Mitochondrial DNA variation of the striped hyena (*Hyaena hyaena*) in Algeria and further insights into the species' evolutionary history

Authors: Louiza Derouiche[a*] (0000-0002-9379-9514), Mónica Rodrigues[b,c] (0000-0001-9059-0216), Hafida Benameur-Hasnaoui[d], Ridah Hadj Aissa[e], Yasaman Hassan-Beigi[b], Seyed Massoud Madjdzadeh[f] (0000-0002-1505-5456), Zuhair Amr[g] (0000-0002-4778-2308), Aimee Cokayne[h], Paul Vercammen[i] (0009-0004-1568-7463), Carlos Rodríguez Fernandes[b,c*] (0000-0002-1386-8103)

Affiliations: [a]École Supérieure des Sciences de l'Aliment et Industries Agroalimentaires (ESSAIA), Avenue Ahmed Hamidouche, route de Beaulieu, El Harrach, 16200 Algiers, Algeria; [b]CE3C - Centre for Ecology, Evolution and Environmental Changes & CHANGE - Global Change and Sustainability Institute, Departamento de Biologia Animal, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal; [c]Faculdade de Psicologia, Universidade de Lisboa, Alameda da Universidade, 1649-013 Lisboa, Portugal; [d]Parc National de Tlemcen, Plateau Lalla Setti, 13000 Tlemcen, Algeria; [e]Département des Sciences Vétérinaires de l'Université de Blida, BP 27, 09000 Blida, Algeria; [f]Zoological Museum and Department of Biology, Faculty of Sciences, Shahid Bahonar University of Kerman, 76169-14111 Kerman, Iran; [g]Department of Biology, Jordan University of Science and Technology, Jordan; [h]Sir Bani Yas Island, P.O Box 126888, Abu Dhabi, United Arab Emirates; [i]Breeding Centre for Endangered Arabian Wildlife, P.O. Box 29922, Sharjah, United Arab Emirates

25 *Corresponding authors

26 E-mail: derouiche_fatma@yahoo.fr; cafernandes@fc.ul.pt

27

28 **Abstract**

29 The striped hyena (*Hyaena hyaena*) occurs in a wide range from north and east
30 Africa, through southwest Asia to India, but its distribution is increasingly patchy
31 and many of its populations are in decline due to intense human pressure. Its
32 genetic diversity and structure, phylogeography, and evolutionary history,
33 remain poorly understood. In this study, we investigated mitochondrial DNA
34 variation in Algerian striped hyenas. Moreover, with the aim of contributing to
35 our understanding of the evolutionary history of the species, we also examined
36 samples from other geographic regions and compared our results with those of
37 the only previous study in which individuals from across the range of the species
38 were analysed. In particular, we performed a wide range of analyses of
39 demographic history and estimation of the age of the extant mitochondrial DNA
40 variation. The Algerian population sample was monomorphic. Overall, the global
41 patterns of genetic diversity and the results of some demographic history
42 analyses supported a scenario of population growth in the species, estimated to
43 have occurred in the Late Pleistocene, but many of the analyses did not detect a
44 significant signal of growth, most likely a result of the limited power provided by
45 a small number of segregating sites. The estimates, from three different methods,
46 for the time to the most recent common ancestor (TMRCA) of the mitochondrial
47 DNA variation hovered around 400 ka, coinciding with one of the longest and

warmest interglacials of the last 800,000 years, with environmental conditions similar to the Holocene.

Keywords: hyena, phylogeography, demographic history, TMRCA, subspecies, hyaenid

Introduction

The striped hyena *Hyaena hyaena* (Linnaeus, 1758), the only extant representative of its genus (Werdelin and Solounias 1991; Koepfli et al. 2006), is a large mammalian carnivore distributed from north and east Africa, through southwest Asia to India (Rieger 1981; Mills and Hofer 1998; Kingdon 2001; Holekamp and Kolowski 2009; Abi-Said and Dloniak 2015). Although once abundant and widespread across this range, its distribution is now mostly patchy, with the species extinct in many areas and the remaining populations generally in decline (Mills and Hofer 1998; Low Cunningham 2004; Kasparek et al. 2004; Khorozyan et al. 2011; Tourani et al. 2012; Abi-Said and Dloniak 2015; Akash et al. 2021; Bhandari et al. 2021; Çoğal et al. 2021). The global population is estimated at less than 10,000 individuals and, given the intense human pressure (persecution, poaching, poisoning, conflict with pastoralists, target of negative feelings and superstitious fears, prey base depletion, use as folk medicine, road mortality) to which it is subject, it is listed as 'Near Threatened' on the International Union for the Conservation of Nature (IUCN) Red List of Threatened Species (Abi-Said and Dloniak 2015).

The striped hyena is one of the four living species in the family Hyaenidae, the other three being the brown hyena (*Parahyaena brunnea*), the spotted hyena (*Crocuta crocuta*), and the aardwolf (*Proteles cristata*). The latter represents the oldest lineage of extant hyaenids, being quite distinct from the other three (Werdelin and Solounias 1991; Jenks and Werdelin 1998; Koepfli et al. 2006; Westbury et al. 2019). It has been estimated that the lineage ancestral to *Crocuta* diverged from the lineage ancestral to *Hyaena* and *Parahyaena* 9-11 Ma (Koepfli

93 et al. 2006; Westbury et al. 2019, 2020), while the latter two lineages diverged
94 from each other about 4-5 Ma (Rohland et al. 2005; Koepfli et al. 2006; Hu et al.
95 2021; Westbury et al. 2021). Traditionally, based on differences in body size and
96 skull and pelage characteristics (Pocock 1934), the following five subspecies
97 were accepted (Rieger 1979, 1981; Mills and Hofer 1998; Kasparek et al. 2004):
98 *barbara* in Northwest Africa (type locality: Oran, Algeria); *dubbah* in East Africa
99 (type locality: Atbara, Sudan); *syriaca* in Syria and Anatolia (type locality:
100 Antakya, Turkey); *sultana* in southern Arabia (type locality: Mt. Qara, Oman); and
101 *hyaena* from Iran to India (type locality: Benna Mountains, Larestan, Iran).
102 However, the distributions and limits of these classically recognized subspecies
103 are not precisely known (Rieger 1981), and attempts to delineate their
104 geographic ranges and boundaries should be seen as tentative (Mills and Hofer
105 1998). Jenks and Werdelin (1998) noted that molecular data could be very useful
106 for assessing subspecies designations in the striped hyena. In fact, a study that
107 sequenced a 340-bp fragment of the mitochondrial Cytochrome b (*Cyt b*) gene in
108 13 striped hyena samples from scattered locations throughout the historical and
109 current distribution of the species, found low genetic diversity (four haplotypes
110 with two variable sites) and a general lack of obvious phylogeographic structure,
111 with three of the four haplotypes shared between Africa and Eurasia (Rohland et
112 al. 2005). Thus, although the species exhibits marked geographic variation in size
113 and coat colour patterns, the currently accepted view is that there is no evidence
114 for the recognition of subspecies (Kingdon 2001; Wozencraft 2005).

115 In Algeria, the striped hyena had the same history of decline over the past few
116 centuries as observed in many other regions of its distribution, and the species
117 remains vulnerable and threatened by anthropogenic causes of mortality and in

118 need of conservation efforts, but it has managed to persist throughout much of
119 the country (Kowalski and Rzebik-Kowalska 1991; Derouiche et al. 2020).

120 In this study, we sequenced a large sample of striped hyenas from Algeria (i.e.,
121 representing the traditional subspecies *barbara*) for a *Cyt b* fragment containing
122 the shorter region analysed by Rohland et al. (2005) to allow comparisons. We
123 also examined samples from geographic regions and traditional subspecies not
124 included in that study, such as southern Arabia (subspecies *sultana*) and
125 southern Iran (subspecies *hyaena*). Our objectives were to assess mitochondrial
126 DNA (mtDNA) diversity of striped hyenas in Algeria and, expanding on the
127 pioneering and seminal study of Rohland et al. (2005), contribute to increase our
128 knowledge of the evolutionary history of the species.

129

130 **Materials and methods**

131 *Sampling and laboratory procedures*

132 To investigate the mtDNA diversity of the striped hyena in Algeria, we collected
133 samples from 25 specimens found dead on roads or in the field. We also analysed
134 samples of striped hyenas from Jordan (2), Saudi Arabia (3), Oman (5), and Iran
135 (2) (Fig. 1 and Table 1). DNA was extracted from muscle tissue and blood
136 samples using the EZNA Tissue DNA kit (Omega Bio-Tek), while that from dry
137 hair samples was extracted following a simple polymerase chain reaction (PCR)
138 buffer-based extraction protocol (Allen et al. 1998; Vigilant 1999). For every
139 batch of samples, we used DNA extraction blanks to monitor contamination. We
140 amplified a fragment containing the first 753 base pairs (bp) of the *Cyt b* gene
141 with the primers L14724 (5'- TGATATGAAAAACCATCGTTG-3'; Irwin et al. 1991)

142 and H15791 (5'- AATGTAGTTGTCTGGGTC-3'; Fernandes et al. 2008). For four
143 hair samples (one from Algeria, one from Saudi Arabia, and the two from Jordan)
144 we were unsuccessful in amplifying this fragment, which we interpreted as a
145 result of DNA degradation. Thus, for those samples we amplified a 363-bp part of
146 the fragment (corresponding to positions 16-378 of *Cyt b*), containing the 340-bp
147 region analysed by Rohland et al. (2005), in two overlapping pieces, using
148 respectively the two following primer pairs: Hyena Cytb F1 (5'-
149 CCAATGACCAACATTCGA-3')/Hyena Cytb R1 (5'-TCAGCCATAGTTGACGTC-3')
150 (for a fragment of 198 bp) and Hyena Cytb F2 (5'-AACCGCCTTTTCATCAGT-
151 3')/Hyena Cytb R2 (5'-GACGTAACCTATGAATGC-3') (for a fragment of 181 bp).
152 PCRs were carried out in volumes of 15 µl with 1x PCR Buffer (NZYTech), 2 mM
153 $MgCl_2$, 0.2 mM of each dNTP (Bioline), 0.5 µM of each primer, 0.75 U of Supreme
154 NZYTaq DNA Polymerase (NZYTech), and 3 µl of DNA extract. DNA extraction
155 and PCR blanks were included to check for contamination. Thermal cycling
156 conditions consisted of an initial denaturation at 95 °C for 5 min, followed by 45
157 cycles of 30 s at 94 °C, 30 s at 50 °C, 30 s at 72 °C, and a final extension of 7 min at
158 72 °C. The results of the PCR amplifications were visualized on 2% agarose gels
159 to verify PCR quality and absence of contamination, and the PCR products were
160 purified with an Exo-SAP protocol (Hanke and Wink 1994; Werle et al. 1994)
161 and sequenced by the Sanger method at Macrogen Inc.

162

163 *Data analyses*

164 Sequences were edited, assembled, aligned, and checked for the absence of
165 indels and stop codons using SEQUENCHER 4.7 (Gene Codes Corporation) and

166 the Translate program of the SEQUENCE MANIPULATION SUITE 2 (Stothard
167 2000). We downloaded striped hyena *Cyt b* sequences from GenBank (mainly the
168 sequences from Rohland et al. (2005) and 39 sequences of Iranian individuals
169 from Dadashi-Jourdehi et al. (2022)) and included these in our alignments (Table
170 1). Trimming our sequences to 648 bp allowed us to include those Iranian
171 sequences from GenBank, but in order to incorporate other shorter sequences
172 from diverse geographic origins, in particular those from Rohland et al. (2005),
173 we created a smaller alignment of 340 bp (Table 1). Sequence alignments were
174 analysed with FABOX 1.5 (Villesen 2007) to identify samples with identical
175 sequences. File format conversions of sequence alignments for use in different
176 computer programs were done using ALTER (Glez-Peña et al. 2010).

177 To characterize patterns of mitochondrial genetic diversity, and what they
178 suggest in terms of demographic history of the striped hyena, we computed the
179 number of polymorphic or segregating sites (S), the number of haplotypes (n_H),
180 haplotype diversity (h), nucleotide diversity (π), the mismatch distribution
181 (Rogers and Harpending 1992; Schneider and Excoffier 1999), Tajima's *D*
182 (Tajima 1989), Fu's F_S (Fu 1997) and the R_2 statistic (Ramos-Onsins and Rozas
183 2002) in ARLEQUIN 3.5.2.2 (Excoffier and Lischer 2010) and DNASP 5.10.1
184 (Librado and Rozas 2009). We also computed Fu and Li's (1993) D* and F* tests,
185 as they are particularly powerful for detecting background selection (Fu 1997). If
186 only these tests are significant, this suggests the action of background selection;
187 if, on the contrary, they are not significant, but, for example, the Fu's F_S test is,
188 then this is more compatible with a history of population growth (Fu 1997).

189 To visualize and examine the genealogical and geographical relationships among
190 haplotypes, we estimated haplotype networks using the integer neighbour-
191 joining method (IntNJ) in POPART 1.7 (Leigh and Bryant 2015). The IntNJ
192 method was run with the reticulation tolerance parameter α set to zero. We used
193 outgroup rooting (Jansen et al. 2002; Cassens et al. 2003; Dubach et al. 2013) to
194 estimate the ancestral node of the ingroup. The outgroups used here, and in the
195 phylogenetic tree reconstructions described next, were the brown hyena (*P.*
196 *brunnea*) and the spotted hyena (*C. crocuta*).

197 To infer evolutionary relationships among haplotypes and assess the presence of
198 clades within *H. hyaena*, we constructed phylogenetic trees. The tree
199 reconstructions were conducted with the data partitioned by codon position, as
200 this was the best-fit partitioning scheme according to the corrected Akaike
201 information criterion (AICc) (Akaike 1974; Sugiura 1978; Hurvich and Tsai
202 1989) in PARTITIONFINDER 2.1.1 (Lanfear et al. 2017) using PHYML (Guindon
203 et al. 2010). Phylogenetic analyses were performed using Bayesian inference (BI)
204 and maximum likelihood (ML) as implemented in MRBAYES 3.2.6 (Ronquist et al.
205 2012) and IQ-TREE 1.6.12 (Nguyen et al. 2015), respectively. Analyses in
206 MRBAYES were conducted with two parallel Markov Chain Monte Carlo (MCMC)
207 runs, each with four Markov chains (one cold and three heated), default heating
208 parameter (t = 0.1), and 20 million generations. The first five million generations
209 were discarded as burn-in and, thereafter, chains were sampled every 500
210 generations. The entire general time-reversible (GTR; Lanave et al. 1984)
211 substitution model space was sampled within the analyses (Huelsenbeck et al.
212 2004). Convergence was indicated by an average standard deviation of split
213 frequencies between parallel runs of less than 0.01. For all model parameters,

214 the effective sample size (ESS) was greater than 700 and the potential scale
215 reduction factor (PSRF) was 1.0. Support for tree nodes was determined
216 according to the values of Bayesian posterior probability (BPP) obtained in a
217 majority-rule consensus tree (Holder et al. 2008; Huggins et al. 2011). In IQ-
218 TREE we used the nucleotide substitution models estimated as best fit for each
219 codon position, according to the Bayesian information criterion (BIC; Schwarz
220 1978), by ModelFinder (Kalyaanamoorthy et al. 2017), a model-selection method
221 implemented in IQ-TREE (648 bp dataset, according to codon position: Kimura
222 two-parameter model (K2P; Kimura 1980); Hasegawa-Kishino-Yano model
223 (Hasegawa et al. 1985) with a proportion of invariable sites (HKY+I); HKY) (340
224 bp dataset, according to codon position: K2P+I; HKY; Tamura-Nei model (TN93;
225 Tamura and Nei 1993)). Tree search runs were performed using IQ-TREE default
226 settings, with the exception of the '-allnni' option which turns on a more
227 thorough Nearest Neighbour Interchange (NNI) tree search, and support for each
228 node was evaluated by 1000 non-parametric bootstrap replicates (Felsenstein
229 1985). Majority-rule consensus trees (Berry and Gascuel 1996; Holder et al.
230 2008) were computed with SUMTREES 4.5.2 of the DendroPy library version
231 4.5.2 (Sukumaran and Holder 2010) and visualized and edited with FIGTREE
232 1.4.4 (available at https://github.com/rambaut/figtree/releases).

233 To try to learn more about the recent evolutionary history of the species, we
234 performed several analyses specifically on demographic history. In these
235 analyses we used only the 648 bp dataset (76 striped hyenas; 13 segregating
236 sites), as the 340 bp dataset (50 individuals) only contains five segregating sites.
237 We constructed skyline plots (Ho and Shapiro 2011), which do not depend on a
238 prespecified parametric model of demographic history, using the following four

239 MCMC sampling-based methods: Bayesian Skyline Plot (BSP; Drummond et al.
240 2005), Extended Bayesian Skyline Plot (EBSP; Heled and Drummond 2008),
241 Bayesian Skyride (Minin et al. 2008), and Bayesian Skygrid (Gill et al. 2013).
242 These skyline plot methods were performed in BEAST 1.10.4 ('Bayesian
243 Evolutionary Analysis Sampling Trees'; Suchard et al. 2018), using the BEAGLE
244 3.1.0 library (Ayres et al. 2019), with the data partitioned by codon position. All
245 BEAST input files were created in BEAUti, available in the BEAST package.
246 Methodological details of the MCMC sampling-based skyline analyses are given
247 in Supplementary Information. We compared different skyline models, and also
248 simple parametric models (such as constant population size or exponential
249 growth) (see Villanea et al. 2020), in terms of their relative fit to the data, using
250 the Bayes factor (Jeffreys 1935; Kass and Raftery 1995), the ratio of the marginal
251 likelihoods of the two models under comparison. The marginal likelihoods were
252 estimated in BEAST with path sampling (Lartillot and Philippe 2006) and
253 stepping-stone sampling (Xie et al. 2011), based on 100 path steps and a chain
254 length of two million iterations. We also tested skyline plot methods based on a
255 single tree (Ho and Shapiro 2011): classical (Pybus et al. 2000), generalized
256 (Strimmer and Pybus 2001), and Bayesian multiple-change-point (MCP; Opgen-
257 Rhein et al. 2005). The required ultrametric binary tree was a maximum clade
258 credibility tree (MCCT) from an analysis in MRBAYES. MCMC in MRBAYES was
259 performed as described above. The analysis used a strict clock model with a
260 uniform prior probability distribution on branch lengths, a normally distributed
261 clock rate prior with mean 0.01 and standard deviation 0.0025 (in substitutions
262 per site per Ma) (how we obtained this substitution rate estimate is explained
263 below), and a truncated normal prior probability distribution on the tree age

264 with a mean of 0.4 Ma and a standard deviation of 0.15 (the choice of these
265 values reflects the estimates obtained for the age of the mtDNA diversity,
266 presented below). The single-tree skyline plot analyses were performed in the R
267 package APE version 5.7.1 (Paradis and Schliep 2019). The analyses performed
268 in the R environment used R version 4.2.3 (R Core Team 2023). Bayesian MCP
269 analyses were performed assuming both a constant population size prior
270 demographic function (Felsenstein 1992) and a skyline plot prior demographic
271 function (Opgen-Rhein et al. 2005). We also estimated the fit of different
272 demographic models (constant population size, exponential growth, expansion
273 growth, logistic growth, with both continuous and piecewise variants) (Pybus
274 and Rambaut 2002) to the MCCT in the R package genieR version 0.1.0 (Xiang et
275 al. 2019).

276 We also used the Monte Carlo likelihood approach of Weiss and von Haeseler
277 (1998), as implemented in IPHULA 1.16 (Schmidt et al. 2007), for inference on
278 demographic history. In this analysis we used the HKY model, which was
279 selected as best-fitting the 648 bp alignment by BIC in ModelFinder and also in
280 JMODELTEST 2.1.10 (Darriba et al. 2012), and a transition/transversion rate
281 ratio (kappa, κ) of 28 estimated in TREE-PUZZLE 5.3.rc16 (Schmidt et al. 2002).
282 The method of Weiss and von Haeseler (1998) involves a model with three
283 parameters (theta, θ: mutation-scaled effective population size (N_e) in the past
284 ('initial'), i.e. before an eventual N_e change; tau, τ: mutation-scaled time since N_e
285 started to change exponentially; rho, ρ: ratio of current to initial N_e), which are
286 estimated based on the number of segregating sites and the mean number of
287 pairwise nucleotide differences (k). The method approximates the likelihood
288 surface for the three parameters by coalescent simulations on a three-

289 dimensional grid. After preliminary simulations to explore the likelihood surface
290 and find the parameter ranges that enclose the global maximum likelihood peak,
291 we ran 10 replicate runs, zoomed into grid areas of high likelihoods, for
292 parameter estimation. In these final runs we defined five grid points, equally
293 spaced between 1 and 20, for ρ, 10 grid points, equally spaced between 0.1 and 5,
294 for θ, and 10 grid points, equally spaced between 0.1 and 2, for τ. We ran 50,000
295 simulations for each grid point (i.e. parameter combination), and the maximum
296 allowed difference between simulated and observed k was set to less than 0.03
297 (i.e. ± 1% around the observed mean k of 1.72).

