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

Authors: Louiza Derouiche[a*], Mónica Rodrigues[b,c], Hafida Benameur-Hasnaoui[d], Ridah Hadj Aissa[e], Yasaman Hassan-Beigi[b], Seyed Massoud Madjdzadeh[f], Zuhair Amr[g], Aimee Cokayne[h], Paul Vercammen[i], Carlos Rodríguez Fernandes[b,c*]

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

[*]**Corresponding authors**

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

Abstract

The striped hyena (*Hyaena hyaena*) occurs in a wide range from north and east Africa, through southwest Asia to India, but its distribution is increasingly patchy and many of its populations are in decline due to intense human pressure. Its genetic diversity and structure, phylogeography, and evolutionary history, remain poorly understood. In this study, we investigated mitochondrial DNA variation in Algerian striped hyenas, by sequencing a fragment of the cytochrome *b* gene encompassing the one used in a pioneering study that remains the only one in which individuals from across the range of the species were analysed, to allow for comparisons. With the aim of contributing to our understanding of the evolutionary history of the species, we also examined samples from geographic regions not included in that study, and using a larger global dataset we performed a wide range of analyses of demographic history and estimation of the age of the extant mitochondrial DNA variation. The Algerian population sample was monomorphic. Overall, the global patterns of genetic diversity and the results of some demographic history analyses supported a scenario of population growth in the species, estimated to have occurred in the Late Pleistocene, but many of the analyses did not detect a significant signal of growth, most likely a result of the limited power provided by a small number of segregating sites. The estimates, from three different methods, for the time to the most recent common ancestor (TMRCA) of the mitochondrial 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: striped hyenas, 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

et al., 2006; Westbury et al., 2019, 2020), while the latter two lineages diverged from each other about 4-5 Ma (Rohland et al., 2005; Koepfli et al., 2006; Hu et al., 2021; Westbury et al., 2021). Traditionally, based on differences in body size and skull and pelage characteristics (Pocock, 1934), the following five subspecies were accepted (Rieger, 1979, 1981; Mills and Hofer, 1998; Kasparek et al., 2004): *barbara* in Northwest Africa (type locality: Oran, Algeria); *dubbah* in East Africa (type locality: Atbara, Sudan); *syriaca* in Syria and Anatolia (type locality: Antakya, Turkey); *sultana* in southern Arabia (type locality: Mt. Qara, Oman); and *hyaena* from Iran to India (type locality: Benna Mountains, Larestan, Iran). Jenks and Werdelin (1998) noted that molecular data could be very useful for assessing subspecies designations in the striped hyena. In fact, a study that sequenced a 340-bp fragment of the mitochondrial Cytochrome b (*Cyt b*) gene in 13 striped hyena samples from scattered locations throughout the historical and current distribution of the species, found low genetic diversity (four haplotypes with two variable sites) and a general lack of obvious phylogeographic structure, with three of the four haplotypes shared between Africa and Eurasia (Rohland et al., 2005). Thus, although the species exhibits marked geographic variation in size and coat colour patterns, the currently accepted view is that there is no evidence for the recognition of subspecies (Kingdon, 2001; Wozencraft, 2005).

In Algeria, the striped hyena had the same history of decline over the past few centuries as observed in many other regions of its distribution, and the species remains vulnerable and threatened by anthropogenic causes of mortality and in need of conservation efforts, but it has managed to persist throughout much of the country (Kowalski and Rzebik-Kowalska, 1991; Derouiche et al., 2020).

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

127

128 **Materials and methods**

129 *Sampling and laboratory procedures*

130 To investigate the mtDNA diversity of the striped hyena in Algeria, we collected
131 samples from 25 specimens found dead on roads or in the field. We also analysed
132 samples of striped hyenas from Jordan, Saudi Arabia, Oman, and Iran (Fig. 1 and
133 Table 1). DNA was extracted from muscle tissue and blood samples using the
134 EZNA Tissue DNA kit (Omega Bio-Tek), while that from dry hair samples was
135 extracted following a simple polymerase chain reaction (PCR) buffer-based
136 extraction protocol (Allen et al. 1998; Vigilant 1999). We amplified a fragment of
137 753 base pairs (bp) of the *Cyt b* gene with the primers L14724 (5'-
138 TGATATGAAAAACCATCGTTG-3'; Irwin et al., 1991) and H15791 (5'-
139 AATGTAGTTGTCTGGGTC-3'; Fernandes et al., 2008). For a few hair samples we
140 were unsuccessful in amplifying this fragment, which we interpreted as a result
141 of DNA degradation. Thus, for these samples we amplified a 363-bp part of the

