu J, Cui Y, Hofreiter M, Hou X, Sheng G, Lai X, Yuan

469   J. 2019. Different maternal lineages revealed by ancient mitochondrial genome of

470   *Camelus bactrianus* from China. *Mitochondrial DNA Part A* **30**:786-793. DOI:

471   https://doi.org/10.1080/24701394.2019.1659250, PMID: 31542986

472   Cui P, Ji R, Ding F, Qi D, Gao H, Meng H, Yu J, Hu S, Zhang H. 2007. A complete

473   mitochondrial genome sequence of the wild two-humped camel (*Camelus*

474   *bactrianus ferus*): an evolutionary history of Camelidae. *BMC Genomics* **8**:241.

475   DOI: https://doi.org/10.1186/1471-2164-8-241, PMID: 17640355

476   Díaz-Maroto P, Rey-Iglesia A, Cartajena I, Núñez L, Westbury MV, Varas V, Moraga

477   M, Campos PF, Orozco-terWengel P, Marín JC, Hansen AJ. 2021. Ancient DNA

478   reveals the lost domestication history of South American camelids in Northern

479   Chile   and   across   the   Andes.   *eLife*   **10**:e63390.   DOI:

480   https://doi.org/10.7554/eLife.63390, PMID: 33724183

481   Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with

482   BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**:1969-1973.

483   DOI: http://doi.org/10.1093/molbev/mss075, PMID: 22367748

484   Felkel S, Wallner B, Chuluunbat B, Yadamsuren A, Faye B, Brem G, Walzer C, Burger

485   PA. 2019. A first Y-chromosomal haplotype network to investigate male-driven

486   population dynamics in domestic and wild Bactrian camels. *Frontiers in Genetics*

487   **10**:423. DOI: https://doi.org/10.3389/fgene.2019.00423, PMID: 31178891

488   Fitak RR, Mohandesan E, Corander J, Burger PA. 2016. The *de novo* genome assembly

489   and annotation of a female domestic dromedary of North African origin.

490   *Molecular Ecology Resources* **16**:314-324. DOI: http://doi.org/10.1111/1755-

491   0998.12443, PMID: 26178449

492   Fitak RR, Mohandesan E, Corander J, Yadamsuren A, Chuluunbat B, Abdelhadi O,

493   Raziq A, Nagy P, Walzer C, Faye B, Burger PA. 2020. Genomic signatures of

494   domestication in Old World camels. *Communications Biology* **3**:316. DOI:

495   http://doi.org/10.1038/s42003-020-1039-5, PMID: 32561887

496 Geraads D, Barr WA, Reed D, Laurin M, Alemseged Z. 2019. New remains of *Camelus*

497   *grattardi* (Mammalia, Camelidae) from the Plio-Pleistocene of Ethiopia and the

498   phylogeny of the genus. *Journal of Mammalian Evolution*. DOI:

499   https://doi.org/10.1007/ s10914-019-09489-2.

500 Geraads D, Didier G, Barr WA, Reed D, Laurin M. 2020. The fossil record of camelids

501   demonstrates a late divergence between Bactrian camel and dromedary. *Acta*

502   *Palaeontologica   Polonica   **65**:251-260.   DOI:

503   https://doi.org/10.4202/app.00727.2020

504 Heintzman PD, Zazula GD, Cahill JA, Reyes AV, MacPhee RDE, Shapiro B. 2015.

505   Genomic data from extinct North American *Camelops* revise camel evolutionary

506   history. *Molecular Biology and Evolution* **32**:2433-2440. DOI:

507   https://doi.org/10.1093/molbev/msv128, PMID: 26037535

508 Hu J, Westbury MV, Yuan J, Zhang Z, Chen S, Xiao B, Hou X, Ji H, Lai X, Hofreiter

509   M, Sheng G. 2021. Ancient mitochondrial genomes from Chinese cave hyenas

510   provide insights into the evolutionary history of the genus *Crocuta*. *Proceedings*

511   *of   the   Royal   Society   B   **288**:20202934.   DOI:

512   https://doi.org/10.1098/rspb.2020.2934, PMID: 33499784

513 Ji R, Cui P, Ding F, Geng J, Gao H, Zhang H, Yu J, Hu S, Meng H. 2009. Monophyletic

514   origin of domestic Bactrian camel (*Camelus bactrianus*) and its evolutionary

515   relationship with the extant wild camel (*Camelus bactrianus ferus*). *Animal*

516   *Genet*ics **40**:377-82. DOI: https://doi.org/10.1111/j.1365-2052.2008.01848.x,

517   PMID: 19292708

518   Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid

519   multiple sequence alignment based on fast Fourier transform. *Nucleic Acids*

520   *Research* **30**:3059-3066. http://doi.org/10.1093/nar/gkf436, PMID: 12136088