298 Inference on demographic history parameters was also performed with the
299 MCMC coalescent genealogy sampler LAMARC 2.1.10 (Kuhner 2006; Kuhner and
300 Smith 2007), assuming either an exponential growth or shrinkage model or a
301 simpler model of constant population size. We used the Bayesian version of
302 LAMARC, with a default starting value of θ per site of 0.01 and a logarithmic
303 prior between 10^{-5} and 10; in the growth model analysis, we specified for the
304 growth rate parameter (*g*) a starting value of 1 and a linear prior bounded
305 between -500 and 5000. Percentile profile likelihoods were computed for both θ
306 per site and *g*. The HKY model is not implemented in LAMARC, so we used the
307 closest model available, Felsenstein84 (F84; Felsenstein and Churchill 1996),
308 which in any case is very closely related to the HKY model. We ran four
309 replicates of two final chains, in which for each chain we used an initial burn-in
310 period of 200,000 discarded genealogies, followed by two million genealogies
311 sampled every 50 for parameter estimation. We used adaptive heating, with four
312 tree search temperatures (initial values: 1, 1.2, 1.5, 2) in each chain, and tree
313 swapping between temperatures attempted at every step of the chain. TRACER

314 1.7.2 (Rambaut et al. 2018) was used to analyse the resulting trace files, assess
315 run convergence, and calculate summary statistics and ESS values of parameters.

316 We also tested constant population size and exponential growth models in
317 GENETREE 9.01, which implements a Markov chain technique for ancestral
318 inference that simulates gene trees conditional on the mutation pattern in a
319 sample of DNA sequences (Griffiths and Tavaré 1994a,b,c, 1995). GENETREE
320 requires data compatible with the infinite sites model (Kimura 1969; Griffiths
321 and Tavaré 1995). To meet this requirement, in the 648 bp dataset we needed to
322 remove one polymorphic site (site 546) inconsistent with the model, which
323 caused the haplotypes H1 and H3, as well as H7 and H10, to no longer differ from
324 each other. In the constant population model, after preliminary runs with
325 different values of the generating parameter for θ to optimize settings and
326 identify the region of maximum likelihood of θ, we generated likelihood surfaces
327 for θ based on 10 million simulations and 1000 points for the empirical
328 distribution. In the exponential growth model, after preliminary runs to identify
329 the likelihood surface regions of θ and growth rate yielding the highest joint
330 likelihood, as the two parameters are not independent of each other, we
331 generated likelihood surfaces separately for θ and growth rate, in both cases
332 based on 100,000 simulations and 1000 points for the empirical distribution. A
333 likelihood ratio test (LRT) was used to determine whether there was a
334 significant difference in the fit to the data between the simpler constant
335 population size model and the exponential growth model.

336 To estimate the time to the most recent common ancestor (TMRCA) of the
337 observed mitochondrial genetic variation, we obtained an estimate of the

338 substitution rate for the analysed *Cyt b* fragment as follows. Given our observed
339 *Cyt b* divergence between striped hyena and its closest relative, the brown hyena,
340 and the divergence between the two lineages that the fossil record suggests may
341 have occurred 4-5.2 Ma ago (Werdelin and Solounias 1991; Werdelin et al. 1994;
342 Werdelin 2003; Werdelin and Manthi 2012; Westbury et al. 2020, 2021), an
343 estimate of the substitution rate can be obtained using equation $d_{xy} = 2\mu T$ (Nei
344 1985, 1987), where d_{xy} is the mean sequence divergence per site, μ is the mean
345 substitution rate per site, and T is the time since divergence. Genetic estimates of
346 divergence time between *Hyaena* and *Parahyaena* also fall at about 4-5 Ma
347 (Rohland et al. 2005; Koepfli et al. 2006; Hu et al. 2021; Westbury et al. 2021).
348 We confirmed the assumption of nucleotide substitution rate homogeneity
349 among the striped hyena haplotypes and between these and *Parahyaena*, using
350 *Crocuta* as outgroup, via generalized relative rate tests with topological
351 weighting (Robinson et al. 1998) in RRTREE 1.1.11 (Robinson-Rechavi and
352 Huchon 2000), and with Tajima's relative rate test (1993) as implemented in
353 MEGA 11.0.9 (Tamura et al. 2021). In RRTREE, statistical significance ($P < 0.05$)
354 of rate heterogeneity among lineages was assessed using both synonymous
355 substitution rates (*Ks*) and nonsynonymous substitution rates (*Ka*). The d_{xy}
356 between striped hyenas and *Parahyaena* was computed in MEGA using the
357 substitution model that best fits the data, according to the BIC in both IQ-TREE
358 and MEGA, among the models available in MEGA for distance estimation (TN93
359 for the 648 bp alignment; K2P for the 340 bp alignment), and the standard error
360 was estimated using 1000 bootstrap replicates. As a point value for the
361 divergence time between *Hyaena* and *Parahyaena* we used 4.6 Ma, the midpoint
362 in the 4-5.2 Ma interval suggested by the fossil record. Similar point values have

363 been used or estimated for the divergence of the two lineages in recent genetic
364 studies of hyaenids (Westbury et al. 2020; Hu et al. 2021; Westbury et al. 2021).
365 The d_{xy} estimate for the 648 bp alignment was 0.091 ± 0.013 (mean ± standard
366 error), and for the 340 bp alignment was 0.086 ± 0.017 (mean ± standard error).
367 Approximating 95% confidence intervals (CIs) based on 1.96 standard errors
368 yields 95% CIs of [0.066, 0.116] and [0.053, 0.119], respectively. Using the above
369 equation $d_{xy} = 2\mu T$ with T = 4.6 Ma, the d_{xy} estimates yield substitution rate
370 estimates of 0.99% per Ma (95% CI: 0.72-1.26%) and 0.93% per Ma (95% CI:
371 0.58-1.29%), respectively; i.e. ≈ 1% per Ma, the traditional average rate for
372 mtDNA (Brown et al. 1979; Wilson et al. 1985). Also integrating the uncertainty
373 in the fossil record about the divergence time between *Hyaena* and *Parahyaena*
374 results, respectively, in the following CIs: 0.63-1.45% and 0.51-1.49%.

375 We estimated the TMRCA of the extant striped hyena mitochondrial genetic
376 variation, based on both 648-bp and 340-bp datasets and using different
377 methods implemented in Thomson's estimator (Thomson et al. 2000),
378 GENETREE and BEAST. An attractive feature of Thomson's estimator is that it
379 does not rely on a specific population genetic model (Thomson et al. 2000). On
380 the other hand, the estimator requires data consistent with the infinite sites
381 model. Accordingly, as mentioned above for the GENETREE analysis, in the 648
382 bp dataset we removed a polymorphic site (site 546). In the 340 bp dataset it
383 was also necessary to remove one polymorphic site incompatible with the said
384 model (site 317), which caused the haplotypes H1 and H7, as well as H4 and H5,
385 to no longer differ from each other. We computed Thomson's estimator and its
386 variance (the latter according to equation 4 in Hudson 2007) using NumPy code
387 kindly provided by Dr. Helmut Simon (https://github.com/helmutsimon), and

388 95% CIs (as in equation 5 in Hudson 2007). The required estimate of the
389 unfolded site frequency spectrum was obtained using the site.spectrum function
390 of the R package PEGAS (Paradis 2010) and an estimate of the ancestral
391 sequence of the ingroup. We estimated this ancestral sequence using the two
392 outgroups (*P. brunnea* and *C. crocuta*) in two different computer programs,
393 FASTML (http://fastml.tau.ac.il/; Ashkenazy et al. 2012) and IQ-TREE, for
394 comparison purposes, and both provided the same estimate of the ingroup's
395 ancestral sequence for each of the two datasets. In GENETREE, we used both the
396 constant population model and the exponential growth model, with the ML
397 estimates of θ and growth rate (obtained as described above) as generating
398 parameters. Both the 648-bp and 340-bp datasets were made compatible with
399 the infinite sites model by removing one inconsistent site in each dataset, as
400 described above. The ingroup's ancestral sequences used, one for each of the two
401 datasets, were those mentioned above, estimated using FASTML and IQ-TREE.
402 We ran 100 million simulations to obtain TMRCA estimates and likelihood
403 surfaces with 10,000 points. In BEAST we ran analyses with both outgroups and
404 only with *P. brunnea*. The outgroups correspond to the molecular clock
405 calibrations used: i) the divergence between the *Crocuta* lineage and the lineage
406 ancestral to *Hyaena* and *Parahyaena*, for which we used a normally distributed
407 calibration prior with mean 9.5 Ma and standard deviation 2 Ma (Jenks and
408 Werdelin 1998; Rohland et al. 2005; Westbury et al. 2019, 2020; Hu et al. 2021);
409 and ii) the split between *Hyaena* and *Parahyaena*, with a normal calibration prior
410 with mean 4.6 Ma and standard deviation 0.3 Ma (Rohland et al. 2005; Koepfli et
411 al. 2006; Hu et al. 2021; Westbury et al. 2020, 2021). The datasets were
412 partitioned by codon position. We used the piecewise-constant skyline tree prior

413 (Drummond et al. 2005), which has been shown to perform well in molecular
414 dating analyses involving a mixture of inter- and intraspecific data (Ritchie et al.
415 2017; Mello et al. 2021), and the approximate continuous time Markov chain
416 (CTMC) rate reference prior (Ferreira and Suchard 2008). In the analyses of the
417 648 bp dataset with two outgroups, we used relaxed clock models (uncorrelated
418 lognormal model, UCLN, Drummond et al. 2006; random local clocks model, RLC,
419 Drummond and Suchard 2010) because a LRT (performed in MEGA) of the strict
420 molecular clock hypothesis rejected it ($P < 0.001$). In contrast, this hypothesis
421 was not rejected ($P = 0.060$) for the dataset with only *P. brunnea* as outgroup, in
422 agreement with the results from RRTREE and Tajima's relative rate test, and
423 therefore we assumed a strict molecular clock in the analysis of this dataset.
424 Similarly, the strict clock hypothesis was rejected for the 340 bp dataset with
425 two outgroups ($P = 0.007$), but not for the one with only *P. brunnea* as outgroup
426 ($P = 0.226$). For each BEAST analysis, we ran four independent replicate MCMC
427 simulations, each with 50 million generations and sampled every 5000
428 generations following a pre-burnin of 10%. We also ran two replicate
429 simulations without data to obtain estimates from the prior distribution and test
430 the influence of priors on posterior distributions (Drummond et al. 2006; Brown
431 and Smith 2018). For each analysis, we used TRACER with the default burn-in to
432 assess convergence of the chain to the stationary distribution, obtain estimates
433 and ESS of parameters, and plot marginal posterior densities. After verifying
434 convergence, and confirming that posterior distributions of estimates were
435 markedly different from the prior distributions, the tree files from the four runs
436 with data were combined using SUMTREES, with the first 25% (2500) trees from
437 each run discarded as burn-in, into a MCCT with median node ages. This method

438 has a good overall performance, in terms of accuracy in estimating ages and
439 model fit, compared to other tree summary approaches (Morrison 2008;
440 Battistuzzi et al. 2010; Heled and Bouckaert 2013; Bromham et al. 2018). After
441 analysing the results of the runs with the relaxed clock models, we noticed that
442 in general the TMRCA estimates for the striped hyena mitochondrial variation
443 seemed too high (> 800 ka) when compared with those obtained with Thomson's
444 estimator (range of mean estimates: ~ 410-420 ka) and with GENETREE (range
445 of mean estimates: ~ 240-330 ka) and the only previously published estimate
446 that we are aware of (340 ka; Rohland et al. 2005). Thus, we considered these
447 high estimates likely to be erroneous. The apparent poor performance of the
448 UCLN clock model does not seem to be due to the presence of significant rate
449 autocorrelation among lineages, since, for both datasets with two outgroups, the
450 95% highest posterior density (HPD) interval of the covariance between parent
451 and child branch rates contained zero (Drummond et al. 2006). The CorrTest
452 method (Tao et al. 2019), implemented in MEGA, also did not reject the
453 hypothesis of independent evolutionary rates among lineages (P > 0.05).
454 However, we also noted that the standard deviation of the UCLN clock
455 (ucld.stdev parameter) was low in the analyses of both datasets (648 bp dataset:
456 mean = 0.049, median = 0.010, 95% HPD interval = [3.7 x 10^{-7}, 0.243]; 340 bp
457 dataset: mean = 0.138, median = 0.055, 95% HPD interval = [9.9 x 10^{-6}, 0.557]),
458 which suggests that the data is quite clock-like. The same indication was given by
459 the results for the posterior distribution of the number of rate changes (K)
460 (rateChangeCount parameter) in the RLC analyses, as a greater posterior than
461 prior probability for K = 0 supports the global clock hypothesis, while small or
462 negligible posterior probability for K = 0 strongly rejects the hypothesis

463 (Drummond and Suchard 2010). In the analysis with the 648 bp dataset, the
464 posterior distribution for K (mean: 0.619; median and mode: 0; 95 % HPD
465 interval: [0,2]) showed a higher probability for $K = 0$ than the prior distribution.
466 This was not so with the 340 bp dataset (mean: 0.769; median and mode: 1;
467 95 % HPD interval: [0,2]), for which 43% of the probability mass of the posterior
468 distribution was for $K = 1$, but yet 42% of the probability was for $K = 0$. In such
469 cases, a strict clock (SC) may perform better in dating estimation (Ho et al. 2005;
470 Brown and Yang 2011). Marginal likelihood estimates, obtained using stepping-
471 stone sampling in BEAST with 100 path steps and a chain length of 500,000, also
472 supported the use of a SC (648 bp dataset: log marginal likelihood = -1344.5; 340
473 bp dataset: log marginal likelihood = -711.2), compared with the RLC (648 bp
474 dataset: log marginal likelihood = -1344.5; 340 bp dataset: log marginal
475 likelihood = -711.1) and the UCLN (648 bp dataset: log marginal likelihood = -
476 1347.4; 340 bp dataset: log marginal likelihood = -713.2). Therefore, we decided
477 to also use the SC in the analyses with two outgroups.

478

479 **Results**

480 *Genetic variation, haplotype relationships and phylogenetic analyses*

481 The new striped hyena DNA sequences discovered in this study were deposited
482 in GenBank under accession numbers PP328974-PP328975 (Table 1). Several
483 observations indicate that the generated sequences are mitochondrial and not
484 nuclear-integrated copies of mtDNA (Arctander 1995; Zhang and Hewitt 1996;
485 Bensasson et al. 2001). First, PCRs consistently yielded single products of the
486 expected size. Second, independent replicate PCRs produced identical sequences.

487 Third, the sequences were unambiguous and did not contain indels or stop
488 codons (Smith et al. 1992; Mirol et al. 2000; Triant and DeWoody 2007).

489 Genetic diversity measures and results of mutation-drift equilibrium tests, for
490 both the 648 bp and 340 bp alignments of striped hyenas, are presented in Table
491 2. In the longer alignment the number of polymorphic sites was 13, with nine
492 being parsimony-informative and four singletons, and in the shorter alignment
493 the number of polymorphic sites was five, four of them parsimony-informative
494 and one a singleton. In both datasets, none of the mutation-drift equilibrium
495 tests used rejected at the 95% confidence level the null hypothesis of a constant
496 size population under the neutral model; note that in the F_S statistic test, an α =
497 0.02 corresponds to a 5% significance level (Fu 1997). The D* and F* statistics
498 were respectively -0.739 (P = 0.233) and -0.983 (P = 0.180) for the 648 bp
499 alignment, and 0.134 (P = 0.529) and -0.228 (P = 0.396) for the 340 bp alignment.
500 Analysis of the mismatch distribution for the 648 bp dataset under a model of
501 sudden demographic expansion (Schneider and Excoffier 1999), implemented in
502 ARLEQUIN, did not reject the expansion hypothesis (P = 0.678; 20,000 bootstrap
503 replicates). In agreement, the Harpending's (1994) raggedness index was low
504 (0.032; P = 0.552). The model parameter for the expansion time was estimated at
505 1.877 (in mutational time units) (95% CI: 0.715-3.736), which, assuming a
506 nucleotide substitution rate of 1% per Ma, translates into 144,830 years (95%
507 CI: 55,170-288,272). For the 340 bp dataset, the analysis of the mismatch
508 distribution under the sudden expansion model failed to converge, possibly due
509 to the low number of polymorphic sites in this dataset. However, analyses in
510 DNASP showed that the observed mismatch distribution fits better with the
511 expected distribution under a sudden expansion model (Rogers and Harpending

512 1992) than under a model of constant population size (Fig. S1). We also
513 attempted in ARLEQUIN to test the mismatch distribution under a model of
514 spatial expansion (Ray et al. 2003; Excoffier 2004), but the least-square
515 procedure to fit the expected and observed mismatch distributions did not
516 converge.

517 Both the haplotype networks (Fig. 2) and the BI and ML phylogenetic trees (Figs.
518 S2-S4), for both the 648 bp and 340 bp datasets, indicate a shallow genetic
519 divergence among the striped hyena haplotypes, and essentially an overall lack
520 of phylogeographic structure.

521

522 *Demographic simulations*

523 The different MCMC sampling-based skyline plot analyses conducted in BEAST
524 did not identify signs of significant population size change in the Late Quaternary
525 demographic history of the striped hyena. In the BSP analyses using a normally
526 distributed prior for the substitution rate with mean 0.01 and standard deviation
527 0.0025 (in substitutions per site per Ma), truncated to be positive
528 (corresponding to a 95% interquantile range of 0.005-0.015), the estimated
529 mean tree root height was 281 ka (median: 248 ka; 95% HPD interval: 61-571
530 ka) for the piecewise-constant model (ESS > 1500 for all parameters), and 264
531 ka (median: 232 ka; 95% HPD interval: 60-536 ka) for the piecewise-linear
532 model (ESS > 2200 for all parameters). The respective BSPs for the two models
533 are shown in Fig. S5, in which it can be observed that the rise of the population
534 size curve at the left end of the plots never exceeds the 95% HPD limits along the
535 rest of the graph (Ho and Shapiro 2011; Grant 2015), where the curve is

536 essentially flat. The results of analyses using either a uniformly or exponentially
537 distributed prior for the substitution rate exhibited a similar pattern of no
538 significant population size changes (not shown). In the EBSP analyses using the
539 same normally distributed prior for the substitution rate as above, the estimated
540 mean tree root height was 363 ka (median: 331 ka; 95% HPD interval: 85-717
541 ka) for the stepwise model (ESS > 2200 for all parameters), and 341 ka (median:
542 304 ka; 95% HPD interval: 58-688 ka) for the linear model (ESS > 2800 for all
543 parameters). For the 'population size changes' parameter in each model, the
544 estimated mean was 0.867 (median and mode were 1, and the 95% HPD interval
545 was 0-2) and 1.024 (median and mode were 1, and the 95% HPD interval was 0-
546 3), respectively. The respective EBSPs for the two models are presented in Fig.
547 S6. Results of analyses using a uniformly distributed prior for the substitution
548 rate also did not indicate significant population size changes in recent
549 demographic history (not shown). The skyride (ESS > 6500 for all parameters)
550 and skygrid (ESS > 650 for all parameters) analyses also yielded the same result
551 (Fig. S7); their respective mean tree root height estimates were 72 ka (median:
552 66 ka; 95% HPD interval: 24-133 ka) and 277 ka (median: 245 ka; 95% HPD
553 interval: 67-564 ka). In the analyses with the constant population size (uniform
554 population size prior; ESS > 4700 for all parameters) and exponential growth
555 (uniform population size and growth rate priors; ESS > 1500 for all parameters)
556 models, the estimated mean tree root heights were 419 ka (median: 349 ka; 95%
557 HPD interval: 111-809 ka) and 508 ka (median: 349 ka; 95% HPD interval: 103-
558 814 ka). The 95% HPD interval of the posterior distribution of the exponential
559 growth rate parameter ([-1.803, 4.018]) included zero, i.e. did not reject a
560 constant population size. The skygrid model was the one with the highest log

561 marginal likelihood estimate (stepping-stone sampling MLE = -1096; path
562 sampling MLE = -1095), but not significantly different from that of the piecewise-
563 linear version of the BSP (MLE = -1097 by both stepping-stone sampling and
564 path sampling). Still using Bayes factors, the skygrid model had positive support
565 against the piecewise-constant version of the BSP (stepping-stone sampling MLE
566 = -1099; path sampling MLE = -1098) and very strong support (2 ln Bayes factor
567 > 10) against other models. The results of the single-tree skyline plot analyses
568 done in APE suggested population growth during the Late Pleistocene (Figs. S8-
569 S9). The fit of the different demographic models tested in genieR was compared
570 using the Akaike information criterion (AIC). The constant population size model
571 was rejected (ΔAIC = 90), while the exponential growth (AIC = -1575.919) and
572 expansion growth (AIC = -1573.919) models were equally supported (ΔAIC ≤ 2).