142 fragment, containing the 340-bp region analysed by Rohland et al. (2005), in two
143 overlapping pieces, using respectively the two following primer pairs: Hyena
144 Cytb F1 (5'-CCAATGACCAACATTCGA-3')/Hyena Cytb R1 (5'-
145 TCAGCCATAGTTGACGTC-3') (for a fragment of 198 bp) and Hyena Cytb F2 (5'-
146 AACCGCCTTTTCATCAGT-3')/Hyena Cytb R2 (5'-GACGTAACCTATGAATGC-3')
147 (for a fragment of 181 bp). PCRs were carried out in volumes of 15 µl with 1x
148 PCR Buffer (NZYTech), 2 mM $MgCl_2$, 0.2 mM of each dNTP (Bioline), 0.5 µM of
149 each primer, 0.75 U of of Supreme NZYTaq DNA Polymerase (NZYTech), and 3 µl
150 of DNA extract. Thermal cycling conditions consisted of an initial denaturation at
151 95 °C for 5 min, followed by 45 cycles of 30 s at 94 °C, 30 s at 50 °C, 30 s at 72 °C,
152 and a final extension of 7 min at 72 °C. The results of the PCR amplifications were
153 visualized on 2% agarose gels to verify PCR quality, and the PCR products were
154 purified with an Exo-SAP protocol (Hanke and Wink, 1994; Werle et al., 1994)
155 and sequenced by the Sanger method on a 3730 xl DNA Analyzer (Applied
156 Biosystems) at Macrogen Inc.

157

158 *Data analyses*

159 Sequences were edited, assembled, aligned, and checked for the absence of
160 indels and stop codons using SEQUENCHER 4.7 (Gene Codes Corporation) and the
161 TRANSLATE program of the SEQUENCE MANIPULATION SUITE 2 (Stothard 2000). We
162 downloaded available striped hyena *Cyt b* sequences from GenBank and included
163 these in our alignments (Table 1). Trimming our sequences to 648 bp allowed us
164 to include some Iranian striped hyena sequences from GenBank, but in order to
165 incorporate other shorter sequences from diverse geographic origins, in

166 particular those from Rohland et al. (2005), we created a smaller alignment of
167 340 bp (Table 1). Sequence alignments were analysed with FABOX 1.5 (Villesen,
168 2007) to identify samples with identical sequences. File format conversions of
169 sequence alignments for use in different computer programs were done using
170 ALTER (Glez-Peña et al., 2010).

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

183 To visualize and examine the genealogical and geographical relationships among
184 haplotypes, we estimated haplotype networks using the integer neighbour-
185 joining method (IntNJ) in POPART 1.7 (Leigh and Bryant, 2015). The IntNJ
186 method was run with the reticulation tolerance parameter α set to zero. We used
187 outgroup rooting (Jansen et al., 2002; Cassens et al., 2003; Dubach et al., 2013) to
188 estimate the ancestral node of the ingroup. The outgroups used here, and in the

189 phylogenetic tree reconstructions described next, were the brown hyena (*P.*
190 *brunnea*) and the spotted hyena (*C. crocuta*).