521   Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: analysis of next generation

522   sequencing data. *BMC Bioinformatics* **15**:356. DOI:

523   https://doi.org/10.1186/s12859-014-0356-4, PMID: 25420514

524   Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics

525   analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*

526   **33**:1870-1874. DOI: https://doi.org/10.1093/molbev/msw054, PMID: 27004904

527   Lado S, Elbers JP, Doskocil A, Scaglione D, Trucchi E, Banabazi MH, Almathen F,

528   Saitou N, Ciani E, Burger PA. 2020. Genome-wide diversity and global migration

529   patterns in dromedaries follow ancient caravan routes. *Communications Biol*ogy

530   **3**:387. DOI: https://doi.org/10.1038/s42003-020-1098-7, PMID: 32678279

531   Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2016. PartitionFinder 2:

532   new methods for selecting partitioned models of evolution for molecular and

533   morphological phylogenetic analyses. *Molecular Biology and Evolution* **34**:772-

534   773. DOI: http://dx.doi.org/10.1093/molbev/msw260, PMID: 28013191

535   Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler

536   transform. *Bioinformatics* **26**:589-595. DOI:

537   http://doi.org/10.1093/bioinformatics/btp698, PMID: 20080505

538 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G,

539   Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence

540   alignment/map format and samtools. *Bioinformatics* **25**:2078-2079. DOI:

541   https://doi.org/10.1093/bioinformatics/btp352, PMID: 19505943

542 Magoč T, Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve

543   genome   assemblies.   *Bioinformatics*   **27**:2957-2963.   DOI:

544   https://doi.org/10.1093/bioinformatics/btr507, PMID: 21903629

545 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput

546   sequencing   reads.   *EMBnet*   *Journal*   **17**:10-12.   DOI:

547   https://doi.org/10.14806/ej.17.1.200

548 Meisner J, Albrechtsen A. 2018. Inferring population structure and admixture

549   proportions   in   low-depth   NGS   data.   *Genetics*   **210**:719-731.   DOI:

550   https://doi.org/10.1534/genetics.118.301336, PMID: 30131346

551 Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly

552   multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*

553   **2010**:pdb prot5448. DOI: https://doi.org/10.1101/pdb.prot5448, PMID: 20516186

554 Miller W, Schuster SC, Welch AJ, Ratan A, Bedoya-Reina OC, Zhao F, Kim HL,

555   Burhans RC, Drautz DI, Wittekindt NE, Tomsho LP, Ibarra-Laclette E, Herrera-

556   Estrella L, Peacock E, Farley S, Sage GK, Rode K, Obbard M, Montiel R,

557   Bachmann L, Ingólfsson Ó, Aars J, Mailund T, Wiig Ø, Talbot SL, Lindqvist C.

558   2012. Polar and brown bear genomes reveal ancient admixture and demographic

559   footprints of past climate change. *PNAS* **109**:E2382-E2390. DOI:

560   https://doi.org/10.1073/pnas.1210506109, PMID: 22826254

561 Ming L, Yuan L, Yi L, Ding G, Hasi S, Chen G, Jambl T, Hedayat-Evright N, Batmunkh

562   M, Badmaevna GK, Gan-Erdene T, Ts B, Zhang W, Zulipikaer A, Hosblig, Erdemt,

563   Natyrov A, Mamay P, Narenbatu, Meng G, Narangerel C, Khongorzul O, He J,

564   Hai L, Lin W, Sirendalai, Sarentuya, Aiyisi, Li Y, Wang Z, Jirimutu. 2020. Whole-

565   genome sequencing of 128 camels across Asia reveals origin and migration of

566   domestic Bactrian camels. *Communications Biol*ogy **3**:1-9. DOI:

567   https://doi.org/10.1038/s42003-019-0734-6, PMID: 31925316

568 Mohandesan E, Fitak RR, Corander J, Yadamsuren A, Chuluunbat B, Abdelhadi O,

569   Raziq A, Nagy P, Stalder G, Walzer C, Faye B, Burger PA. 2017. Mitogenome

570   sequencing in the genus *Camelus* reveals evidence for purifying selection and

571   long-term divergence between wild and domestic Bactrian camels. *Scientific*

572   *Rep*ort **7**:9970. DOI: https://doi.org/10.1038/s41598-017-08995-8, PMID:

573   28855525

574 Orlando L. 2016. Back to the roots and routes of dromedary domestication. *PNAS*

575   **113**:6588-6590. DOI: https://doi.org/10.1073/pnas.1606340113, PMID:

576   27226295

577 Peters J, von den Driesh A. 1997. The two-humped camel (*Camelus bactrianus*): new

578   light on its distribution, management and medical treatment in the past. *Journal*