573 The implementation of Weiss and Von Haeseler's (1998) method in IPHULA
574 estimated a maximum likelihood model parameter combination (ρ = 10, θ = 2.8,
575 τ = 0.3) that suggests slow exponential growth since the Late Pleistocene.
576 Specifically, assuming a nucleotide substitution rate of 1% per Ma and a
577 generation time for the striped hyena of 9.2 years (Pacifici et al. 2013), a τ of 0.3
578 translates to about 23 ka. In turn, a θ of 2.8 corresponds to a female N_e of 23484
579 at that time. However, the 95% confidence set of models, based on the likelihood
580 ratio of each combination of demographic parameters tested and the maximum
581 likelihood model, included scenarios of constant population size.

582 In LAMARC, analyses using the exponential growth or shrinkage model
583 suggested a slow exponential growth scenario. The highest likelihood estimate of
584 *g* was 910 (95% HPD interval: [-83, 4704]; ESS > 700 in all individual replicate

585 runs). Therefore, the 95% HPD interval crossed zero, not ruling out scenarios of
586 constant population size or even contraction. Moreover, LAMARC estimates of *g*
587 tend to be biased upwards when only a few loci are analysed (Kuhner et al.
588 1998). The highest likelihood estimate of current θ per site was 0.008 (95% HPD
589 interval: [0.002, 0.020]; ESS > 350 in all individual replicate runs), which
590 corresponds to a θ of 5.2 for the 648 bp fragment and, again assuming a
591 nucleotide substitution rate of 1% per Ma and a generation time of 9.2 years, can
592 be translated into a female N_e (N_ef) of 43478. However, LAMARC estimates of θ
593 in the varying population size model may also be slightly biased upward, due to
594 correlation with *g*, when only a few loci are analysed (Kuhner et al. 1998). In the
595 analyses with the constant population size model, the highest likelihood estimate
596 of θ per site was 0.004 (95% HPD interval: [0.002, 0.006]; ESS > 850 in all
597 individual replicate runs), which corresponds to a θ of 2.6 for the 648 bp
598 fragment and, assuming a nucleotide substitution rate of 1% per Ma and a
599 generation time of 9.2 years, can be translated into a N_ef of 21739. This θ
600 estimate was very similar to that obtained using Watterson's (1975) estimator
601 based on the number of segregating sites: 2.7.

602 In GENETREE, the analyses using the constant population size model also yielded a
603 ML θ estimate of 2.7 (thus, a N_ef of 22645), while in the exponential growth
604 model the ML θ estimate was 3.1 (i.e. N_ef = 26000); the latter model suggested a
605 slow exponential growth rate. However, a LRT did not detect a significant
606 difference in the fit to the data between the simpler constant population size
607 model and the exponential growth model (P = 0.460).

608

609 *Time to most recent common ancestor (TMRCA)*

610 Thomson's method estimates for the TMRCA of the observed genetic variation in
611 *H. hyaena* were 419 ± 112 ka (mean ± standard error; 95% CI: 225-672 ka)
612 based on the 648 bp dataset and a nucleotide substitution rate of 1% per Ma, and
613 410 ± 160 ka (mean ± standard error; 95% CI: 153-793 ka) based on the 340 bp
614 dataset and a nucleotide substitution rate of 0.93% per Ma (Table 3). Adding the
615 uncertainty in the fossil record and in the d_{xy} estimate for the divergence
616 between *Hyaena* and *Parahyaena* into the nucleotide substitution rate estimate
617 (i.e. 95% CIs of 0.63-1.45% and 0.51-1.49% for the 648 bp and 340 bp datasets,
618 respectively) yields the following respective 95% CIs for the TMRCA: [155, 1066]
619 and [95, 1447]. In GENETREE, analyses of the 648 bp dataset with the constant
620 population size model and the ML estimate of θ of 2.7 gave a TMRCA estimate, in
621 coalescent time units, of 1.6 ± 0.5 (mean ± standard deviation), which, using a
622 nucleotide substitution rate of 1% per Ma and a generation time of 9.2 years,
623 corresponds to 333 ± 104 ka; approximating the 95% CI by ± 2 standard
624 deviations (Harding et al. 1997) gives 125-542 ka (86-860 ka including
625 uncertainty in the fossil record and in the d_{xy} estimate for the divergence
626 between *Hyaena* and *Parahyaena* in the nucleotide substitution rate estimation).
627 With the exponential growth model and the respective ML estimate of θ of 3.1,
628 the TMRCA estimate in coalescent time units was 1.27 ± 0.35 (mean ± standard
629 deviation), i.e. 304 ± 84 ka (95% CI: 136-471 ka) (94-748 ka incorporating the
630 aforementioned uncertainty in the nucleotide substitution rate estimate).
631 Analyses of the 340 bp dataset with the constant population size model and a ML
632 estimate of θ of 0.9 yielded a TMRCA estimate in coalescent time units of 1.9 ±
633 0.8 (mean ± standard deviation), which, using a nucleotide substitution rate of

634 0.93% per Ma and a generation time of 9.2 years, translates into 279 ± 118 ka
635 (95% CI: 44-515 ka) (28-939 ka considering the uncertainty in the nucleotide
636 substitution rate estimate). Analyses of the same dataset with the exponential
637 growth model and a ML estimate of θ of 1.1 resulted in a TMRCA estimate in
638 coalescent time units of 1.38 ± 0.45 (mean ± standard deviation), i.e. 240 ± 78 ka
639 (95% CI: 83-397 ka) (52-723 ka considering the uncertainty in the nucleotide
640 substitution rate estimate) (Table 3). The fact that GENETREE TMRCA estimates
641 using two very different demographic models were quite similar suggests a
642 minor effect of the underlying demographic model on the estimates. TMRCAs
643 estimates from the BEAST analyses for different combinations of dataset,
644 calibrations, and clock model, are given in Table 3; ESS values were > 3000 for all
645 parameters in all individual runs under the strict clock, and > 200 under a
646 random local clock. Across the different analyses, the TMRCA point estimate for
647 the observed mtDNA variation in striped hyenas ranged from 330-480 ka, with
648 the clock rate point estimate between 1.1% and 1.6% per Ma, hovering around
649 1.3% per Ma (Table 3), which is quite close to the rate estimates of ≈ 1% that we
650 inferred based on the d_{xy} distance estimate between *Hyaena* and *Parahyaena*.

651

652 **Discussion**

653 Topics such as genetic diversity and structure, phylogeography, and evolutionary
654 history of the striped hyena remain poorly investigated (Wilkinson et al. 2024).
655 To date, our knowledge of the mitochondrial genetic diversity and
656 phylogeographic structure of the striped hyena comes essentially from the
657 analysis of a 340-bp *Cyt b* fragment in 13 individuals from across the historical

658 and present distribution of the species, which was part of a ground-breaking
659 study on the evolutionary history of extant hyenas (Rohland et al. 2005). That
660 analysis found low genetic diversity and no evident geographic structure of this
661 diversity. In our study we wanted to investigate the mitochondrial genetic
662 variation of the striped hyena in Algeria, using for this purpose a *Cyt b* fragment
663 encompassing the one analysed by Rohland et al. (2005). It was also our
664 objective, with the addition of samples from the Arabian Peninsula and Iran
665 (regions not covered in Rohland et al. (2005) and where classic morphological
666 subspecies have been described) and using a diverse set of demographic history
667 and coalescence time inference analyses, to conduct a more exhaustive and
668 better-sampled global analysis of the species than in Rohland et al. (2005). The
669 finding that a sample of 24 Algerian striped hyenas was monomorphic for a 753-
670 bp *Cyt b* fragment was striking, and agrees with the notion of low mitochondrial
671 diversity in the species. The apparent absence of phylogeographic structure in
672 the species seemingly extends also to southern Arabia and Iran (Figs. 2 and S2-
673 S4), with for example haplotype sharing between Oman and Iran for the 648-bp
674 fragment (Fig. 2a), which therefore also calls into question the validity of the
675 *sultana* and *hyaena* subspecies.

676 *Demographic history*

677 The observed patterns of moderate haplotype diversity and low nucleotide
678 diversity (Table 2) are compatible with a historical population bottleneck in the
679 species, followed by demographic growth and the appearance/accumulation of
680 new haplotypes (Grant and Bowen 1998). Moreover, the mismatch distributions
681 for both 648 bp and 340 bp datasets showed a significant fit with expected

682 distributions under a sudden expansion model. Nevertheless, it should be taken
683 into account that tests based on the mismatch distribution have been noted as
684 conservative in rejecting the null hypothesis of population expansion (Ramos-
685 Onsins and Rozas 2002; Fahey et al. 2014). On the other hand, none of the
686 mutation-drift equilibrium tests used rejected the hypothesis of constant
687 population size (Table 2). A reduced number of segregating sites, as is the case
688 here, can limit the power of these tests (Ramos-Onsins and Rozas 2002). In
689 particular, the values of the F_S statistic were clearly negative, a result consistent
690 with expectations in a scenario of population growth, which may suggest a lack
691 of power in the datasets due to a small number of segregating sites. However, a
692 limited number of segregating sites does not necessarily prevent significant
693 results in those tests (Pilkington et al. 2008; Kvist et al. 2011; Cabanne et al.
694 2013; Álvarez-Varas et al. 2015; Hassan-Beigi et al. 2022). These tests also have
695 less power if the demographic expansion is historically very recent (e.g.
696 Holocene), especially if it was not very intense (Fu 1997; Ramos-Onsins and
697 Rozas 2002). Rohland et al. (2005) also did not detect a statistical signal of
698 demographic expansion in their data (the authors did not report which analyses
699 and tests of demographic history were performed), but the observed
700 phylogeographic pattern was inferred to be the result of a recent and rapid range
701 expansion, probably within the last 100 ka and possibly as recently as the
702 Holocene, from a small refuge population, most likely located in Africa as,
703 according to Werdelin and Solounias (1991), no Late Pleistocene fossils of
704 striped hyenas are know outside of Africa. The results of the single-tree skyline
705 plot analyses and demographic model fit comparisons all supported the
706 population growth hypothesis. However, these single-tree methods have the

707 limitation of not taking into account uncertainty in the tree topology (Strimmer
708 and Pybus 2001; Pybus and Rambaut 2002); we attempted to mitigate this
709 problem by obtaining the tree to be used by Bayesian analysis and summarizing
710 the posterior into a MCCT, the tree that maximizes the product of the posterior
711 clade probabilities. All the most sophisticated demographic inference methods
712 used (MCMC sampling-based skyline plots, IPHULA, LAMARC, GENETREE) could
713 not reject with high confidence (> 95%) the hypothesis of constant population
714 size. MCMC sampling-based skyline methods may allow to correctly identify
715 demographic trajectories based on single-locus data (Gill et al. 2013), but such
716 datasets may also lack the power to reject the hypothesis of constant
717 demography (Heled and Drummond 2008). For example, although the sample
718 size we analysed cannot be considered small (n = 76), it has been noted in the
719 literature that the BSP method may lack the power to correctly detect
720 demographic expansions using single locus mitochondrial data when the sample
721 size is < 100 (Grant 2015). Again, also for those more complex methods it is
722 obviously preferable that the number of segregating sites is not small (Ho and
723 Shapiro 2011), but they have been applied to datasets with numbers of
724 segregating sites similar to our data, and produced significant results (e.g.
725 Griffiths and Tavaré 1994b, 1999; Kvist et al. 2011; Cabanne et al. 2013; Horreo
726 et al. 2013; Kuo et al. 2014; Álvarez-Varas et al. 2015; Ely et al. 2017). A small
727 number of segregating sites may not be the only challenge for the methods
728 implemented in IPHULA, LAMARC, and GENETREE. Another factor is the
729 assumed population growth model, which in the case of these three methods is
730 the exponential model. It is conceivable that the hypothesis of constant
731 population size may not be rejected with confidence if an exponential growth

732 model is very far from the true demographic history, which could be, for example,
733 one of slow, more or less linear growth, or a complex history alternating
734 recurrent episodes of expansion, stability, and/or decline (e.g. Excoffier and
735 Schneider 1999; Strimmer and Pybus 2001). For example, in the BEAST analysis
736 with the exponential growth coalescent model as tree prior, the marginal
737 posterior distribution of the exponential growth rate parameter included zero,
738 implying that a contrasting model of constant population size could not be
739 rejected. Overall, the observed patterns of haplotype and nucleotide diversity,
740 the mismatch distributions, and the single-tree skyline plot analyses and
741 demographic model fit comparisons supported the hypothesis of population
742 growth, confirming the inference by Rohland et al. (2005). This hypothesis was
743 not statistically significant in the more sophisticated and exacting methods used,
744 probably largely as a result of the limited power provided by a small number of
745 segregating sites.

746 It was surprising to observe only one haplotype in 24 individuals from Algeria,
747 whereas for example 41 animals from Iran showed eight haplotypes (648 bp
748 fragment; Table 1, Fig. 2a). It is therefore possible that the post-bottleneck
749 demographic growth within Africa was not as rapid and intense as that
750 associated with the colonization of Eurasia. If this was the case, then signatures
751 of demographic expansion will perhaps be easier to detect using large Asian
752 population samples. In the same way that in the human species, signatures of
753 demographic expansion are easier to detect in non-African populations than in
754 African ones (Ingman et al. 2000; Atkinson et al. 2008) or in pooled or global
755 datasets (e.g. Atkinson et al. 2009). Given the typical substitution rates in

756 mitochondrial DNA, it is expected that recent demographic expansions (e.g.
757 postglacial), particularly in organisms whose generation time is not short, may
758 be difficult to detect with data from relatively short mtDNA fragments (e.g. <
759 1000 bp), even using efficient methods in extracting the information present in
760 the data (Felsenstein 1992). Data from mitogenomes can be very useful in such
761 cases (Keis et al. 2013; Jensen et al. 2018). Future mitogenomic studies will likely
762 help to clarify open questions regarding the nature and age of the Late
763 Quaternary striped hyena demographic and geographic expansion(s) suggested
764 by the currently available mitochondrial data.

765 *Age of the extant mitochondrial DNA variation*

766 For the estimation of the TMRCA of the sampled genetic variation we also used
767 very different methods, always based on the same rationale that if different
768 analytical approaches yield congruent results, this increases our confidence in
769 them. The Thomson's estimator is simple and does not require the assumption of
770 a model for the demographic history of the population. GENETREE computes
771 TMRCA likelihood distributions conditional on the gene tree mutation structure,
772 a given value of θ, and the assumed model of population size history. For both
773 approaches the TMRCA estimates were translated into absolute time using the
774 estimated substitution rates for each fragment (340 bp and 648 bp). In contrast,
775 in BEAST the time information to calculate the TMRCA in years was derived from
776 fossil calibrations for the divergence between the striped hyena lineage and the
777 brown hyena and spotted hyena lineages. The TMRCA point estimates of the
778 observed genetic variation ranged between about 300 and 420 ka based on the
779 648 bp dataset, and between about 240 and 480 ka based on the 340 bp dataset

780 (Table 3); hence there was relative agreement between the estimates from the
781 two datasets. The variability of the estimates was lower for the dataset with
782 longer sequence length and more segregating sites, as would perhaps be
783 expected. For both datasets, estimates were higher with the Thomson's method
784 (648 bp: 419 ka; 340 bp: 410 ka) and with BEAST (648 bp: 334-387 ka; 340 bp:
785 441-484 ka) than with GENETREE (648 bp: 304-333 ka; 340 bp: 240-279 ka).
786 This pattern is in line with arguments that GENETREE results should be
787 interpreted as estimated lower bounds for the true values (Hammer et al. 1998;
788 Karafet et al. 1999). In fact, most estimates from Thomson's method and BEAST
789 tended to hover around 400 ka, coinciding with one of the longest and warmest
790 interglacials of the last 800,000 years (within Marine Isotope Stage 11, MIS 11:
791 374-424 Ka, Lisiecki and Raymo 2005), with environmental conditions similar to
792 the Holocene (Howard 1997; Kleinen et al. 2014; Tzedakis et al. 2022). Rohland
793 et al. (2005), using the phylogenetic program r8s (Sanderson 2003) and a point
794 estimate of 10 Ma for the divergence between spotted and striped/brown
795 hyenas as a calibration date, estimated the TMRCA of striped hyenas at about
796 340 ka, which is quite congruent with our estimates. One of our goals in this
797 study was precisely to provide additional estimates, using different methods, for
798 comparison with that one obtained with r8s. For instance, it has been suggested
799 that BEAST may be more accurate than r8s for clade age estimation, particularly
800 with single locus mitochondrial data (Mulcahy et al. 2012).

801 *Nuclear DNA perspective*

802 A recent whole nuclear genome sequencing study on the evolutionary histories
803 of extant hyenas (Westbury et al. 2021), in which the striped hyena was

804 represented by a single captive individual of unknown geographic origin,
805 indicated a very low genetic diversity also in the nuclear genome of the species,
806 and estimated a demographic growth in the lineage starting ~ 500 ka that was
807 eventually followed by a sharp decline since ~ 100 ka. A priority in the near
808 future should be a genome-wide study with a broad geographic sampling
809 throughout Africa and Asia to allow a better characterization of genetic diversity,
810 and the impact on it of the anthropogenic reduction of the species in the last
811 centuries, and a more powerful and informative assessment of population
812 structure and demographic history, as already attempted for its closest extant
813 species, the brown hyena (Westbury et al. 2018).

814

815

816

817

818

819

820

821

822

823

824

825 **Acknowledgments**

826 Carlos Rodríguez Fernandes thanks the support of CE3C and CHANGE through a
827 FCT-Tenure auxiliary researcher contract (FCiência.ID grant 1441) and an
828 auxiliary researcher contract (FCiência.ID grant 0801), and of FPUL for a
829 contract of invited assistant professor. We thank Yassine Righi for a striped
830 hyena sample from Belezma National Park, Algeria. We are extremely grateful to
831 Dr. Helmut Simon for all his advice on TMRCA estimation, and for the NumPy
832 code to compute Thomson's estimator. We thank Professor Robert Griffiths for
833 his advice on running GENETREE and analysing the results. We thank Dr.
834 Emmanuel Paradis for his advice on how to compute the unfolded site frequency
835 spectrum in his R package PEGAS. We thank Dr. Heiko A. Schmidt for his advice
836 in installing the IPHULA software. Lastly, we would like to thank Mammalian
837 Biology's Editor in Chief, Professor Heiko Georg Rödel, and two anonymous
838 reviewers for helpful comments and suggestions to improve the manuscript.

839

840 **Author contributions**

841 LD and MR performed the laboratory work for genetic analysis. LD and CRF
842 analysed the data and led the writing of the manuscript. All other authors read
843 previous versions of the manuscript and approved the final version. LD, HBH,
844 RHA, SMM, ZA, AC, and PV contributed striped hyena biological samples. YHB
845 produced Figure 1.

846

847

Funding

The costs associated with the sampling were supported by the private resources of Louiza Derouiche. This study was financed by Portuguese National Funds, through FCT - Fundação para a Ciência e a Tecnologia, within projects granted to CE3C: UID/BIA/00329/2019, UIDB/00329/2020 (https://doi.org/10.54499/UIDB/00329/2020), UIDP/00329/2020 (https://doi.org/10.54499/UIDP/00329/2020), and UIDP/00329/2025.