191 To infer evolutionary relationships among haplotypes and assess the presence of
192 clades within *H. hyaena*, we constructed phylogenetic trees. The tree
193 reconstructions were conducted with the data partitioned by codon position, as
194 this was the best-fit partitioning scheme according to the corrected Akaike
195 information criterion (AICc) (Akaike, 1974; Sugiura, 1978; Hurvich and Tsai,
196 1989) in PARTITIONFINDER 2.1.1 (Lanfear et al., 2017) using PHYML (Guindon et al.,
197 2010). Phylogenetic analyses were performed using Bayesian inference (BI) and
198 maximum likelihood (ML) as implemented in MRBAYES 3.2.6 (Ronquist et al.,
199 2012) and IQ-TREE 1.6.12 (Nguyen et al., 2015), respectively. Analyses in
200 MRBAYES were conducted with two parallel Markov Chain Monte Carlo (MCMC)
201 runs, each with four Markov chains (one cold and three heated), default heating
202 parameter (t = 0.1), and 20 million generations. The first five million generations
203 were discarded as burn-in and, thereafter, chains were sampled every 500
204 generations. The entire general time-reversible (GTR; Lanave et al., 1984)
205 substitution model space was sampled within the analyses (Huelsenbeck et al.,
206 2004). Convergence was indicated by an average standard deviation of split
207 frequencies between parallel runs of less than 0.01. For all model parameters,
208 the effective sample size (ESS) was greater than 700 and the potential scale
209 reduction factor (PSRF) was 1.0. Support for tree nodes was determined
210 according to the values of Bayesian posterior probability (BPP) obtained in a
211 majority-rule consensus tree (Holder et al., 2008; Huggins et al., 2011). In IQ-
212 TREE we used the nucleotide substitution models estimated as best fit for each
213 codon position, according to the Bayesian information criterion (BIC; Schwarz

214 1978), by ModelFinder (Kalyaanamoorthy et al. 2017), a model-selection method
215 implemented in IQ-TREE (648 bp dataset, according to codon position: Kimura
216 two-parameter model (K2P; Kimura 1980); Hasegawa-Kishino-Yano model
217 (Hasegawa et al. 1985) with a proportion of invariable sites (HKY+I); HKY) (340
218 bp dataset, according to codon position: K2P+I; HKY; Tamura-Nei model (TN93;
219 Tamura and Nei 1993)). Tree search runs were performed using IQ-TREE default
220 settings, with the exception of the '-allnni' option which turns on a more
221 thorough Nearest Neighbour Interchange (NNI) tree search, and support for each
222 node was evaluated by 1000 non-parametric bootstrap replicates (Felsenstein,
223 1985). Majority-rule consensus trees (Berry and Gascuel, 1996; Holder et al.,
224 2008) were computed with SUMTREES 4.5.2 of the DendroPy library version 4.5.2
225 (Sukumaran and Holder, 2010) and visualized and edited with FIGTREE 1.4.4
226 (available at https://github.com/rambaut/figtree/releases).

227 To try to learn more about the recent evolutionary history of the species, we
228 performed several analyses specifically on demographic history. In these
229 analyses we used only the 648 bp dataset (76 striped hyenas; 13 segregating
230 sites), as the 340 bp dataset (50 individuals) only contains five segregating sites.
231 We constructed skyline plots (Ho and Shapiro 2011), which do not depend on a
232 prespecified parametric model of demographic history, using the following four
233 MCMC sampling-based methods: Bayesian Skyline Plot (BSP; Drummond et al.
234 2005), Extended Bayesian Skyline Plot (EBSP; Heled and Drummond, 2008),
235 Bayesian Skyride (Minin et al. 2008), and Bayesian Skygrid (Gill et al. 2013).
236 These skyline plot methods were performed in BEAST 1.10.4 ('Bayesian
237 Evolutionary Analysis Sampling Trees'; Suchard et al. 2018), using the BEAGLE
238 3.1.0 library (Ayres et al. 2019), with the data partitioned by codon position. All