579   *of Zoology* **242**:651-679. DOI: https://doi.org/10.1111/j.1469-

580   7998.1997.tb05819.x

581   Rambaut A, Suchard MA, Xie D, Baele G, Suchard MA. 2018. Posterior

582   summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*

583   **67**:901-904. DOI: https://doi.org/10.1093/sysbio/syy032, PMID: 29718447

584   Rowan J, Martini P, Likius A, Merceron G, Boisserie JR. 2018. New Pliocene remains

585   of *Camelus grattardi* (Mammalia, Camelidae) from the Shungura Formation,

586   Lower Omo Valley, Ethiopia, and the evolution of African camels. *Historical*

587   *Biology* **31**:1123-1134. DOI: https://doi.org/10.1080/08912963.2017.1423485

588   Sheng G, Basler N, Ji X, Paijmans JLA, Alberti F, Preick M, Hartmann S, Westbury

589   MV, Yuan J, Jablonski NG, Xenikoudakis G, Hou X, Xiao B, Liu J, Hofreiter M,

590   Lai X, Barlow A. 2019. Paleogenome reveals genetic contribution of extinct giant

591   panda to extant populations. *Current Biology* **29**:1-6. DOI:

592   https://doi.org/10.1016/j.cub.2019.04.021, PMID: 31080081

593   Silbermayr K, Orozco-terWengel P, Charruau P, Enkhbileg D, Walzer C, Vogl C,

594   Schwarzenberger F, Kaczensky P, Burger PA. 2009. High mitochondrial

595   differentiation levels between wild and domestic Bactrian camels: a basis for rapid

596   detection of maternal hybridization. *Animal Genetics* **41**:315-318. DOI:

597   https://doi.org/10.1111/j.1365-2052.2009.01993.x, PMID: 19968638

598   Skotte L, Korneliussen TS, Albrechtsen A. 2013. Estimating individual admixture

599   proportions from next generation sequencing data. *Genetics* **195**:693-702. DOI:

600   https://doi.org/10.1534/genetics.113.154138, PMID: 24026093

601   Stanley HF, Kadwell M, Wheeler JC. 1994. Molecular evolution of the family

602   Camelidae: a mitochondrial DNA study. *Proceedings of the Royal Society B* **256**:1-

603   6. DOI: https://doi.org/10.2307/49586, PMID: 8008753

604 Sun J, Wang S, Wang Y, Zhou Y, Lin Z, Chen S. 1985. Paleoenvironment of the last

605   glacial stage in Northeast China. *Quaternary Science* **6**:82-89.

606 Taron UH, Lell M, Barlow A, Paijmans JLA. 2018. Testing of alignment parameters for

607   ancient samples: eevaluating and optimizing mapping parameters for ancient

608   samples using the TAPAS tool. *Genes* **9**:157. DOI:

609   https://doi.org/10.3390/genes9030157, PMID: 29533977

610 Taron UH, Paijmans JLA, Barlow A, Preick M, Iyengar A, Drăgusin V, Vasile S,

611   Marciszak A, Roblícková M, Hofreiter M. 2021. Ancient DNA from the Asiatic

612   Wild Dog (*Cuon alpinus*) from Europe. *Genes* **12**:144. DOI: https://doi.org/

613   10.3390/genes12020144, PMID: 33499169

614 Tibary A, El Allali K. 2020. Dromedary camel: a model of heat resistant livestock

615   animal. *Theriogenology* **154**:203-211. DOI:

616   https://doi.org/10.1016/j.theriogenology.2020.05.046, PMID: 32663620

617 Titov VV. 2008. Habitat conditions for *Camelus knoblochi* and factors in its extinction.

618   *Quaternary International* **179**:120-125. DOI:

619   https://doi.org/10.1016/j.quaint.2007.10.022

620 Tulgat R, Schaller GB. 1992. Status and distribution of wild Bactrian camels *Camelus*

621   *bactrianus ferus*. *Biological Conservation* **62**:11-19. DOI:

622   http://doi.org/10.1016/0006-3207(92)91147-k

623 van der Made J, Morales J, Sen S, Aslan F. 2002. The first camel from the Upper

624   Miocene of Turkey and the dispersal of the camels into the Old World. *Comptes*

30

625   *Rendus Palevol* **1**:117-122. DOI: http://doi.org/10.1016/S1631-0683(02)00012-X

626   Vilstrup JT, Seguin-Orlando A, Stiller M, Ginolhac A, Raghavan M, Nielsen SCA,

627   Weinstock J, Froese D, Vasiliev SK, Ovodov ND, Clary J, Helgen KM, Fleischer

628   RC, Cooper A, Shapiro B, Orlando L. 2013. Mitochondrial phylogenomics of

629   modern and ancient equids. *PLoS ONE* **8**:e55950. DOI:

630   https://doi.org/10.1371/journal.pone.0055950, PMID: 23437078

631   Wu H, Guang X, Al-Fageeh MB, Cao J, Pan S, Zhou H, Zhang L, Abutarboush MH,

632   Xing Y, Xie Z, Alshanqeeti AS, Zhang Y, Yao Q, Al-Shomrani BM, Zhang D, Li J,

633   Manee MM, Yang Z, Yang L, Liu Y, Zhang J, Altammami MA, Wang S, Yu L,

634   Zhang W, Liu S, Ba L, Liu C, Yang X, Meng F, Wang S, Li L, Li E, Li X, Wu K,

635   Zhang S, Wang J, Yin Y, Yang H, Al-Swailem AM, Wang J. 2014. Camelid

636   genomes reveal evolution and adaptation to desert environments. *Nature*

637   *Communications* **5**:1-9. DOI: http://doi.org/10.1038/ncomms6188, PMID:

638   25333821

639   Xue Y, Li D, Xiao W, Liu F, Zhang Y, Wang X, Jia H. 2015. Activity patterns of wild

640   Bactrian camels (*Camelus bactrianus*) in the northern piedmont of the Altun

641   Mountains, China. *Animal Biology* **65**:209-217. DOI:

642   https://doi.org/10.1163/15707563-00002471

643   Zhang C, Ren Z, Yang X. 2015. Genetic diversity of mtDNA Cyt b gene and D-loop

644   sequence of Sunite domestic Bactrian camel. *Journal of Northwest A&F*

645   *University* **43**:32-42. DOI: https://doi.org/10.13207/j.cnki.jnwafu.2015.03.029

646

Acknowledgements

Author contributions

Junxia Yuan, Xulong Lai and Guilian Sheng conceived the study; Shungang Chen, Linying Wang, Jiaming Hu, Bo Xiao and Xindong Hou performed the experiments; Junxia Yuan, Guilian Sheng and Michael V. Westbury guided the experiment and bioinformatics analyses; Shungang Chen, Michael V. Westbury, Linying Wang and Junxia Yuan analyzed the data; Fengli Zhang, Siren Wang, Zhen Zhang and Fuqiang Li helped to collect the samples; Wenhui Liu carried out morphological analyses of the samples; Junxia Yuan, Michael V. Westbury, Guilian Sheng, Wenhui Liu, Jiaming Hu, Shungang Chen, Linying Wang and Xulong Lai wrote the paper. All authors read and gave comments to the final version of the manuscript.

Data availability

The obtained mitochondrial genomes reported in this article have been deposited in the Genbank under the accession nos. MZ430515- MZ430521.

Supplementary Materials

Appendix 1-table 1. The information of the *C. knoblochi* specimens analyzed in this

32

669   study.

670   **Appendix 1-table 2.** Raw reads mapping results of the *C. knoblochi* specimens

671   analyzed in this study.

672   **Appendix 1-table 3.** Sequences downloaded from NCBI for mitochondrial data

673   analyses in this study.

674   **Appendix 1-table 4.** Raw reads mapping statistic results of modern Old World camels

675   download from NCBI.

676   **Appendix 2-figure 1.** Cytosine deamination frequency inferred from *C. knoblochi*

677   samples analyzed in this study.

678   **Appendix 2-figure 2.** Estimated endogenous fragment length distributions for *C.*

679   *knoblochi* samples analyzed in this study.

680   **Appendix 2-figure 3.** Coverage plots of the mitochondrial genomes obtained in this

681   study.

682   **Appendix 2-figure 4.** Maximum-likelihood phylogenetic tree of camelids based on

683   complete mitochondrial genomes.

684   **Appendix 2-figure 5.** Principal component analysis computed using genotype

685   likelihoods with *C. dromedarius*.

686   **Appendix 2-figure 6.** Principal component analysis computed using genotype

687   likelihoods without *C. dromedarius*.

688   **Appendix 2-figure 7.** Admixture proportions analyses computed using genotype

689   likelihoods and an *a priori* population numbers (K) of three and four.

690   **Appendix 2-figure 8.** Admixture proportions analysis computed using genotype

691   likelihoods and an *a priori* population number (K) of four with the complete

692   dataset.

693   **Appendix 2-figure 9.** The axial skeleton of *C. knoblochi* from Northeastern China.

694   **Appendix 2-figure 10.** The complete forearm of *C. knoblochi* from Northeastern China.

695   **Appendix 2-figure 11.** The incomplete forearm of *C. knoblochi* from Northeastern

696   China.

697   **Appendix 2-figure 12.** The incomplete femur of *C. knoblochi* from Northeastern China.

698   **Appendix 2-figure 13.** The metatarsus of *C.knoblochi* from Northeastern China.