Data availability

Data from this study are available upon reasonable request to the corresponding author.

Conflict of interest

The authors have no conflicts of interest to declare.

868 **References**

869 Abi-Said M, Dloniak SMD (2015) *Hyaena hyaena*. The IUCN Red List of Threatened
870 Species. https://dx.doi.org/10.2305/IUCN.UK.2015-2.RLTS.T10274A45195080.en.

871 Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat
872 Contr 19:716–723. https://doi.org/10.1109/TAC.1974.1100705

873 Akash M, Dheer A, Dloniak SM, Jacobson AP (2021) The faded stripes of Bengal: a
874 historical perspective on the easternmost distribution of the striped hyena. Eur J Wildl
875 Res 67:108. https://doi.org/10.1007/s10344-021-01552-9

876 Allen M, Engström A-S, Meyers S, Handt O, Saldeen T, Von Haeseler A, Pääbo S,
877 Gyllensten U (1998) Mitochondrial DNA sequencing of shed hairs and saliva on robbery
878 caps: sensitivity and matching probabilities. J Forensic Sci 43:453-464.
879 https://doi.org/10.1520/JFS16169J

880 Álvarez-Varas R, González-Acuña D, Vianna JA (2015) Comparative phylogeography of
881 co-distributed *Phrygilus* species (Aves, Thraupidae) from the Central Andes. Mol
882 Phylogenet Evol 90:150–163. https://doi.org/10.1016/j.ympev.2015.04.009

883 Arctander P (1995) Comparison of a mitochondrial gene and a corresponding nuclear
884 pseudogene. Proc Biol Sci 262:13–19. https://doi.org/10.1098/rspb.1995.0170

885 Ashkenazy H, Penn O, Doron-Faigenboim A, Cohen O, Cannarozzi G, Zomer O, Pupko T
886 (2012) FastML: a web server for probabilistic reconstruction of ancestral sequences.
887 Nucleic Acids Res 40:W580–W584. https://doi.org/10.1093/nar/gks498

888 Atkinson QD, Gray RD, Drummond AJ (2009) Bayesian coalescent inference of major
889 human mitochondrial DNA haplogroup expansions in Africa. Proc Biol Sci 276:367–373.
890 https://doi.org/10.1098/rspb.2008.0785

Atkinson QD, Gray RD, Drummond AJ (2008) mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. Mol Biol Evol 25:468–474. https://doi.org/10.1093/molbev/msm277

Ayres DL, Cummings MP, Baele G, Darling AE, Lewis PO, Swofford DL, Huelsenbeck JP, Lemey P, Rambaut A, Suchard MA (2019) BEAGLE 3: improved performance, scaling, and usability for a high-performance computing library for statistical phylogenetics. Syst Biol 68:1052–1061. https://doi.org/10.1093/sysbio/syz020

Battistuzzi FU, Filipski A, Hedges SB, Kumar S (2010) Performance of relaxed-clock methods in estimating evolutionary divergence times and their credibility intervals. Mol Biol Evol 27:1289–1300. https://doi.org/10.1093/molbev/msq014

Bensasson D, Zhang D-X, Hartl DL, Hewitt GM (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends Ecol Evol 16:314–321. https://doi.org/10.1016/s0169-5347(01)02151-6

Berry V, Gascuel O (1996) On the interpretation of bootstrap trees: appropriate threshold of clade selection and induced gain. Mol Biol Evol 13:999–1011. https://doi.org/10.1093/molbev/13.7.999

Bhandari S, Youlatos D, Thapamagar T, Bhusal DR (2021) Shrinking striped hyena (*Hyaena hyaena* Linnaeus, 1758) distribution in Nepal. Eur J Wildl Res 67:10. https://doi.org/10.1007/s10344-020-01437-3

Bromham L, Duchêne S, Hua X, Ritchie AM, Duchêne DA, Ho SYW (2018) Bayesian molecular dating: opening up the black box. Biol Rev 93:1165–1191. https://doi.org/10.1111/brv.12390

913 Brown JW, Smith SA (2018) The past sure is tense: on interpreting phylogenetic
914 divergence time estimates. Syst Biol 67:340–353.
915 https://doi.org/10.1093/sysbio/syx074

916 Brown RP, Yang Z (2011) Rate variation and estimation of divergence times using strict
917 and relaxed clocks. BMC Evol Biol 11:271. https://doi.org/10.1186/1471-2148-11-271

918 Brown WM, George M, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA.
919 Proc Natl Acad Sci USA 76:1967–1971. https://doi.org/10.1073/pnas.76.4.1967

920 Cabanne GS, Sari EHR, Meyer D, Santos FR, Miyaki CY (2013) Matrilineal evidence for
921 demographic expansion, low diversity and lack of phylogeographic structure in the
922 Atlantic forest endemic Greenish *Schiffornis Schiffornis virescens* (Aves: Tityridae). J
923 Ornithol 154:371–384. https://doi.org/10.1007/s10336-012-0901-8

924 Cassens I, Van Waerebeek K, Best PB, Crespo EA, Reyes J, Milinkovitch MC (2003) The
925 phylogeography of dusky dolphins (*Lagenorhynchus obscurus*): a critical examination of
926 network methods and rooting procedures. Mol Ecol 12:1781–1792.
927 https://doi.org/10.1046/j.1365-294X.2003.01876.x

928 Çoğal M, İlemin Y, Sözen M (2021) Status and distribution of the striped hyaena, *Hyaena*
929 *hyaena*, in Turkey: an updated assessment (Carnivora: Mammalia). Turk J Zool 45:131–
930 141. https://doi.org/10.3906/zoo-2007-8

931 Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new
932 heuristics and parallel computing. Nat Methods 9:772.
933 https://doi.org/10.1038/nmeth.2109

934 Derouiche L, Irzagh A, Rahmouni R, Tahri R, Hadjeloum M, Bouhadad R, Fernandes C
935 (2020) Deep mitochondrial DNA phylogeographic divergence in the threatened aoudad

936 *Ammotragus lervia* (Bovidae, Caprini). Gene 739:144510.
937 https://doi.org/10.1016/j.gene.2020.144510

938 Drummond AJ (2005) Bayesian coalescent inference of past population dynamics from
939 molecular sequences. Mol Biol Evol 22:1185–1192.
940 https://doi.org/10.1093/molbev/msi103

941 Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and
942 dating with confidence. PLoS Biol 4:e88. https://doi.org/10.1371/journal.pbio.0040088

943 Drummond AJ, Suchard MA (2010) Bayesian random local clocks, or one rate to rule
944 them all. BMC Biol 8:114. https://doi.org/10.1186/1741-7007-8-114

945 Dubach JM, Briggs MB, White PA, Ament BA, Patterson BD (2013) Genetic perspectives
946 on "Lion Conservation Units" in Eastern and Southern Africa. Conserv Genet 14:741–755.
947 https://doi.org/10.1007/s10592-013-0453-3

948 Ely CR, Wilson RE, Talbot SL (2017) Genetic structure among greater white-fronted
949 goose populations of the Pacific flyway. Ecol Evol 7:2956–2968.
950 https://doi.org/10.1002/ece3.2934

951 Excoffier L (2004) Patterns of DNA sequence diversity and genetic structure after a
952 range expansion: lessons from the infinite-island model. Mol Ecol 13:853–864.
953 https://doi.org/10.1046/j.1365-294x.2003.02004.x

954 Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to
955 perform population genetics analyses under Linux and Windows. Mol Ecol Resour
956 10:564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

957 Excoffier L, Schneider S (1999) Why hunter-gatherer populations do not show signs of
958 Pleistocene demographic expansions. Proc Natl Acad Sci USA 96:10597–10602.
959 https://doi.org/10.1073/pnas.96.19.10597

960 Fahey AL, Ricklefs RE, Dewoody JA (2014) DNA-based approaches for evaluating
961 historical demography in terrestrial vertebrates: evaluating historical demography. Biol
962 J Linn Soc 112:367–386. https://doi.org/10.1111/bij.12259

963 Felsenstein J (1992) Estimating effective population size from samples of sequences:
964 inefficiency of pairwise and segregating sites as compared to phylogenetic estimates.
965 Genet Res 59:139–147. https://doi.org/10.1017/S0016672300030354

966 Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap.
967 Evolution 39:783–791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x

968 Felsenstein J, Churchill GA (1996) A Hidden Markov Model approach to variation among
969 sites in rate of evolution. Mol Biol Evol 13:93–104.
970 https://doi.org/10.1093/oxfordjournals.molbev.a025575

971 Fernandes CA, Ginja C, Pereira I, Tenreiro R, Bruford MW, Santos-Reis M (2008) Species-
972 specific mitochondrial DNA markers for identification of non-invasive samples from
973 sympatric carnivores in the Iberian Peninsula. Conserv Genet 9:681–690.
974 https://doi.org/10.1007/s10592-007-9364-5

975 Ferreira MAR, Suchard MA (2008) Bayesian analysis of elapsed times in
976 continuous-time Markov chains. Can J Statistics 36:355–368.
977 https://doi.org/10.1002/cjs.5550360302

978 Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth,
979 hitchhiking and background selection. Genetics 147:915–925.
980 https://doi.org/10.1093/genetics/147.2.915

981 Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. Genetics 133:693–709.
982 https://doi.org/10.1093/genetics/133.3.693

983 Gill MS, Lemey P, Faria NR, Rambaut A, Shapiro B, Suchard MA (2013) Improving
984 Bayesian population dynamics inference: a coalescent-based model for multiple loci.
985 Mol Biol Evol 30:713–724. https://doi.org/10.1093/molbev/mss265

986 Glez-Peña D, Gómez-Blanco D, Reboiro-Jato M, Fdez-Riverola F, Posada D (2010) ALTER:
987 program-oriented conversion of DNA and protein alignments. Nucleic Acids Res
988 38:W14–W18. https://doi.org/10.1093/nar/gkq321

989 Grant WAS, Bowen BW (1998) Shallow population histories in deep evolutionary
990 lineages of marine fishes: insights from sardines and anchovies and lessons for
991 conservation. J Hered 89:415–426. https://doi.org/10.1093/jhered/89.5.415

992 Grant WS (2015) Problems and cautions with sequence mismatch analysis and Bayesian
993 skyline plots to infer historical demography. J Hered 106:333–346.
994 https://doi.org/10.1093/jhered/esv020

995 Griffiths RC, Tavaré S (1994c) Sampling theory for neutral alleles in a varying
996 environment. Philos Trans R Soc Lond B Biol Sci 344:403–410.
997 https://doi.org/10.1098/rstb.1994.0079

998 Griffiths RC, Tavaré S (1999) The ages of mutations in gene trees. Ann Appl Probab
999 9:567–590 https://doi.org/10.1214/aoap/1029962804

1000 Griffiths RC, Tavaré S (1994b) Ancestral inference in population genetics. Stat Sci 9:307–
1001 319 https://doi.org/10.1214/ss/1177010378

1002 Griffiths RC, Tavare S (1994a) Simulating probability distributions in the coalescent.
1003 Theor Popul Biol 46:131–159. https://doi.org/10.1006/tpbi.1994.1023

1004 Griffiths RC, Tavaré S (1995) Unrooted genealogical tree probabilities in the infinitely-
1005 many-sites model. Math Biosci 127:77–98. https://doi.org/10.1016/0025-
1006 5564(94)00044-Z

1007 Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New
1008 algorithms and methods to estimate maximum-likelihood phylogenies: assessing the
1009 performance of PhyML 3.0. Syst Biol 59:307–321.
1010 https://doi.org/10.1093/sysbio/syq010

1011 Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, Jenkins T, Griffiths RC,
1012 Templeton AR, Zegura SL (1998) Out of Africa and back again: nested cladistic analysis
1013 of human Y chromosome variation. Mol Biol Evol 15:427–441.
1014 https://doi.org/10.1093/oxfordjournals.molbev.a025939

1015 Hanke M, Wink M (1994) Direct DNA sequencing of PCR-amplified vector inserts
1016 following enzymatic degradation of primer and dNTPs. Biotechniques 17:858–860

1017 Harding RM, Fullerton SM, Griffiths RC, Bond J, Cox MJ, Schneider JA, Moulin DS, Clegg JB
1018 (1997) Archaic African and Asian lineages in the genetic ancestry of modern humans.
1019 Am J Hum Genet 60:772–789

1020 Harpending HC (1994) Signature of ancient population growth in a low-resolution
1021 mitochondrial DNA mismatch distribution. Hum Biol 66:591–600

1022 Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular
1023 clock of mitochondrial DNA. J Mol Evol 22:160–174.
1024 https://doi.org/10.1007/BF02101694

1025 Hassan-Beigi Y, Rodrigues M, Adibi MA, Madjdzadeh SM, Vercammen P, Farhadinia MS,
1026 Fernandes CR (2023) Mitochondrial DNA variation of the caracal (*Caracal caracal*) in
1027 Iran and range-wide phylogeographic comparisons. Mamm Biol 103:83–106.
1028 https://doi.org/10.1007/s42991-022-00328-1

1029 Heled J, Bouckaert RR (2013) Looking for trees in the forest: summary tree from
1030 posterior samples. BMC Evol Biol 13:221. https://doi.org/10.1186/1471-2148-13-221

1031 Heled J, Drummond AJ (2008) Bayesian inference of population size history from
1032 multiple loci. BMC Evol Biol 8:289. https://doi.org/10.1186/1471-2148-8-289

1033 Ho SYW, Phillips MJ, Drummond AJ, Cooper A (2005) Accuracy of rate estimation using
1034 relaxed-clock models with a critical focus on the early metazoan radiation. Mol Biol Evol
1035 22:1355–1363. https://doi.org/10.1093/molbev/msi125

1036 Ho SYW, Shapiro B (2011) Skyline-plot methods for estimating demographic history
1037 from nucleotide sequences. Mol Ecol Resour 11:423–434.
1038 https://doi.org/10.1111/j.1755-0998.2011.02988.x

1039 Holder MT, Sukumaran J, Lewis PO (2008) A justification for reporting the majority-rule
1040 consensus tree in Bayesian phylogenetics. Syst Biol 57:814–821.
1041 https://doi.org/10.1080/10635150802422308

1042 Holekamp KE, Kolowski JM (2009) Hyaenidae. In: Wilson DE, Mittermeier RA (eds),
1043 Handbook of the mammals of the world. Volume 1: Carnivores. Lynx Edicions, Madrid,
1044 Spain, pp 234–260

1045 Horreo JL, Palacín C, Alonso JC, Milá B (2013) A link between historical population
1046 decline in the threatened great bustard and human expansion in Iberia: evidence from
1047 genetic and demographic data: bustards' historical population decline. Biol J Linn Soc
1048 110:518–527. https://doi.org/10.1111/bij.12152

1049 Howard WR (1997) A warm future in the past. Nature 388:418–419.
1050 https://doi.org/10.1038/41201

1051 Hu J, Westbury MV, Yuan J, Zhang Z, Chen S, Xiao B, Hou X, Ji H, Lai X, Hofreiter M, Sheng
1052 G (2021) Ancient mitochondrial genomes from Chinese cave hyenas provide insights
1053 into the evolutionary history of the genus *Crocuta*. Proc R Soc B 288:20202934.
1054 https://doi.org/10.1098/rspb.2020.2934

1055 Hudson RR (2007) The variance of coalescent time estimates from DNA sequences. J Mol
1056 Evol 64:702–705. https://doi.org/10.1007/s00239-006-0261-1

1057 Huelsenbeck JP, Larget B, Alfaro ME (2004) Bayesian phylogenetic model selection using
1058 reversible jump Markov chain Monte Carlo. Mol Biol Evol 21:1123–1133.
1059 https://doi.org/10.1093/molbev/msh123

1060 Huggins PM, Li W, Haws D, Friedrich T, Liu J, Yoshida R (2011) Bayes estimators for
1061 phylogenetic reconstruction. Syst Biol 60:528–540.
1062 https://doi.org/10.1093/sysbio/syr021

1063 Hurvich CM, Tsai C-L (1989) Regression and time series model selection in small
1064 samples. Biometrika 76:297–307. https://doi.org/10.1093/biomet/76.2.297

1065 Ingman M, Kaessmann H, Pääbo S, Gyllensten U (2000) Mitochondrial genome variation
1066 and the origin of modern humans. Nature 408:708–713.
1067 https://doi.org/10.1038/35047064

1068 Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the cytochrome b gene of
1069 mammals. J Mol Evol 32:128–144. https://doi.org/10.1007/BF02515385

1070 Jansen T, Forster P, Levine MA, Oelke H, Hurles M, Renfrew C, Weber J, Olek K (2002)
1071 Mitochondrial DNA and the origins of the domestic horse. Proc Natl Acad Sci USA
1072 99:10905–10910. https://doi.org/10.1073/pnas.152330099

1073 Jeffreys H (1935) Some tests of significance, treated by the theory of probability. Math
1074 Proc Camb Phil Soc 31:203–222. https://doi.org/10.1017/S030500410001330X

1075 Jenks SM, Werdelin L (1998) Taxonomy and systematics of living hyaenas (Family
1076 Hyaenidae). In: Mills MGL, Hofer H (eds) Hyaenas: status survey and conservation action
1077 plan. IUCN/SSC Hyaena Specialist Group, IUCN, Gland, Switzerland, pp 8–17

1078 Jensen EL, Miller JM, Edwards DL, Garrick RC, Tapia W, Caccone A, Russello MA (2018)
1079 Temporal mitogenomics of the Galapagos giant tortoise from Pinzón reveals potential
1080 biases in population genetic inference. J Hered 109:631–640.
1081 https://doi.org/10.1093/jhered/esy016

1082 Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS (2017)
1083 ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods
1084 14:587–589. https://doi.org/10.1038/nmeth.4285

1085 Karafet TM, Zegura SL, Posukh O, Osipova L, Bergen A, Long J, Goldman D, Klitz W,
1086 Harihara S, de Knijff P, Wiebe V, Griffiths RC, Templeton AR, Hammer MF (1999)
1087 Ancestral Asian source(s) of New World Y-chromosome founder haplotypes. Am J Hum
1088 Genet 64:817–831. https://doi.org/10.1086/302282

1089 Kasparek M, Kasparek A, Gözcelioğlu B, Çolak E, Yiğit N (2004) On the status and
1090 distribution of the striped hyaena, *Hyaena hyaena* , in Turkey. Zool Middle East 33:93–
1091 108. https://doi.org/10.1080/09397140.2004.10638068

1092 Kass RE, Raftery AE (1995) Bayes Factors. J Am Stat Assoc 90:773–795.
1093 https://doi.org/10.1080/01621459.1995.10476572

1094 Keis M, Remm J, Ho SYW, Davison J, Tammeleht E, Tumanov IL, Saveljev AP, Männil P,
1095 Kojola I, Abramov AV, Margus T, Saarma U (2013) Complete mitochondrial genomes and
1096 a novel spatial genetic method reveal cryptic phylogeographical structure and migration
1097 patterns among brown bears in north-western Eurasia. J Biogeogr 40:915–927.
1098 https://doi.org/10.1111/jbi.12043

1099 Khorozyan I, Malkhasyan A, Murtskhvaladze M (2011) The striped hyaena *Hyaena*
1100 *hyaena* (Hyaenidae, Carnivora) rediscovered in Armenia. Folia Zool 60:253–261.
1101 https://doi.org/10.25225/fozo.v60.i3.a10.2011

1102 Kimura M (1969) The number of heterozygous nucleotide sites maintained in a finite
1103 population due to steady flux of mutations. Genetics 61:893–903.
1104 https://doi.org/10.1093/genetics/61.4.893

1105 Kimura M (1980) A simple method for estimating evolutionary rates of base
1106 substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–
1107 120. https://doi.org/10.1007/BF01731581

1108 Kingdon J (2001) The Kingdon Field Guide to African Mammals. Academic Press, London

1109 Kleinen T, Hildebrandt S, Prange M, Rachmayani R, Müller S, Bezrukova E, Brovkin V,
1110 Tarasov PE (2014) The climate and vegetation of Marine Isotope Stage 11 – model
1111 results and proxy-based reconstructions at global and regional scale. Quat Int 348:247–
1112 265. https://doi.org/10.1016/j.quaint.2013.12.028