239 BEAST input files were created in BEAUti, available in the BEAST package. In the
240 BSP analysis we used four groups and performed analyses with both the
241 piecewise-constant and piecewise-linear models, and for the substitution rate
242 parameter we tested normal, exponential, and uniform prior distributions. The
243 lengths of the MCMC chains were 200 million generations, sampled every
244 20000th after a pre-burn-in of 20 million. In the EBSP analysis we performed
245 runs with both the stepwise and linear models, and for the substitution rate
246 parameter we tested normal and uniform prior distributions. A uniform
247 distribution was used for the population prior mean. The lengths of the MCMC
248 chains were 200 million generations, sampled every 20000th after a pre-burn-in
249 of 20 million. For the Skyride model, we used time-aware smoothing, a normal
250 prior distribution for the substitution rate parameter, and MCMC chains with
251 200 million generations, sampled every 20000th after a pre-burn-in of 20 million.
252 For Skygrid, we used 75 grid points, a cutoff value of 1 Ma, a normal prior
253 distribution for the substitution rate parameter, and MCMC chains with 200
254 million generations, sampled every 20000th after a pre-burn-in of 20 million. For
255 each of the four types of MCMC sampling-based skyline plot analyses, we used
256 TRACER 1.7.2 (Rambaut et al. 2018), with a default burn-in of 10% of the chain
257 length, to assess convergence of the chain to the stationary distribution, plot
258 marginal posterior densities, and obtain estimates and ESS of parameters. To
259 draw the skyline plots we also used TRACER, and the Python script
260 popGraphFromCSV.py (obtained from https://github.com/benb/beast-
261 mcmc/tree/master/doc/EBSP/scripts). We compared different skyline models,
262 and also simple parametric models (such as constant population size or
263 exponential growth) (see Villanea et al. 2020), in terms of their relative fit to the

264 data, using the Bayes factor (Jeffreys 1935; Kass and Raftery 1995), the ratio of
265 the marginal likelihoods of the two models under comparison. The marginal
266 likelihoods were estimated in BEAST with path sampling (Lartillot and Philippe
267 2006) and stepping-stone sampling (Xie et al. 2011), based on 100 path steps
268 and a chain length of two million iterations. We also tested skyline plot methods
269 based on a single tree (Ho and Shapiro 2011): classical (Pybus et al. 2000),
270 generalized (Strimmer and Pybus 2001), and Bayesian multiple-change-point
271 (MCP; Opgen-Rhein et al. 2005). The required ultrametric binary tree was a
272 maximum clade credibility tree (MCCT) from an analysis in MRBAYES. MCMC in
273 MRBAYES was performed as described above. The analysis used a strict clock
274 model with a uniform prior probability distribution on branch lengths, a
275 normally distributed clock rate prior with mean 0.01 and standard deviation
276 0.0025 (in substitutions per site per Ma) (how we obtained this substitution rate
277 estimate is explained below), and a truncated normal prior probability
278 distribution on the tree age with a mean of 0.4 Ma and a standard deviation of
279 0.15 (the choice of these values reflects the estimates obtained for the age of the
280 mtDNA diversity, presented below). The single-tree skyline plot analyses were
281 performed in the R package APE version 5.7.1 (Paradis and Schliep 2019).
282 Bayesian MCP analyses were performed assuming both a constant population
283 size prior demographic function (Felsenstein 1992) and a skyline plot prior
284 demographic function (Opgen-Rhein et al. 2005). We also estimated the fit of
285 different demographic models (constant population size, exponential growth,
286 expansion growth, logistic growth, with both continuous and piecewise variants)
287 (Pybus and Rambaut 2002) to the MCCT in the R package genieR version 0.1.0
288 (Xiang et al. 2019). We also used the Monte Carlo likelihood approach of Weiss

289 and von Haeseler (1998), as implemented in IPHULA 1.16 (Schmidt et al. 2007),
290 for inference on demographic history. In this analysis we used the HKY model,
291 which was selected as best-fitting the 648 bp alignment by BIC in ModelFinder
292 and also in JMODELTEST 2.1.10 (Darriba et al. 2012), and a transition/transversion
293 rate ratio (kappa, κ) of 28 estimated in TREE-PUZZLE 5.3.rc16 (Schmidt et al.
294 2002). The method of Weiss and von Haeseler (1998) involves a model with
295 three parameters (theta, θ: mutation-scaled effective population size (N_e) in the
296 past ('initial'), i.e. before an eventual N_e change; tau, τ: mutation-scaled time
297 since N_e started to change exponentially; rho, ρ: ratio of current to initial N_e),
298 which are estimated based on the number of segregating sites and the mean
299 number of pairwise nucleotide differences (k). The method approximates the
300 likelihood surface for the three parameters by coalescent simulations on a three-
301 dimensional grid. After preliminary simulations to explore the likelihood surface
302 and find the parameter ranges that enclose the global maximum likelihood peak,
303 we ran 10 replicate runs, zoomed into grid areas of high likelihoods, for
304 parameter estimation. In these final runs we defined five grid points, equally
305 spaced between 1 and 20, for ρ, 10 grid points, equally spaced between 0.1 and 5,
306 for θ, and 10 grid points, equally spaced between 0.1 and 2, for τ. We ran 50,000
307 simulations for each grid point (i.e. parameter combination), and the maximum
308 allowed difference between simulated and observed k was set to less than 0.03
309 (i.e. ± 1% around the observed mean k of 1.72). Inference on demographic
310 history parameters was also performed with the MCMC coalescent genealogy
311 sampler LAMARC 2.1.10 (Kuhner 2006; Kuhner and Smith 2007), assuming
312 either an exponential growth or shrinkage model or a simpler model of constant
313 population size. We used the Bayesian version of LAMARC, with a default