1113 Koepfli K-P, Jenks SM, Eizirik E, Zahirpour T, Valkenburgh BV, Wayne RK (2006)
1114 Molecular systematics of the Hyaenidae: relationships of a relictual lineage resolved by a
1115 molecular supermatrix. Mol Phylogenet Evol 38:603–620.
1116 https://doi.org/10.1016/j.ympev.2005.10.017

1117 Kowalski K, Rzebik-Kowalska B (1991) Mammals of Algeria. Polish Academy of Science,
1118 Institute of Systematics and Evolution of Animals, Wrocław

1119 Kuhner MK (2006) LAMARC 2.0: maximum likelihood and Bayesian estimation of
1120 population parameters. Bioinformatics 22:768–770.
1121 https://doi.org/10.1093/bioinformatics/btk051

1122 Kuhner MK, Smith LP (2007) Comparing likelihood and Bayesian coalescent estimation
1123 of population parameters. Genetics 175:155–165.
1124 https://doi.org/10.1534/genetics.106.056457

1125 Kuhner MK, Yamato J, Felsenstein J (1998) Maximum likelihood estimation of
1126 population growth rates based on the coalescent. Genetics 149:429–434.
1127 https://doi.org/10.1093/genetics/149.1.429

1128 Kuo H-C, Chen S-F, Fang Y-P, Flanders J, Rossiter SJ (2014) Comparative rangewide
1129 phylogeography of four endemic Taiwanese bat species. Mol Ecol 23:3566–3586.
1130 https://doi.org/10.1111/mec.12838

1131 Kvist L, Giralt D, Valera F, Hoi H, Kristin A, Darchiashvili G, Lovaszi P (2011) Population
1132 decline is accompanied by loss of genetic diversity in the Lesser Grey Shrike *Lanius*
1133 *minor*. Ibis 153:98–109. https://doi.org/10.1111/j.1474-919X.2010.01091.x

1134 Lanave C, Preparata G, Sacone C, Serio G (1984) A new method for calculating
1135 evolutionary substitution rates. J Mol Evol 20:86–93.
1136 https://doi.org/10.1007/BF02101990

1137 Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B (2016) PartitionFinder 2: new
1138 methods for selecting partitioned models of evolution for molecular and morphological
1139 phylogenetic analyses. Mol Biol Evol msw260.
1140 https://doi.org/10.1093/molbev/msw260

1141 Lartillot N, Philippe H (2006) Computing Bayes Factors using thermodynamic
1142 integration. Syst Biol 55:195–207. https://doi.org/10.1080/10635150500433722

1143 Leigh JW, Bryant D (2015) POPART: full-feature software for haplotype network
1144 construction. Methods Ecol Evol 6:1110–1116. https://doi.org/10.1111/2041-
1145 210X.12410

1146 Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA
1147 polymorphism data. Bioinformatics 25:1451–1452.
1148 https://doi.org/10.1093/bioinformatics/btp187

1149 Lisiecki LE, Raymo ME (2005) A Pliocene-Pleistocene stack of 57 globally distributed
1150 benthic $\delta^{18}O$ records. Paleoceanography 20:2004PA001071.
1151 https://doi.org/10.1029/2004PA001071

1152 Low Cunningham P (2004) Checklist and status of the terrestrial mammals from the
1153 United Arab Emirates. Zool Middle East 33:7–20.
1154 https://doi.org/10.1080/09397140.2004.10638059

1155 Mello B, Tao Q, Barba-Montoya J, Kumar S (2021) Molecular dating for phylogenies
1156 containing a mix of populations and species by using Bayesian and RelTime approaches.
1157 Mol Ecol Resour 21:122–136. https://doi.org/10.1111/1755-0998.13249

1158 Mills MGL, Hofer H (1998) Hyaenas: status survey and conservation action plan.
1159 IUCN/SSC Hyaena Specialist Group, IUCN, Gland, Switzerland

1160 Minin VN, Bloomquist EW, Suchard MA (2008) Smooth skyride through a rough skyline:
1161 Bayesian coalescent-based inference of population dynamics. Mol Biol Evol 25:1459–
1162 1471. https://doi.org/10.1093/molbev/msn090

1163 Mirol PM, Mascheretti S, Searle JB (2000) Multiple nuclear pseudogenes of
1164 mitochondrial cytochrome b in *Ctenomys* (Caviomorpha, Rodentia) with either great
1165 similarity to or high divergence from the true mitochondrial sequence. Heredity 84:538–
1166 547. https://doi.org/10.1046/j.1365-2540.2000.00689.x

1167 Morrison DA (2008) How to summarize estimates of ancestral divergence times. Evol
1168 Bioinform Online 4:EBO.S545. https://doi.org/10.4137/EBO.S545

1169 Mulcahy DG, Noonan BP, Moss T, Townsend TM, Reeder TW, Sites JW, Wiens JJ (2012)
1170 Estimating divergence dates and evaluating dating methods using phylogenomic and
1171 mitochondrial data in squamate reptiles. Mol Phylogenet Evol 65:974–991.
1172 https://doi.org/10.1016/j.ympev.2012.08.018

1173 Nei M (1987) Molecular Evolutionary Genetics. Columbia University Press, New York

1174 Nei M (1985) Human evolution at the molecular level. In: Ohta T, Aoki K (eds)
1175 Population Genetics and Molecular Evolution. Springer-Verlag, Berlin, pp 41–64.

1176 Nguyen L-T, Schmidt HA, Von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective
1177 stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol
1178 32:268–274. https://doi.org/10.1093/molbev/msu300

1179 Opgen-Rhein R, Fahrmeir L, Strimmer K (2005) Inference of demographic history from
1180 genealogical trees using reversible jump Markov chain Monte Carlo. BMC Evol Biol 5:6.
1181 https://doi.org/10.1186/1471-2148-5-6

1182 Pacifici M, Santini L, Marco MD, Baisero D, Francucci L, Marasini GG, Visconti P,
1183 Rondinini C (2013) Generation length for mammals. Nature Conservation 5:89–94.
1184 https://doi.org/10.3897/natureconservation.5.5734

1185 Paradis E (2010) pegas: an R package for population genetics with an integrated–
1186 modular approach. Bioinformatics 26:419–420.
1187 https://doi.org/10.1093/bioinformatics/btp696

1188 Paradis E, Schliep K (2019) ape 5.0: an environment for modern phylogenetics and
1189 evolutionary analyses in R. Bioinformatics 35:526–528.
1190 https://doi.org/10.1093/bioinformatics/bty633

1191 Pilkington MM, Wilder JA, Mendez FL, Cox MP, Woerner A, Angui T, Kingan S, Mobasher
1192 Z, Batini C, Destro-Bisol G, Soodyall H, Strassmann BI, Hammer MF (2008) Contrasting
1193 signatures of population growth for mitochondrial DNA and Y chromosomes among
1194 human populations in Africa. Mol Biol Evol 25:517–525.
1195 https://doi.org/10.1093/molbev/msm279

1196 Pocock RI (1934) The races of the striped and brown hyænas. Proceedings of the
1197 Zoological Society of London 104:799–825. https://doi.org/10.1111/j.1096-
1198 3642.1934.tb01667.x

1199 Pybus OG, Rambaut A (2002) GENIE: estimating demographic history from molecular
1200 phylogenies. Bioinformatics 18:1404–1405.
1201 https://doi.org/10.1093/bioinformatics/18.10.1404

1202 Pybus OG, Rambaut A, Harvey PH (2000) An integrated framework for the inference of
1203 viral population history from reconstructed genealogies. Genetics 155:1429–1437.
1204 https://doi.org/10.1093/genetics/155.3.1429

1205 R Core Team (2023) R: a language and environment for statistical computing. R
1206 Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

1207 Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarization
1208 in Bayesian phylogenetics using Tracer 1.7. Syst Biol 67:901–904.
1209 https://doi.org/10.1093/sysbio/syy032

1210 Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against
1211 population growth. Mol Biol Evol 19:2092–2100.
1212 https://doi.org/10.1093/oxfordjournals.molbev.a004034

1213 Ray N, Currat M, Excoffier L (2003) Intra-deme molecular diversity in spatially
1214 expanding populations. Mol Biol Evol 20:76–86.
1215 https://doi.org/10.1093/molbev/msg009

1216 Rieger I (1979) A review of the biology of striped hyaenas, *Hyaena hyaena* (Linné, 1758).
1217 Saeugetierkd Mitt 27:81–95

1218 Rieger I (1981) Hyaena hyaena. Mamm Species 150:1–5.
1219 https://doi.org/10.2307/41353899

1220 Ritchie AM, Lo N, Ho SYW (2017) The impact of the tree prior on molecular dating of
1221 data sets containing a mixture of inter- and intraspecies sampling. Syst Biol 66:413–425.
1222 https://doi.org/10.1093/sysbio/syw095

1223 Robinson M, Gouy M, Gautier C, Mouchiroud D (1998) Sensitivity of the relative-rate test
1224 to taxonomic sampling. Mol Biol Evol 15:1091–1098.
1225 https://doi.org/10.1093/oxfordjournals.molbev.a026016

1226 Robinson-Rechavi M, Huchon D (2000) RRTree: Relative-rate tests between groups of
1227 sequences on a phylogenetic tree. Bioinformatics 16:296–297.
1228 https://doi.org/10.1093/bioinformatics/16.3.296

1229 Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of
1230 pairwise genetic differences. Mol Biol Evol 9:552–569.
1231 https://doi.org/10.1093/oxfordjournals.molbev.a040727

1232 Rohland N, Pollack JL, Nagel D, Beauval C, Airvaux J, Pääbo S, Hofreiter M (2005) The
1233 population history of extant and extinct hyenas. Mol Biol Evol 22:2435–2443.
1234 https://doi.org/10.1093/molbev/msi244

1235 Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L,
1236 Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic
1237 inference and model choice across a large model space. Syst Biol 61:539–542.
1238 https://doi.org/10.1093/sysbio/sys029

1239 Sanderson MJ (2003) r8s: inferring absolute rates of molecular evolution and
1240 divergence times in the absence of a molecular clock. Bioinformatics 19:301–302.
1241 https://doi.org/10.1093/bioinformatics/19.2.301

1242 Schmidt HA, Strimmer K, Vingron M, Von Haeseler A (2002) TREE-PUZZLE: maximum
1243 likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics
1244 18:502–504. https://doi.org/10.1093/bioinformatics/18.3.502

1245 Schmidt HA, Von Haeseler A, Buschbom J (2007) pIPHULA—parallel inference of
1246 population parameters using a likelihood approach. Bioinformatics 23:2636–2637.
1247 https://doi.org/10.1093/bioinformatics/btm391

1248 Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the
1249 distribution of pairwise differences when the mutation rates vary among sites:
1250 application to human mitochondrial DNA. Genetics 152:1079–1089.
1251 https://doi.org/10.1093/genetics/152.3.1079

1252 Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6:461–464.
1253 https://doi.org/10.1214/aos/1176344136

1254 Smith MF, Thomas WK, Patton JL (1992) Mitochondrial DNA-like sequence in the
1255 nuclear genome of an akodontine rodent. Mol Biol Evol 9:204–215.
1256 https://doi.org/10.1093/oxfordjournals.molbev.a040714

1257 Stothard P (2000) The Sequence Manipulation Suite: JavaScript programs for analyzing
1258 and formatting protein and DNA sequences. BioTechniques 28:1102–1104.
1259 https://doi.org/10.2144/00286ir01

1260 Strimmer K, Pybus OG (2001) Exploring the demographic history of DNA sequences
1261 using the generalized skyline plot. Mol Biol Evol 18:2298–2305.
1262 https://doi.org/10.1093/oxfordjournals.molbev.a003776

1263 Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A (2018) Bayesian
1264 phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol 4:vey016.
1265 https://doi.org/10.1093/ve/vey016

1266 Sugiura N (1978) Further analysis of the data by Akaike's information criterion and the
1267 finite corrections. Commun Stat - Theory Methods 7:13–26.
1268 https://doi.org/10.1080/03610927808827599

1269 Sukumaran J, Holder MT (2010) DendroPy: a Python library for phylogenetic computing.
1270 Bioinformatics 26:1569–1571. https://doi.org/10.1093/bioinformatics/btq228

1271 Tajima F (1993) Simple methods for testing the molecular evolutionary clock hypothesis.
1272 Genetics 135:599–607. https://doi.org/10.1093/genetics/135.2.599

1273 Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA
1274 polymorphism. Genetics 123:585–595. https://doi.org/10.1093/genetics/123.3.585

1275 Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the
1276 control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–
1277 526. https://doi.org/10.1093/oxfordjournals.molbev.a040023

1278 Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics
1279 Analysis Version 11. Mol Biol Evol 38:3022–3027.
1280 https://doi.org/10.1093/molbev/msab120

1281 Tao Q, Tamura K, U. Battistuzzi F, Kumar S (2019) A machine learning method for
1282 detecting autocorrelation of evolutionary rates in large phylogenies. Mol Biol Evol
1283 36:811–824. https://doi.org/10.1093/molbev/msz014

1284 Thomson R, Pritchard JK, Shen P, Oefner PJ, Feldman MW (2000) Recent common
1285 ancestry of human Y chromosomes: evidence from DNA sequence data. Proc Natl Acad
1286 Sci USA 97:7360–7365. https://doi.org/10.1073/pnas.97.13.7360

1287 Tourani M, Moqanaki EM, Kiabi BH (2012) Vulnerability of striped hyaenas, *Hyaena*
1288 *hyaena* , in a human-dominated landscape of Central Iran. Zool Middle East 56:133–136.
1289 https://doi.org/10.1080/09397140.2012.10648948

1290 Triant DA, DeWoody JA (2007) The occurrence, detection, and avoidance of
1291 mitochondrial DNA translocations in mammalian systematics and phylogeography. J
1292 Mammal 88:908–920. https://doi.org/10.1644/06-MAMM-A-204R1.1

1293 Tzedakis PC, Hodell DA, Nehrbass-Ahles C, Mitsui T, Wolff EW (2022) Marine Isotope
1294 Stage 11c: an unusual interglacial. Quat Sci Rev 284:107493.
1295 https://doi.org/10.1016/j.quascirev.2022.107493

1296 Vigilant L (1999) An evaluation of techniques for the extraction and amplification of
1297 DNA from naturally shed hairs. Biol Chem 380:1329–1331.
1298 https://doi.org/10.1515/BC.1999.169

1299 Villanea FA, Kitchen A, Kemp BM (2019) Applications of Bayesian skyline plots and
1300 approximate Bayesian computation for human demography. Hum Biol 91:279-296.
1301 https://doi.org/10.13110/humanbiology.91.4.04

1302 Villesen P (2007) FaBox: an online toolbox for fasta sequences. Mol Ecol Notes 7:965–
1303 968. https://doi.org/10.1111/j.1471-8286.2007.01821.x

1304 Watterson GA (1975) On the number of segregating sites in genetical models without
1305 recombination. Theor Popul Biol 7:256–276. https://doi.org/10.1016/0040-
1306 5809(75)90020-9

1307 Weiss G, Von Haeseler A (1998) Inference of population history using a likelihood
1308 approach. Genetics 149:1539–1546. https://doi.org/10.1093/genetics/149.3.1539

1309 Werdelin L (2003) Carnivora from the Kanapoi hominid site, Turkana Basin, northern
1310 Kenya, in: Harris JM, Leakey MG (eds), Geology and Vertebrate Paleontology of the Early
1311 Pliocene Site of Kanapoi, northern Kenya. Contrib Sci 498:115–132

1312 Werdelin L, Manthi FK (2012) Carnivora from the Kanapoi hominin site, northern Kenya.
1313 J Afr Earth Sci 64:1–8. https://doi.org/10.1016/j.jafrearsci.2011.11.003

1314 Werdelin L, Solounias N (1991) The Hyaenidae: taxonomy, systematics and evolution.
1315 Fossils and Strata 30:1-104. https://doi.org/10.18261/8200374815-1991

1316 Werdelin L, Turner A, Solounias N (1994) Studies of fossil hyaenids: the genera
1317 *Hyaenictis* Gaudry and *Chasmaporthetes* Hay, with a reconsideration of the Hyaenidae of
1318 Langebaanweg, South Africa. Zool J Linn Soc 111:197–217.
1319 https://doi.org/10.1111/j.1096-3642.1994.tb01483.x

1320 Werle E, Schneider C, Renner M, Völker M, Fiehn W (1994) Convenient single-step, one
1321 tube purification of PCR products for direct sequencing. Nucleic Acids Res 22:4354–
1322 4355. https://doi.org/10.1093/nar/22.20.4354

1323 Westbury MV, Cahsan BD, Dalerum F, Norén K, Hofreiter M (2019) Aardwolf population
1324 diversity and phylogenetic positioning inferred using complete mitochondrial genomes.
1325 Afr J Wildl Res 49:27-33. https://doi.org/10.3957/056.049.0027

1326 Westbury MV, Hartmann S, Barlow A, Preick M, Ridush B, Nagel D, Rathgeber T, Ziegler
1327 R, Baryshnikov G, Sheng G, Ludwig A, Wiesel I, Dalen L, Bibi F, Werdelin L, Heller R,
1328 Hofreiter M (2020) Hyena paleogenomes reveal a complex evolutionary history of cross-
1329 continental gene flow between spotted and cave hyena. Sci Adv 6:eaay0456.
1330 https://doi.org/10.1126/sciadv.aay0456

1331 Westbury MV, Hartmann S, Barlow A, Wiesel I, Leo V, Welch R, Parker DM, Sicks F,
1332 Ludwig A, Dalén L, Hofreiter M (2018) Extended and continuous decline in effective
1333 population size results in low genomic diversity in the world's rarest hyena species, the
1334 brown hyena. Mol Biol Evol 35:1225–1237. https://doi.org/10.1093/molbev/msy037

1335 Westbury MV, Le Duc D, Duchêne DA, Krishnan A, Prost S, Rutschmann S, Grau JH, Dalén
1336 L, Weyrich A, Norén K, Werdelin L, Dalerum F, Schöneberg T, Hofreiter M (2021)
1337 Ecological specialization and evolutionary reticulation in extant Hyaenidae. Mol Biol
1338 Evol 38:3884–3897. https://doi.org/10.1093/molbev/msab055

Wilkinson CE, Dheer A, Zett T, Torrents-Ticó M, Yarnell RW, Bar Ziv E, Bhandari S, Jacobson A, Dloniak SM (2024) Review of the global research on Hyaenidae and implications for conservation and management. Mammal Rev 54:193-212. https://doi.org/10.1111/mam.12337

Wilson AC, Cann RL, Carr SM, George M, Gyllensten UB, Helm-Bychowski KM, Higuchi RG, Palumbi SR, Prager EM, Sage RD, Stoneking M (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. Biol J Linn Soc 26:375–400. https://doi.org/10.1111/j.1095-8312.1985.tb02048.x

Wozencraft WC (2005) Order Carnivora. In: Wilson DE, Reeder, D.M. (eds), Mammal Species of the World: A Taxonomic and Geographic Reference, 3rd edn. Johns Hopkins University Press, Baltimore, pp 532–628

Xiang F, Dearlove B, Frost S (2019) genieR: An R package for inference of demographic history of phylogenies. J Open Source Softw 4:634. https://doi.org/10.21105/joss.00634

Xie W, Lewis PO, Fan Y, Kuo L, Chen M-H (2011) Improving marginal likelihood estimation for Bayesian phylogenetic model selection. Syst Biol 60:150–160. https://doi.org/10.1093/sysbio/syq085

Zhang DX, Hewitt GM (1996) Nuclear integrations: challenges for mitochondrial DNA markers. Trends Ecol Evol 11:247–251. https://doi.org/10.1016/0169-5347(96)10031-8

Figures and Tables

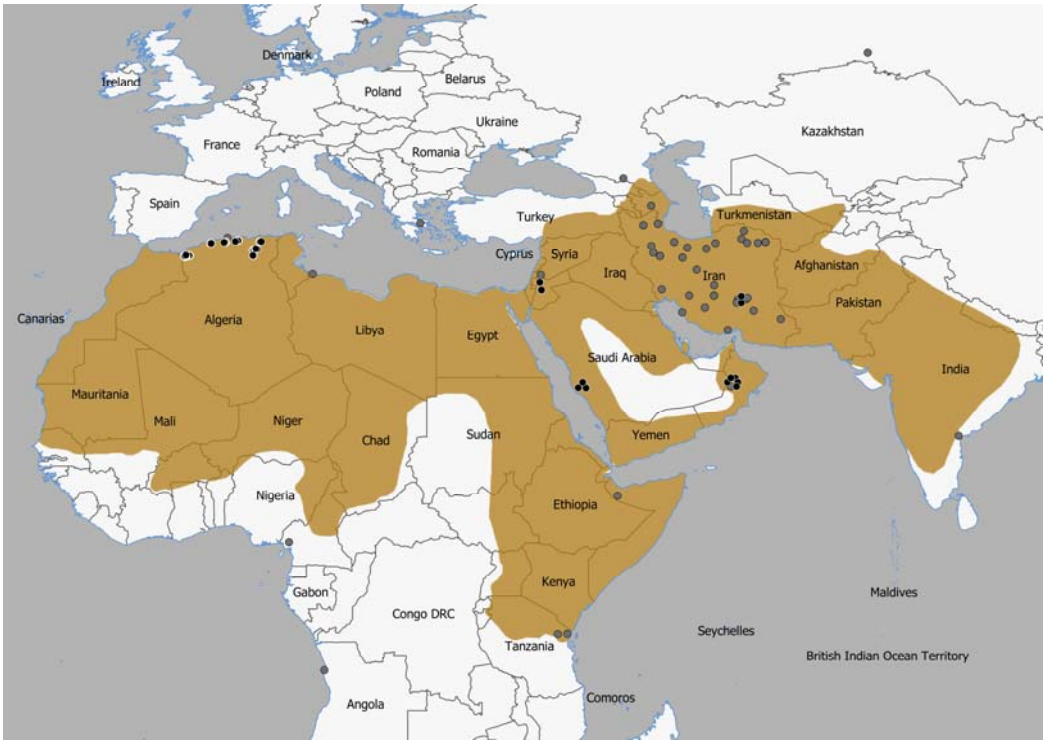


Fig. 1. Map showing the geographic range of the striped hyena according to Abi-Said and Dloniak (2015). The plotted dots indicate the origin of the samples and sequences analysed; black dots in the case of our samples and dark grey dots for Genbank sequences (see Table 1 for details).

Table 1. Information on the samples and previously published *Cyt b* sequences of striped hyenas and outgroups used in this study.

Sample code	Country	Location	Haplotype (648 bp alignment)	Haplotype (340 bp alignment)	GenBank accession number	Reference
Hy1	Algeria	Chlef	H1	H1	PP328974	This study
Hy2	Algeria	Batna	H1	H1	PP328974	This study
Hy3	Algeria	Tlemcen	H1	H1	PP328974	This study
Hy4	Algeria	Tlemcen	H1	H1	PP328974	This study
Hy5	Algeria	Médéa	H1	H1	PP328974	This study
Hy6	Algeria	Bouïra	H1	H1	PP328974	This study
Hy7	Algeria	Tlemcen	H1	H1	PP328974	This study
Hy9	Algeria	Biskra	H1	H1	PP328974	This study
Hy10	Algeria	Chlef	H1	H1	PP328974	This study
Hy11	Algeria	Bouïra	H1	H1	PP328974	This study
Hy13	Algeria	Constantine	H1	H1	PP328974	This study
Hy16	Algeria	Bouïra	H1	H1	PP328974	This study
Hy18	Algeria	Tlemcen	H1	H1	PP328974	This study
Hy21	Algeria	Médéa	H1	H1	PP328974	This study
Hy22	Algeria	Batna	H1	H1	PP328974	This study
Hy27	Algeria	Batna	H1	H1	PP328974	This study
Hy28	Algeria	Tlemcen	H1	H1	PP328974	This study
Hy29	Algeria	Batna	H1	H1	PP328974	This study
Hy30	Algeria	Tlemcen	H1		PP328974	This study
Hy31	Algeria	Tlemcen	H1		PP328974	This study

Sample code	Country	Location	Haplotype (648 bp alignment)	Haplotype (340 bp alignment)	GenBank accession number	Reference
Hy32	Algeria	Tlemcen	H1	H1	PP328974	This study
Hy33	Algeria	Tlemcen	H1	H1	PP328974	This study
Hy34	Algeria	Tlemcen	H1	H1	PP328974	This study
Hy35	Algeria	Tlemcen	H1	H1	PP328974	This study
Belezma_YRhigi	Algeria	Batna		H1	= PP328974	This study
	Algeria	Algiers		H1	DQ157581	Rohland et al. 2005
	Libya	Tripoli		H1	DQ157585	Rohland et al. 2005
	Cameroon			H1	DQ157579	Rohland et al. 2005
	Angola	Luanda		H7	DQ157582	Rohland et al. 2005
	Tanzania	NE Tanzania		H4	DQ157583	Rohland et al. 2005
	Tanzania		H11	H5	AY048787	Albert 2001
	Somalia	Somaliland		H5	DQ157584	Rohland et al. 2005
Hyena_Amr1	Jordan	Qatraneh		H3	= KP789314	This study
Hyena_Amr2	Jordan	Abu Nusair		H3	= KP789314	This study
	Syria			H4	DQ157577	Rohland et al. 2005
	Greece	Thebes		H5	DQ157578	Rohland et al. 2005
00-01C7-1926	Saudi Arabia	Taif	H2		PP328975	This study
00-01C7-F7AE	Saudi Arabia	Taif		H2	= PP328975	This study
Phiri	Saudi Arabia	Taif	H2		PP328975	This study
Arnold	Oman		H3	H1	= MT893967	This study
Buwayter	Oman		H3	H1	= MT893967	This study

Sample code	Country	Location	Haplotype (648 bp alignment)	Haplotype (340 bp alignment)	GenBank accession number	Reference
Dopey	Oman		H3	H1	= MT893967	This study
Luna	Oman		H3	H1	= MT893967	This study
Sirius	Oman		H3	H1	= MT893967	This study
	Oman			H3	KP789314	Unpublished
Hh1_Massoud	Iran	Kerman	H4	H6	= MT893978	This study
20_5_7_Massoud	Iran	Kerman	H4	H6	= MT893978	This study
	Iran	Razavi Khorasan	H3		MT893967	Dadashi-Jourdehi et al. 2022
	Iran	Qorveh	H3		MT893969	Dadashi-Jourdehi et al. 2022
	Iran	Khash	H3		MT893970	Dadashi-Jourdehi et al. 2022
	Iran	Bushehr	H3		MT893972	Dadashi-Jourdehi et al. 2022
	Iran	Nishapur	H3		MT893973	Dadashi-Jourdehi et al. 2022
	Iran	Damghan	H3		MT893974	Dadashi-Jourdehi et al. 2022
	Iran	Khuzestan	H3		MT893975	Dadashi-Jourdehi et al. 2022
	Iran	Kerman	H3		MT893984	Dadashi-Jourdehi et al. 2022
	Iran	Qom	H3		MT893989	Dadashi-Jourdehi et al. 2022
	Iran	Qazvin	H3		MT893991	Dadashi-Jourdehi et al. 2022
	Iran	Tehran	H3		MT893993	Dadashi-Jourdehi et al. 2022
	Iran	Qom	H3		MT893994	Dadashi-Jourdehi et al. 2022
	Iran	Hamedan	H3		MT893997	Dadashi-Jourdehi et al. 2022
	Iran	Razavi Khorasan	H4		MT893978	Dadashi-Jourdehi et al. 2022
	Iran	Razavi Khorasan	H4		MT893979	Dadashi-Jourdehi et al. 2022

Sample code	Country	Location	Haplotype (648 bp alignment)	Haplotype (340 bp alignment)	GenBank accession number	Reference
	Iran	Sabzevar	H4		MT893983	Dadashi-Jourdehi et al. 2022
	Iran	Joghatai	H4		MT893988	Dadashi-Jourdehi et al. 2022
	Iran	North Khorasan	H4		MT893990	Dadashi-Jourdehi et al. 2022
	Iran	Bardsir	H4		MT893996	Dadashi-Jourdehi et al. 2022
	Iran	Khatam	H4		MT893998	Dadashi-Jourdehi et al. 2022
	Iran	Bijar	H4		MT893999	Dadashi-Jourdehi et al. 2022
	Iran	Kerman	H4		MT894000	Dadashi-Jourdehi et al. 2022
	Iran	Khatam	H4		MT894001	Dadashi-Jourdehi et al. 2022
	Iran	Bam	H4		MT894002	Dadashi-Jourdehi et al. 2022
	Iran	Yasuj	H5		MT893965	Dadashi-Jourdehi et al. 2022
	Iran	Ardestan	H5		MT893966	Dadashi-Jourdehi et al. 2022
	Iran	Ardabil	H5		MT893968	Dadashi-Jourdehi et al. 2022
	Iran	East Azerbaijan	H5		MT893971	Dadashi-Jourdehi et al. 2022
	Iran	Khamir	H5		MT893976	Dadashi-Jourdehi et al. 2022
	Iran	Fars	H5		MT893977	Dadashi-Jourdehi et al. 2022
	Iran	Semnan	H5		MT893981	Dadashi-Jourdehi et al. 2022
	Iran	East Azerbaijan	H5		MT893985	Dadashi-Jourdehi et al. 2022
	Iran	Ardabil	H5		MT893987	Dadashi-Jourdehi et al. 2022
	Iran	Hamedan	H5		MT893992	Dadashi-Jourdehi et al. 2022
	Iran	Razavi Khorasan	H6		MT893980	Dadashi-Jourdehi et al. 2022
	Iran	Taft	H7		MT893982	Dadashi-Jourdehi et al. 2022

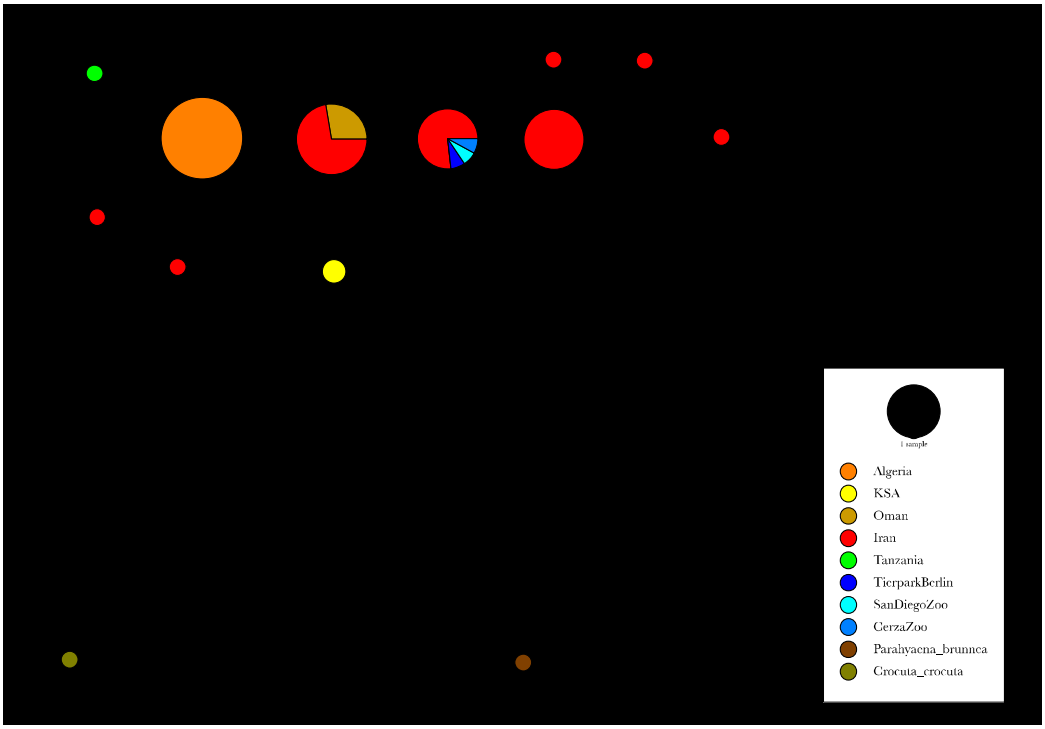
Sample code	Country	Location	Haplotype (648 bp alignment)	Haplotype (340 bp alignment)	GenBank accession number	Reference
	Iran	Khatam	H8		MT893986	Dadashi-Jourdehi et al. 2022
	Iran	Shahdad	H9		MT893995	Dadashi-Jourdehi et al. 2022
	Iran	Hamedan	H10		MT894003	Dadashi-Jourdehi et al. 2022
	Azerbaijan			H1	DQ157587	Rohland et al. 2005
	Russia	Vladikavkaz		H1	DQ157576	Rohland et al. 2005
	Russia	Siberia		H1	DQ157586	Rohland et al. 2005
	India			H1	DQ157580	Rohland et al. 2005
		Tierpark Berlin	H5	H3	AY048788	Albert 2001
		San Diego Zoo	H5	H3	AY928678	Koepfli et al. 2006
		Cerza Zoo	H5	H3	JF894376	Bon et al. 2012
Outgroups						
Parahyaena brunnea				Parahyaena_brunnea	NC038159	Westbury et al. 2018
Crocuta crocuta				Crocuta_crocuta	NC020670	Bon et al. 2012

Table 2. Estimates of mitochondrial genetic diversity and mutation-drift equilibrium tests in striped hyenas, based on the two *Cyt b* alignments with different sequences and lengths

	S	**n_H**	**h**	**π**	***D***	***F_s***	***R_2***
648 bp dataset (n = 76)	13	11	0.794 ± 0.022	0.0026 ± 0.0017	-0.993 (P = 0.166)	-2.662 (P = 0.126)	0.064 (P = 0.181)
340 bp dataset (n = 50)	5	7	0.498 ± 0.082	0.0020 ± 0.0017	-0.922 (P = 0.198)	-3.345 (P = 0.020)	0.071 (P = 0.149)

n, number of samples; S, number of segregating sites; n_H, number of haplotypes; h, haplotype diversity (mean and standard deviation); π, nucleotide diversity (mean and standard deviation); *D*, Tajima's *D* statistic (Tajima 1989); F_S, Fu's F_S statistic (Fu 1997); R_2, Ramos-Onsins and Rozas' R_2 statistic (Ramos-Onsins and Rozas 2002); P, *p*-value of the test.

a)


1 sample
Algeria
KSA
Oman
Iran
Tanzania
TierparkBerlin
SanDiegoZoo
CerzaZoo
Parahyaena_brunnea
Crocuta_crocuta

b)

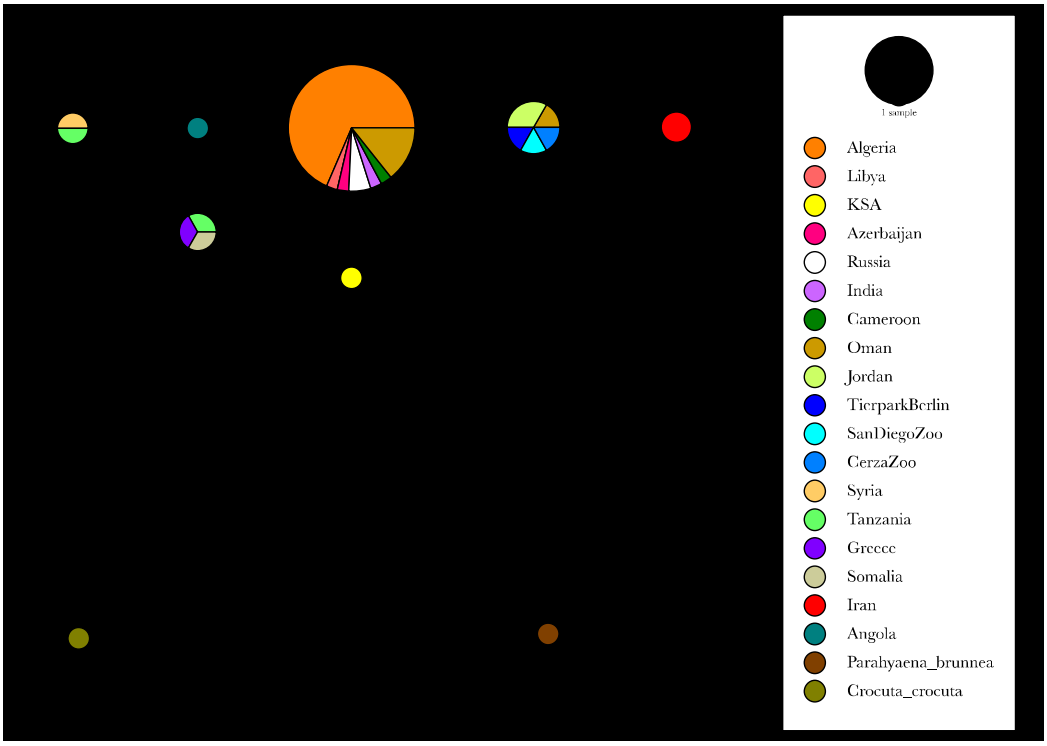

1 sample
Algeria
Libya
KSA
Azerbaijan
Russia
India
Cameroon
Oman
Jordan
TierparkBerlin
SanDiegoZoo
CerzaZoo
Syria
Tanzania
Greece
Somalia
Iran
Angola
Parahyaena_brunnea
Crocuta_crocuta

Fig. 2. IntNJ networks of (a) 648-bp and (b) 340-bp striped hyena *Cyt b* haplotypes, rooted with two outgroups (brown hyena and spotted hyena). The first contains 76 striped hyena sequences and the second contains 50. Circles represent haplotypes and their size is proportional to frequency. Circles are coloured according to where haplotypes were found and to their relative frequency. Small black circles represent hypothetical haplotypes. Dashes on lines connecting haplotypes represent the number of nucleotide substitutions separating them. Haplotype designations and further information are given in Table 1.