314 starting value of θ per site of 0.01 and a logarithmic prior between 10^{-5} and 10;
315 in the growth model analysis, we specified for the growth rate parameter (*g*) a
316 starting value of 1 and a linear prior bounded between -500 and 5000. Percentile
317 profile likelihoods were computed for both θ per site and *g*. The HKY model is
318 not implemented in LAMARC, so we used the closest model available,
319 Felsenstein84 (F84; Felsenstein and Churchill 1996), which in any case is very
320 closely related to the HKY model. We ran four replicates of two final chains, in
321 which for each chain we used an initial burn-in period of 200,000 discarded
322 genealogies, followed by two million genealogies sampled every 50 for
323 parameter estimation. We used adaptive heating, with four tree search
324 temperatures (initial values: 1, 1.2, 1.5, 2) in each chain, and tree swapping
325 between temperatures attempted at every step of the chain. TRACER was used to
326 analyse the resulting trace files, assess run convergence, and calculate summary
327 statistics and ESS values of parameters. We also tested constant population size
328 and exponential growth models in GENETREE 9.01, which implements a Markov
329 chain technique for ancestral inference that simulates gene trees conditional on
330 the mutation pattern in a sample of DNA sequences (Griffiths and Tavaré
331 1994a,b,c, 1995). GENETREE requires data compatible with the infinite sites model
332 (Kimura 1969; Griffiths and Tavaré 1995). To meet this requirement, in the 648
333 bp dataset we needed to remove one polymorphic site (site 546) inconsistent
334 with the model, which caused the haplotypes H1 and H3, as well as H7 and H10,
335 to no longer differ from each other. In the constant population model, after
336 preliminary runs with different values of the generating parameter for θ to
337 optimize settings and identify the region of maximum likelihood of θ, we
338 generated likelihood surfaces for θ based on 10 million simulations and 1000

339 points for the empirical distribution. In the exponential growth model, after
340 preliminary runs to identify the likelihood surface regions of θ and growth rate
341 yielding the highest joint likelihood, as the two parameters are not independent
342 of each other, we generated likelihood surfaces separately for θ and growth rate,
343 in both cases based on 100,000 simulations and 1000 points for the empirical
344 distribution. A likelihood ratio test (LRT) was used to determine whether there
345 was a significant difference in the fit to the data between the simpler constant
346 population size model and the exponential growth model.

347 To estimate the time to the most recent common ancestor (TMRCA) of the
348 observed mitochondrial genetic variation, we obtained an estimate of the
349 substitution rate for the analysed *Cyt b* fragment as follows. Given our observed
350 *Cyt b* divergence between striped hyena and its closest relative, the brown hyena,
351 and the divergence between the two lineages that the fossil record suggests may
352 have occurred 4-5.2 Ma ago (Werdelin and Solounias 1991; Werdelin et al. 1994;
353 Werdelin 2003; Werdelin and Manthi 2012; Westbury et al. 2020, 2021), an
354 estimate of the substitution rate can be obtained using equation $d_{xy} = 2\mu T$ (Nei
355 1985, 1987), where d_{xy} is the mean sequence divergence per site, μ is the mean
356 substitution rate per site, and T is the time since divergence. Genetic estimates of
357 divergence time between *Hyaena* and *Parahyaena* also fall at about 4-5 Ma
358 (Rohland et al., 2005; Koepfli et al., 2006; Hu et al., 2021; Westbury et al., 2021).
359 We confirmed the assumption of nucleotide substitution rate homogeneity
360 among the striped hyena haplotypes and between these and *Parahyaena*, using
361 *Crocuta* as outgroup, via generalized relative rate tests with topological
362 weighting (Robinson et al. 1998) in RRTREE 1.1.11 (Robinson-Rechavi and
363 Huchon 2000), and with Tajima's relative rate test (1993) as implemented in