Table 3. TMRCA estimates for the striped hyena and between striped hyena and the outgroups used for time calibrations in the phylogenetic dating analyses

	Hyaena hyaena	*H. hyaena / P. brunnea*	*(H. hyaena + P. brunnea) / C. crocuta*
Thomson's estimator			
648 bp dataset	418.938 ka[a] (95% CI: 225.273-672.201)		
340 bp dataset	410.256 ka[b] (95% CI: 152.578-792.834)		
GENETREE			
648 bp dataset			
Constant population size model	333.334 ka[a] (95% CI: 125.000-541.668)		
Exponential growth model	303.784 ka[a] (95% CI: 136.344-471.224)		
340 bp dataset			
Constant population size model	279.418 ka[b] (95% CI: 44.118-514.718)		
Exponential growth model	240.043 ka[b] (95% CI: 83.493-396.593)		
BEAST[c]			
648 bp dataset; calibration with two outgroups[d]			
Strict clock (estimated clock rate: mean = 1.3% median = 1.3% 95% HPD = [0.9%, 1.8%])	377.8 ka (mean: 395.9 ka) (95% HPD: 172.7-657.2)	4.598 Ma (mean: 4.599 Ma) (95% HPD: 4.015-5.164)	9.323 Ma (mean: 9.393 Ma) (95% HPD: 6.808-12.118)
Random local clock (estimated clock rate: mean = 1.3% median = 1.3% 95% HPD = [0.9%, 1.8%])	387.3 ka (mean: 422 ka) (95% HPD: 161.5-711.9)	4.600 Ma (mean: 4.602 Ma) (95% HPD: 4.010-5.163)	9.304 Ma (mean: 9.369 Ma) (95% HPD: 6.705-12.070)

	Hyaena hyaena	*H. hyaena / P. brunnea*	*(H. hyaena + P. brunnea) / C. crocuta*
648 bp dataset; calibration with one outgroup[e]			
Strict clock (estimated clock rate: mean = 1.6% median = 1.5% 95% HPD = [0.8%, 2.5%])	333.8 ka (mean: 353.1 ka) (95% HPD: 135.0-613.3)	4.605 Ma (mean: 4.605 Ma) (95% HPD: 4.028-5.199)	
340 bp dataset; calibration with two outgroups[d]			
Strict clock (estimated clock rate: mean = 1.3% median = 1.2% 95% HPD = [0.8%, 1.8%])	440.7 ka (mean: 472.1 ka) (95% HPD: 143.4-872.3)	4.567 Ma (mean: 4.569 Ma) (95% HPD: 3.973-5.136)	9.874 Ma (mean: 9.940 Ma) (95% HPD: 6.908-13.058)
340 bp dataset; calibration with one outgroup[e]			
Strict clock (estimated clock rate: mean = 1.1% median = 1.1% 95% HPD = [0.6%, 1.7%])	483.8 ka (mean: 522.3 ka) (95% HPD: 150.8-974.3)	4.603 Ma (mean: 4.604 Ma) (95% HPD: 4.010-5.183)	

[a] Estimate using a substitution rate of 1% per Ma

[b] Estimate using a substitution rate of 0.93% per Ma

[c] For each of the analyses carried out in BEAST we also present the respective estimated clock rate in % per Ma

[d] Two calibration points: *Hyaena hyaena / Parahyaena brunnea* (4.6 Ma; 95% interquantile range: 4.012-5.188 Ma) and *(H. hyaena + P. brunnea) / Crocuta crocuta* (9.5 Ma; 95% interquantile range: 5.58-13.42 Ma) (normally distributed calibration densities)

[e] One calibration point: *Hyaena hyaena / Parahyaena brunnea* (4.6 Ma; 95% interquantile range: 4.012-5.188 Ma) (normally distributed calibration density)

Supplementary Information

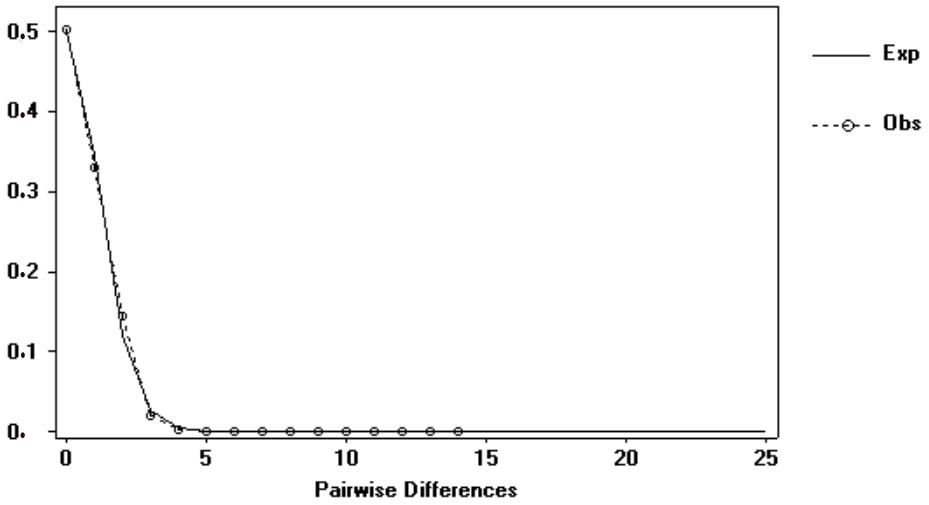
Methodological details of analyses using MCMC sampling based skyline methods in BEAST

In the BSP analysis we used four groups and performed analyses with both the piecewise-constant and piecewise-linear models, and for the substitution rate parameter we tested normal, exponential, and uniform prior distributions. The lengths of the MCMC chains were 200 million generations, sampled every 20000th after a pre-burn-in of 20 million. In the EBSP analysis we performed runs with both the stepwise and linear models, and for the substitution rate parameter we tested normal and uniform prior distributions. A uniform distribution was used for the population prior mean. The lengths of the MCMC chains were 200 million generations, sampled every 20000th after a pre-burn-in of 20 million. For the Skyride model, we used time-aware smoothing, a normal prior distribution for the substitution rate parameter, and MCMC chains with 200 million generations, sampled every 20000th after a pre-burn-in of 20 million. For Skygrid, we used 75 grid points, a cutoff value of 1 Ma, a normal prior distribution for the substitution rate parameter, and MCMC chains with 200 million generations, sampled every 20000th after a pre-burn-in of 20 million. For each of the four types of MCMC sampling-based skyline plot analyses, we used Tracer 1.7.2 (Rambaut et al. 2018), with a default burn-in of 10% of the chain length, to assess convergence of the chain to the stationary distribution, plot marginal posterior densities, and obtain estimates and ESS of parameters. To draw the skyline plots we also used Tracer, and the Python script popGraphFromCSV.py (obtained from https://github.com/benb/beast-mcmc/tree/master/doc/EBSP/scripts).

References

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst Biol 67:901–904. https://doi.org/10.1093/sysbio/syy032

a)



b)

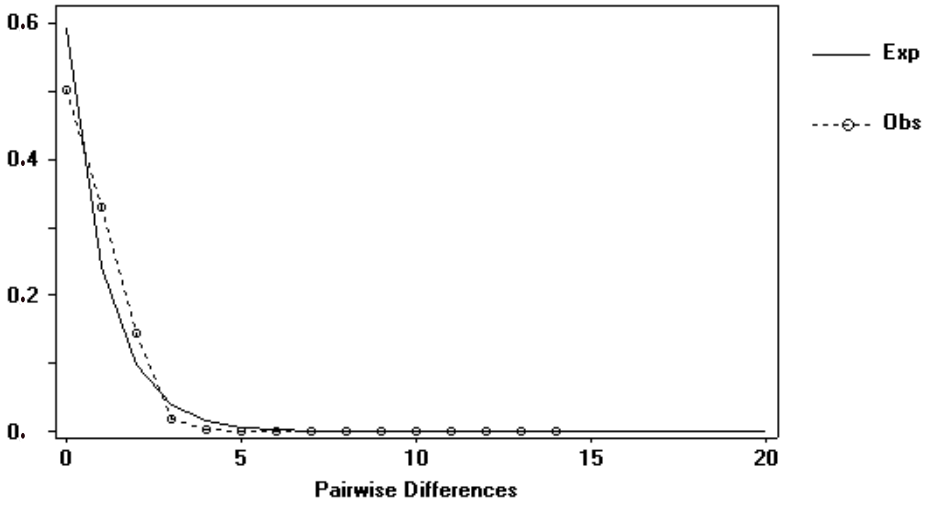
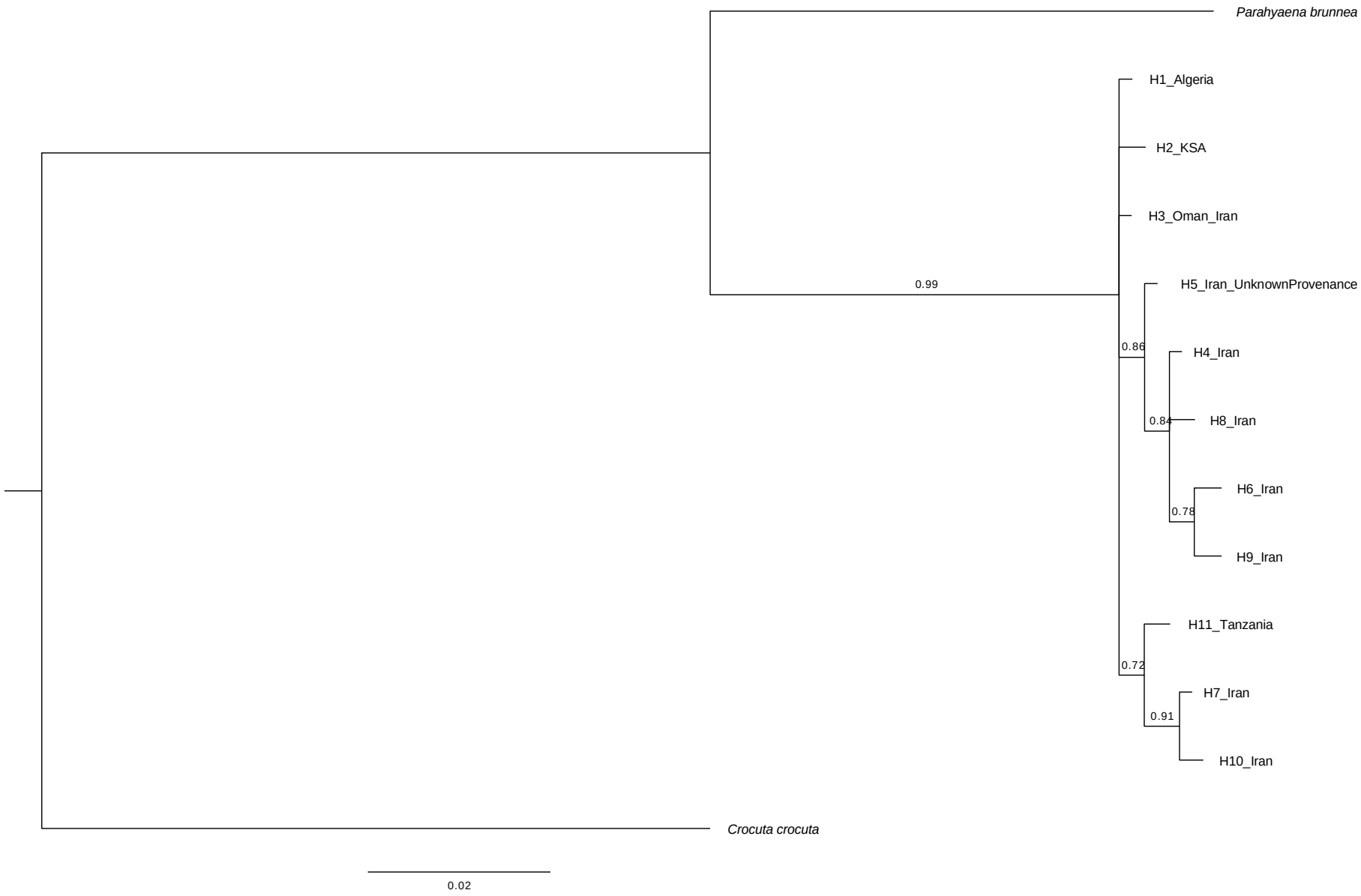


Fig. S1. DNASP graphs of mismatch distributions for the striped hyena 340 bp dataset comparing the observed distribution with (a) the expected values under a model of population growth or decline (Rogers and Harpending 1992), and (b) the expected values under a model of constant population size. The x-axis measures the number of pairwise nucleotide differences between sequences, while the y-axis measures the frequency of such pairwise differences in the dataset.

a)


Parahyaena brunnea
H1_Algeria
H2_KSA
H3_Oman_Iran
0.99
H5_Iran_UnknownProvenance
0.86
H4_Iran
0.84
H8_Iran
H6_Iran
0.78
H9_Iran
H11_Tanzania
0.72
H7_Iran
0.91
H10_Iran
Crocuta crocuta
0.02

b)

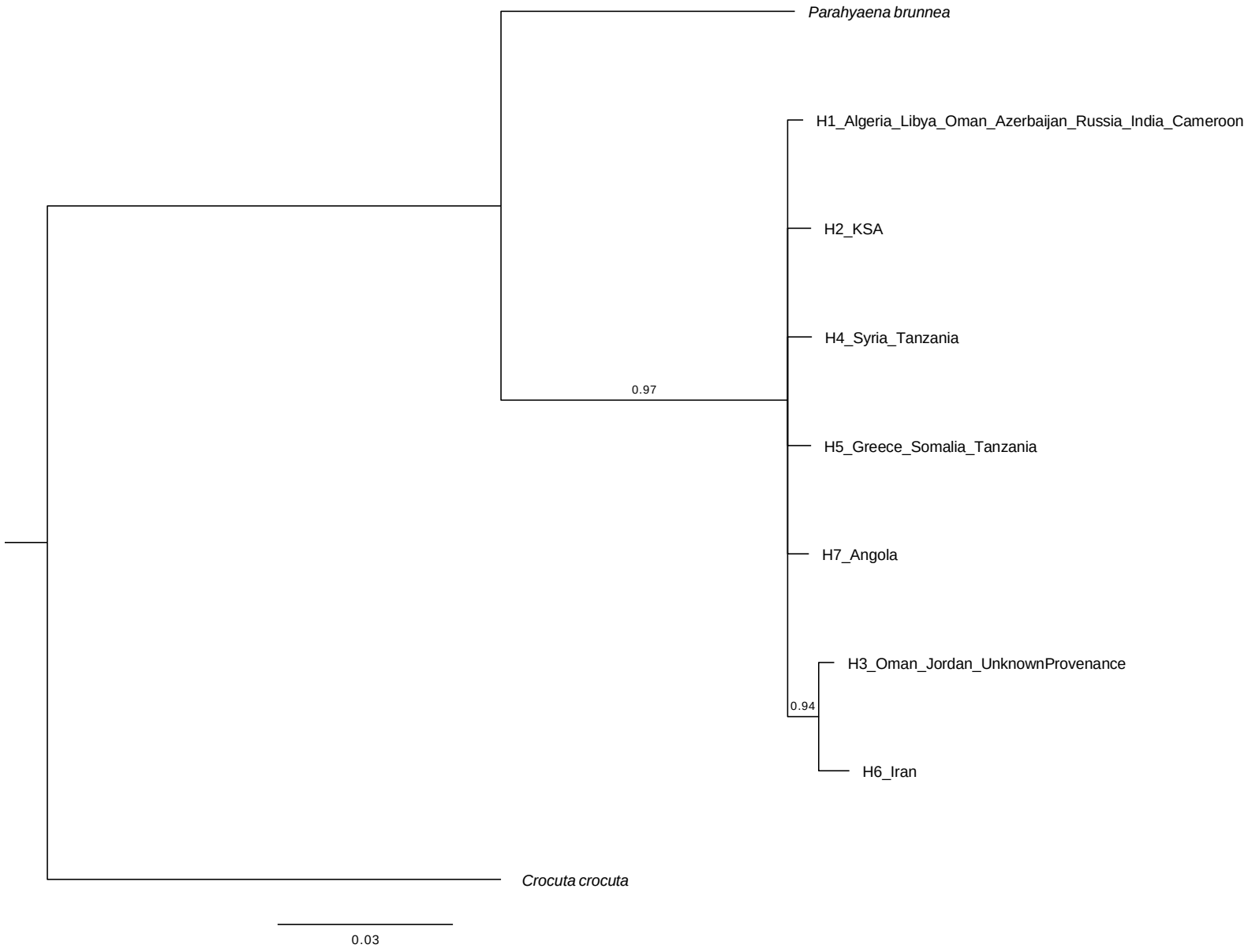

Parahyaena brunnea
H1_Algeria_Libya_Oman_Azerbaijan_Russia_India_Cameroon
H2_KSA
H4_Syria_Tanzania
0.97
H5_Greece_Somalia_Tanzania
H7_Angola
H3_Oman_Jordan_UnknownProvenance
0.94
H6_Iran
Crocuta crocuta
0.03

Fig. S2. Majority-rule consensus phylograms (cut-off 0.7) from the Bayesian inference analyses of the (a) 648-bp and (b) 340-bp *Cyt b* datasets, with two outgroups (*Parahyaena brunnea* and *Crocuta crocuta*). See Table 1 for haplotype information. Numbers above branches are Bayesian posterior probabilities.

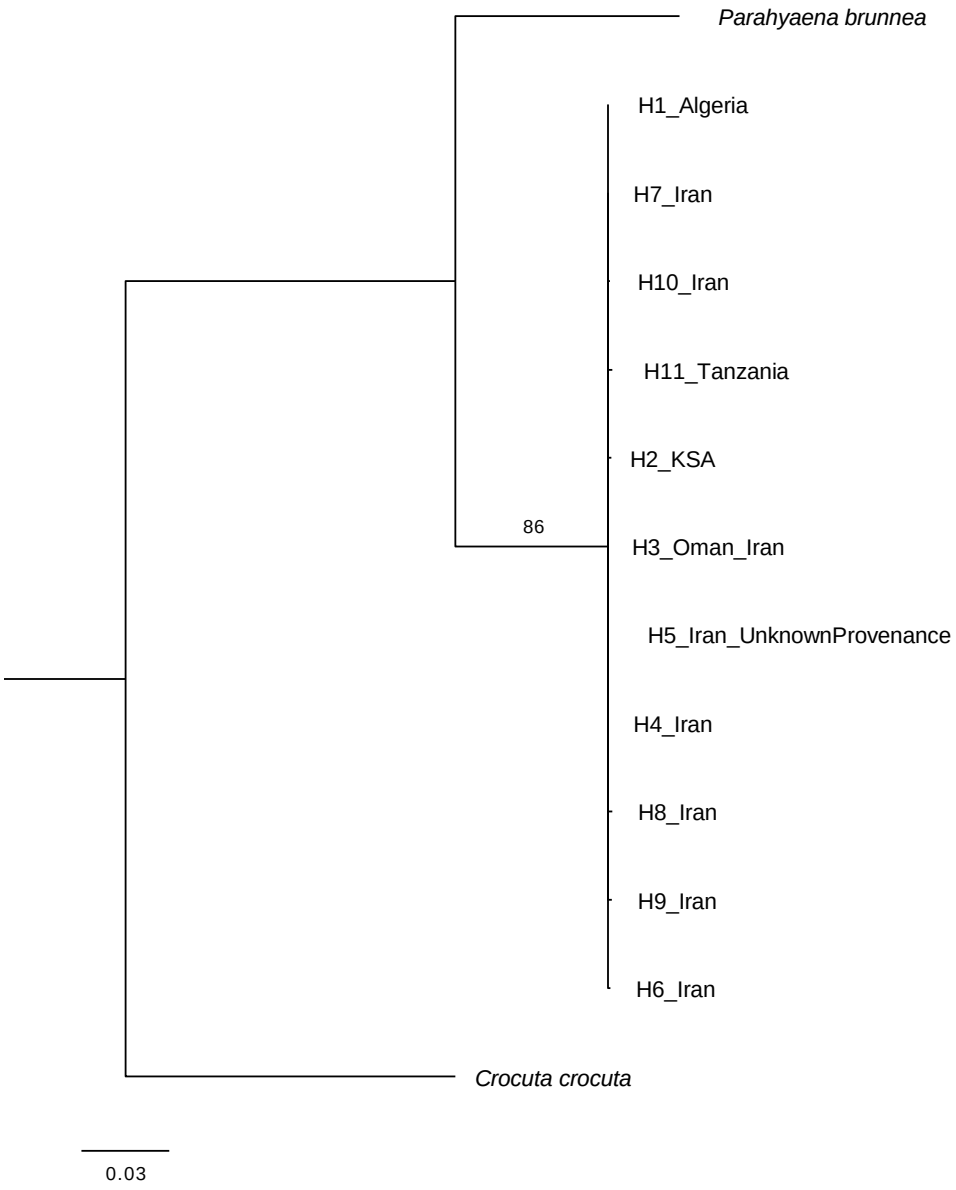


Fig. S3. Majority-rule consensus phylogram (cut-off 70%) of the maximum likelihood analysis of the 648 bp dataset with two outgroups (*Parahyaena brunnea* and *Crocuta crocuta*). See Table 1 for haplotype information. Numbers above branches are bootstrap values.