364 MEGA 11.0.9 (Tamura et al. 2021). In RRTREE, statistical significance ($P < 0.05$) of
365 rate heterogeneity among lineages was assessed using both synonymous
366 substitution rates (*Ks*) and nonsynonymous substitution rates (*Ka*). The d_{xy}
367 between striped hyenas and *Parahyaena* was computed in MEGA using the
368 substitution model that best fits the data, according to the BIC in both IQ-TREE
369 and MEGA, among the models available in MEGA for distance estimation (TN93 for
370 the 648 bp alignment; K2P for the 340 bp alignment), and the standard error
371 was estimated using 1000 bootstrap replicates. As a point value for the
372 divergence time between *Hyaena* and *Parahyaena* we used 4.6 Ma, the midpoint
373 in the 4-5.2 Ma interval suggested by the fossil record. Similar point values have
374 been used or estimated for the divergence of the two lineages in recent genetic
375 studies of hyaenids (Westbury et al. 2020; Hu et al. 2021; Westbury et al. 2021).
376 The d_{xy} estimate for the 648 bp alignment was 0.091 ± 0.013, and for the 340 bp
377 alignment was 0.086 ± 0.017. Approximating 95% confidence intervals (CIs)
378 based on 1.96 standard errors yields 95% CIs of [0.066, 0.116] and [0.053,
379 0.119], respectively. Using the above equation $d_{xy} = 2\mu T$ with T = 4.6 Ma, the d_{xy}
380 estimates yield substitution rate estimates of 0.99% per Ma (95% CI: 0.72-
381 1.26%) and 0.93% per Ma (95% CI: 0.58-1.29%), respectively; i.e. ≈ 1% per Ma,
382 the traditional average rate for mtDNA (Brown et al. 1979; Wilson et al. 1985).
383 Also integrating the uncertainty in the fossil record about the divergence time
384 between *Hyaena* and *Parahyaena* results, respectively, in the following CIs: 0.63-
385 1.45% and 0.51-1.49%.

386 We estimated the TMRCA of the extant striped hyena mitochondrial genetic
387 variation, based on both 648-bp and 340-bp datasets and using different
388 methods implemented in Thomson's estimator (Thomson et al 2000), GENETREE

389 and BEAST. An attractive feature of Thomson's estimator is that it does not rely
390 on a specific population genetic model (Thomson et al 2000). On the other hand,
391 the estimator requires data consistent with the infinite sites model. Accordingly,
392 as mentioned above for the GENETREE analysis, in the 648 bp dataset we removed
393 a polymorphic site (site 546). In the 340 bp dataset it was also necessary to
394 remove one polymorphic site incompatible with the said model (site 317), which
395 caused the haplotypes H1 and H7, as well as H4 and H5, to no longer differ from
396 each other. We computed Thomson's estimator and its variance (the latter
397 according to equation 4 in Hudson 2007) using NumPy code kindly provided by
398 Dr. Helmut Simon (https://github.com/helmutsimon), and 95% CIs (as in
399 equation 5 in Hudson 2007). The required estimate of the unfolded site
400 frequency spectrum was obtained using the site.spectrum function of the R
401 package PEGAS (Paradis 2010) and an estimate of the ancestral sequence of the
402 ingroup. We estimated this ancestral sequence using the two outgroups (*P.*
403 *brunnea* and *C. crocuta*) in two different computer programs, FASTML
404 (http://fastml.tau.ac.il/; Ashkenazy et al 2012) and IQ-TREE, for comparison
405 purposes, and both provided the same estimate of the ingroup's ancestral
406 sequence for each of the two datasets. In GENETREE, we used both the constant
407 population model and the exponential growth model, with the ML estimates of θ
408 and growth rate (obtained as described above) as generating parameters. Both
409 the 648-bp and 340-bp datasets were made compatible with the infinite sites
410 model by removing one inconsistent site in each dataset, as described above. The
411 ingroup's ancestral sequences used, one for each of the two datasets, were those
412 mentioned above, estimated using FASTML and IQ-TREE. We ran 100 million
413 simulations to obtain TMRCA estimates and likelihood surfaces with 10,000

414 points. In BEAST we ran analyses with both outgroups and only with *P. brunnea*.
415 The outgroups correspond to the molecular clock calibrations used: i) the
416 divergence between the *Crocuta* lineage and the lineage ancestral to *Hyaena* and
417 *Parahyaena*, for which we used a normally distributed calibration prior with
418 mean 9.5 Ma and standard deviation 2 Ma (Jenks and Werdelin 1998; Rohland et
419 al. 2005; Westbury et al. 2019, 2020; Hu et al. 2021); and ii) the split between
420 *Hyaena* and *Parahyaena*, with a normal calibration prior with mean 4.6 Ma and
421 standard deviation 0.3 Ma (Rohland et al. 2005; Koepfli et al. 2006; Hu et al.
422 2021; Westbury et al. 2020, 2021). The datasets were partitioned by codon
423 position. We used the piecewise-constant skyline tree prior (Drummond et al.
424 2005), which has been shown to perform well in molecular dating analyses
425 involving a mixture of inter- and intraspecific data (Ritchie et al. 2017; Mello et al.
426 2021), and the approximate continuous time Markov chain (CTMC) rate
427 reference prior (Ferreira and Suchard 2008). In the analyses of the 648 bp
428 dataset with two outgroups, we used relaxed clock models (uncorrelated
429 lognormal model, UCLN, Drummond et al. 2006; random local clocks model, RLC,
430 Drummond and Suchard 2010) because a LRT (performed in MEGA) of the strict
431 molecular clock hypothesis rejected it ($P = 1.8 \times 10^{-5}$). In contrast, this hypothesis
432 was not rejected ($P = 0.06$) for the dataset with only *P. brunnea* as outgroup, in
433 agreement with the results from RRTREE and Tajima's relative rate test, and
434 therefore we assumed a strict molecular clock in the analysis of this dataset.
435 Similarly, the strict clock hypothesis was rejected for the 340 bp dataset with
436 two outgroups ($P = 0.007$), but not for the one with only *P. brunnea* as outgroup
437 ($P = 0.226$). For each BEAST analysis, we ran four independent replicate MCMC
438 simulations, each with 50 million generations and sampled every 5000

439 generations following a pre-burnin of 10%. We also ran two replicate
440 simulations without data to obtain estimates from the prior distribution and test
441 the influence of priors on posterior distributions (Drummond et al., 2006; Brown
442 and Smith, 2018). For each analysis, we used TRACER with the default burn-in to
443 assess convergence of the chain to the stationary distribution, obtain estimates
444 and ESS of parameters, and plot marginal posterior densities. After verifying
445 convergence, and confirming that posterior distributions of estimates were
446 markedly different from the prior distributions, the tree files from the four runs
447 with data were combined using SUMTREES, with the first 25% (2500) trees from
448 each run discarded as burn-in, into a MCCT with median node ages. This method
449 has a good overall performance, in terms of accuracy in estimating ages and
450 model fit, compared to other tree summary approaches (Morrison 2008;
451 Battistuzzi et al. 2010; Heled and Bouckaert, 2013; Bromham et al., 2018). After
452 analysing the results of the runs with the relaxed clock models, we noticed that
453 in general the TMRCA estimates for the striped hyena mitochondrial variation
454 seemed too high (> 800 ka) when compared with those obtained with Thomson's
455 estimator (range of mean estimates: ~ 410-420 ka) and with GENETREE (range of
456 mean estimates: ~ 240-330 ka) and the only previously published estimate that
457 we are aware of (340 ka; Rohland et al 2005). Thus, we considered these high
458 estimates likely to be erroneous. The apparent poor performance of the UCLN
459 clock model does not seem to be due to the presence of significant rate
460 autocorrelation among lineages, since, for both datasets with two outgroups, the
461 95% highest posterior density (HPD) interval of the covariance between parent
462 and child branch rates contained zero (Drummond et al., 2006). The CorrTest
463 method (Tao et al 2019), implemented in MEGA, also did not reject the hypothesis