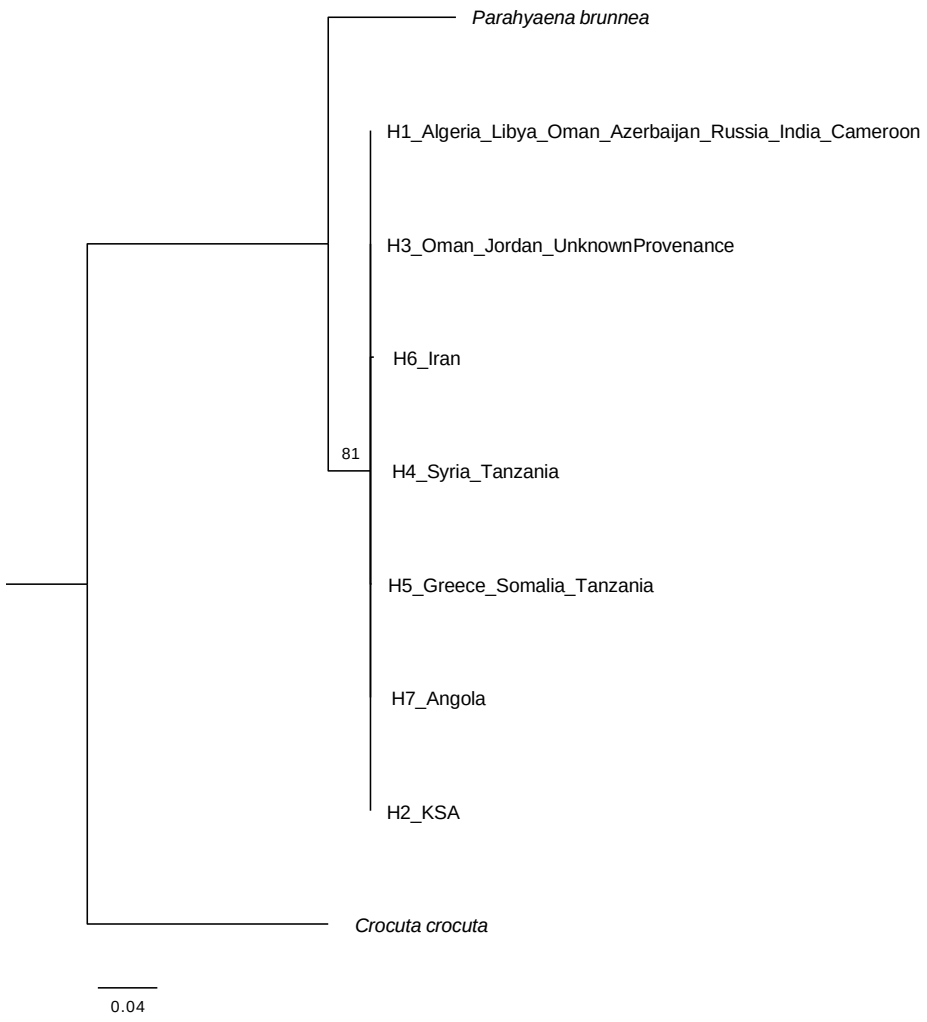
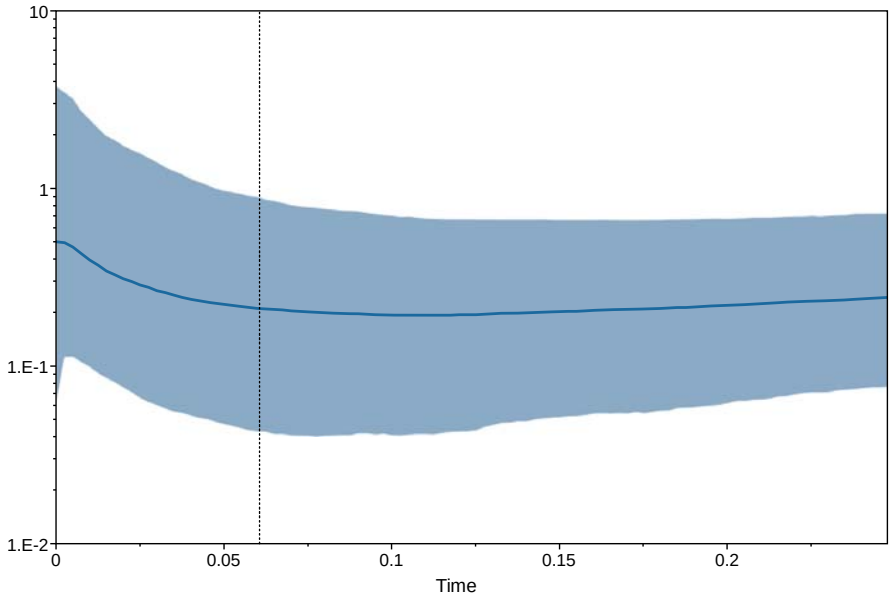


Fig. S4. Majority-rule consensus phylogram (cut-off 70%) of the maximum likelihood analysis of the 340 bp dataset with two outgroups (*Parahyaena brunnea* and *Crocuta crocuta*). See Table 1 for haplotype information. Numbers above branches are bootstrap values.

a)



b)

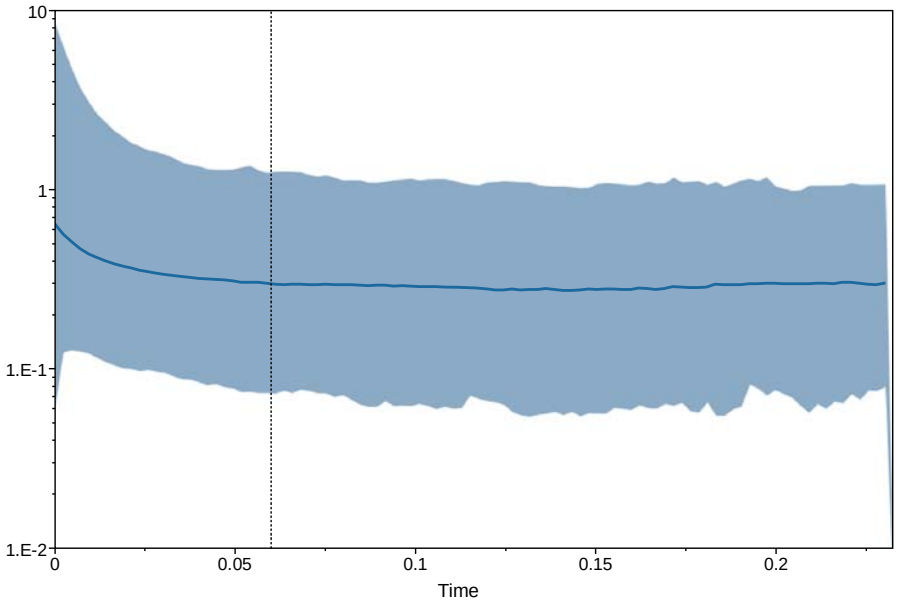


Fig. S5. BSPs for the sample of 76 striped hyena 648-bp *Cyt b* sequences, using the a) piecewise-constant and b) piecewise-linear models, and a normally distributed prior for the substitution rate with mean 0.01 and standard deviation 0.0025 (in substitutions per site per Ma). The thicker solid dark blue line is the median estimate of population size, and the lighter blue shaded region represents the 95% HPD envelope. Maximum time is the root height median, and the black vertical dotted line indicates the lower 95% HPD estimate. The x-axis (in linear scale) measures time in Ma units (going backwards in time from left to right) and the y-axis (in logarithmic scale) measures population size in units of estimated effective population size (N_e) x generation time in years x 10^{-6}.

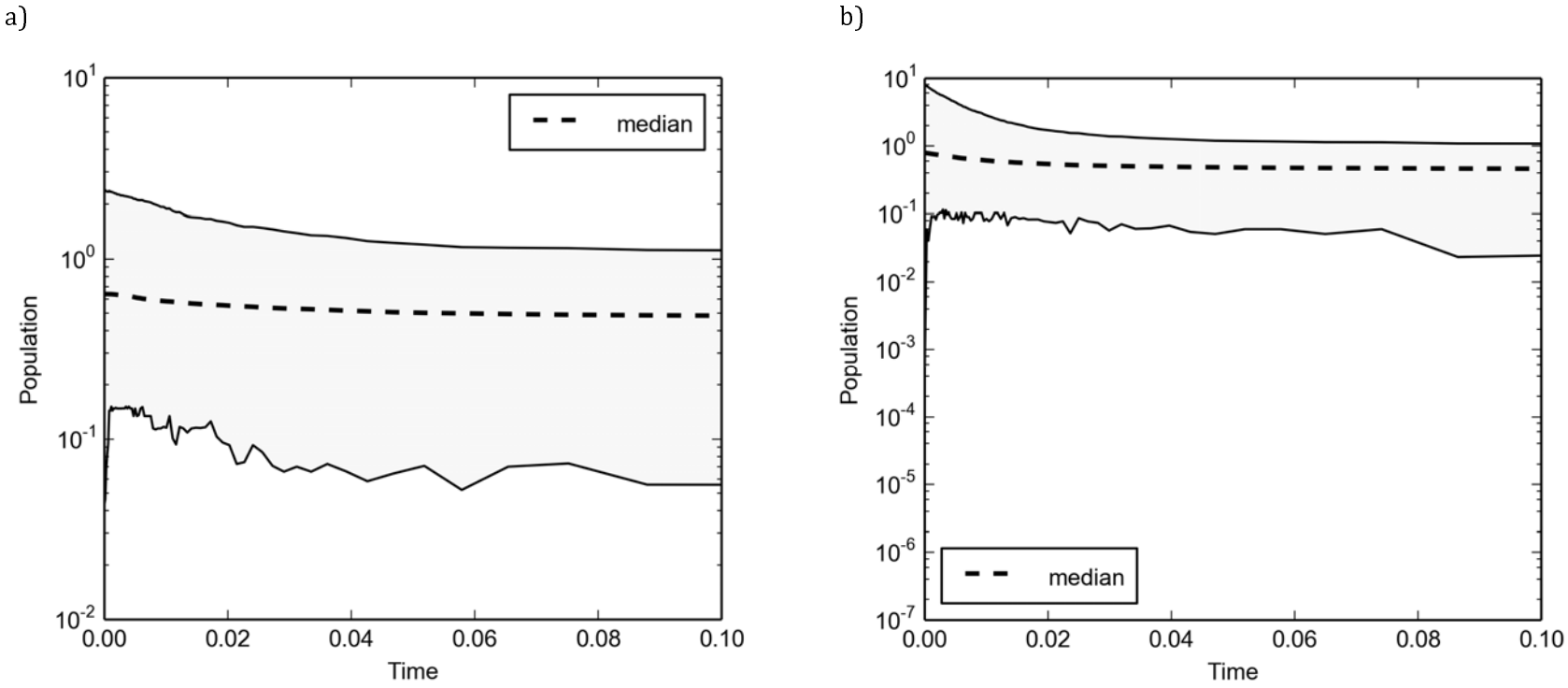


Fig. S6. EBSPs for the sample of 76 striped hyena 648-bp *Cyt b* sequences, using the a) stepwise and b) linear models, and a normally distributed prior for the substitution rate with mean 0.01 and standard deviation 0.0025 (in substitutions per site per Ma). The thicker dashed line is the median estimate of population size, and the shaded region represents the 95% HPD envelope. The x-axis (in linear scale) measures time in Ma units (going backwards in time from left to right) and the y-axis (in logarithmic scale) measures population size in units of estimated effective population size (N_e) x generation time in years x 10^{-6}..

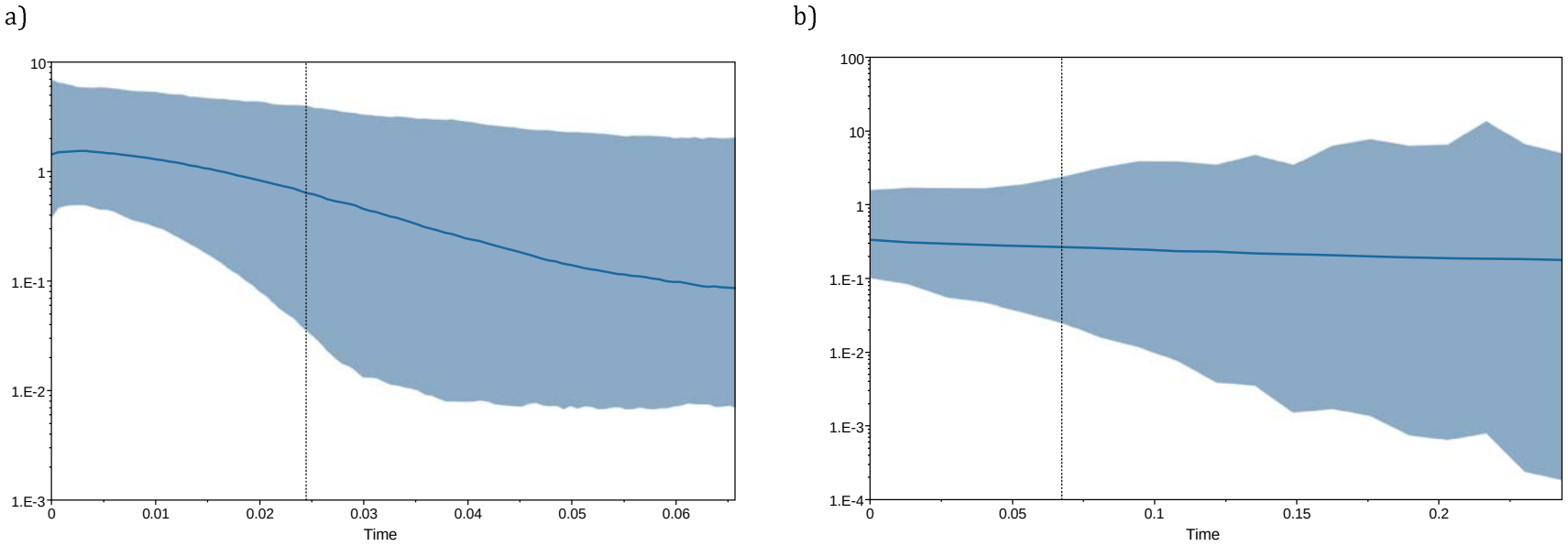


Fig. S7. Skyline plots for the sample of 76 striped hyena 648-bp *Cyt b* sequences, using the a) skyride and b) skygrid methods, and a normally distributed prior for the substitution rate with mean 0.01 and standard deviation 0.0025 (in substitutions per site per Ma). The thicker solid dark blue lines are the median estimates of population size, and the lighter blue shaded regions represent 95% HPD envelopes. Maximum time is the root height median, and the black vertical dotted lines indicate the lower 95% HPD estimates. The x-axis (in linear scale) measures time in Ma units (going backwards in time from left to right) and the y-axis (in logarithmic scale) measures population size in units of estimated effective population size (N_e) x generation time in years x 10^{-6}.

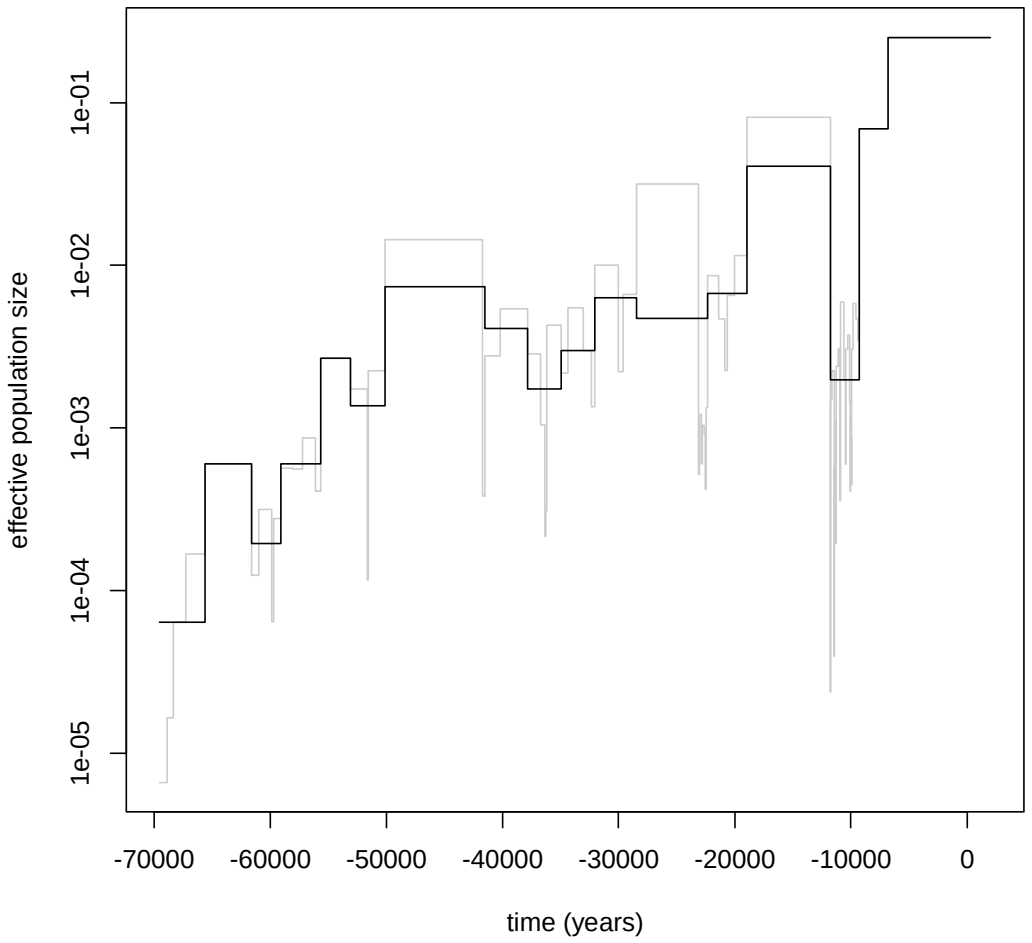


Fig. S8. Generalized skyline plot (black line), and classic skyline plot (grey line) in the background. Time is zero at the present.

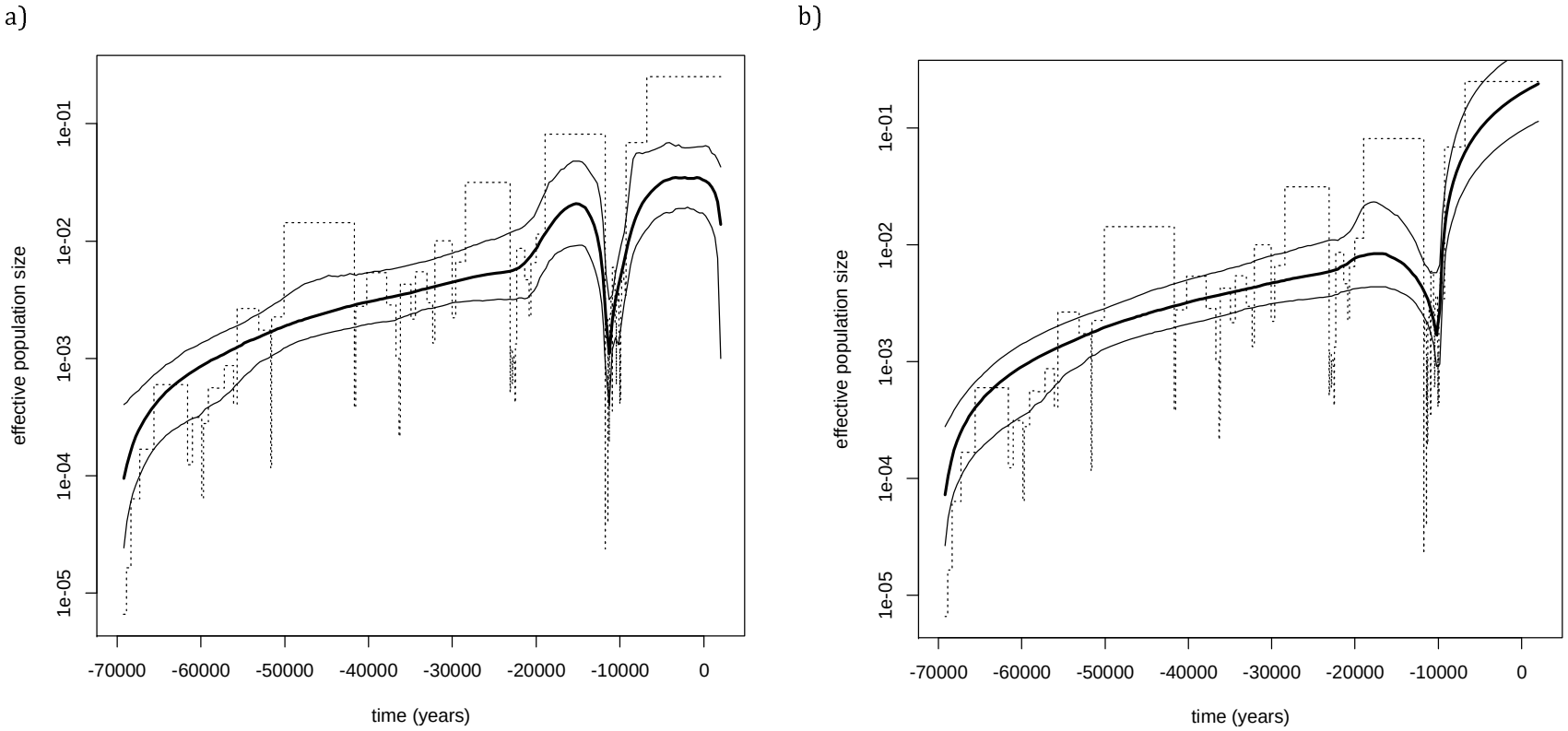


Fig. S9. Bayesian MCP plots estimated using a) a constant population size prior and b) a skyline plot prior. The thicker black lines are the median of the posterior distribution of population size, and the thinner lines represent the 95% confidence intervals. In both figures, for comparison, the classic skyline plot is shown in the background (dashed lines). Time is zero at the present.