464 of independent evolutionary rates among lineages ($P > 0.05$). However, we also
465 noted that the standard deviation of the UCLN clock (ucld.stdev parameter) was
466 low in the analyses of both datasets (648 bp dataset: mean = 0.049, median =
467 0.010, 95% HPD interval = [3.7 x 10^{-7}, 0.243]; 340 bp dataset: mean = 0.138,
468 median = 0.055, 95% HPD interval = [9.9 x 10^{-6}, 0.557]), which suggests that the
469 data is quite clock-like. The same indication was given by the results for the
470 posterior distribution of the number of rate changes (K) (rateChangeCount
471 parameter) in the RLC analyses, as a greater posterior than prior probability for
472 $K = 0$ supports the global clock hypothesis, while small or negligible posterior
473 probability for $K = 0$ strongly rejects the hypothesis (Drummond and Suchard
474 2010). In the analysis with the 648 bp dataset, the posterior distribution for K
475 (mean: 0.619; median and mode: 0; 95 % HPD interval: [0,2]) showed a higher
476 probability for $K = 0$ than the prior distribution. This was not so with the 340 bp
477 dataset (mean: 0.769; median and mode: 1; 95 % HPD interval: [0,2]), for which
478 43% of the probability mass of the posterior distribution was for $K = 1$, but yet
479 42% of the probability was for $K = 0$. In such cases, a strict clock (SC) may
480 perform better in dating estimation (Ho et al 2005; Brown and Yang 2011).
481 Marginal likelihood estimates, obtained using stepping-stone sampling in BEAST
482 with 100 path steps and a chain length of 500,000, also supported the use of a SC
483 (648 bp dataset: log marginal likelihood = -1344.5; 340 bp dataset: log marginal
484 likelihood = -711.2), compared with the RLC (648 bp dataset: log marginal
485 likelihood = -1344.5; 340 bp dataset: log marginal likelihood = -711.1) and the
486 UCLN (648 bp dataset: log marginal likelihood = -1347.4; 340 bp dataset: log
487 marginal likelihood = -713.2). Therefore, we decided to also use the SC in the
488 analyses with two outgroups.

489

490 **Results**

491 *Genetic variation, haplotype relationships and phylogenetic analyses*

492 The new striped hyena DNA sequences discovered in this study were deposited
493 in GenBank under accession numbers PP328974-PP328975 (Table 1). Several
494 observations indicate that the generated sequences are mitochondrial and not
495 nuclear-integrated copies of mtDNA (Arctander 1995; Zhang and Hewitt, 1996;
496 Bensasson et al., 2001). First, PCRs consistently yielded single products of the
497 expected size. Second, independent replicate PCRs produced identical sequences.
498 Third, the sequences were unambiguous and did not contain indels or stop
499 codons (Smith et al. 1992; Mirol et al. 2000; Triant and DeWoody 2007).

500 Genetic diversity measures and results of mutation-drift equilibrium tests, for
501 both the 648 bp and 340 bp alignments of striped hyenas, are presented in Table
502 2. In the longer alignment the number of polymorphic sites was 13, with nine
503 being parsimony-informative and four singletons, and in the shorter alignment
504 the number of polymorphic sites was five, four of them parsimony-informative
505 and one a singleton. In both datasets, none of the mutation-drift equilibrium
506 tests used rejected at the 95% confidence level the null hypothesis of a constant
507 size population under the neutral model; note that in the F_S statistic test, an α =
508 0.02 corresponds to a 5% significance level (Fu 1997). The D* and F* statistics
509 were respectively -0.739 (P = 0.233) and -0.983 (P = 0.180) for the 648 bp
510 alignment, and 0.134 (P = 0.529) and -0.228 (P = 0.396) for the 340 bp alignment.
511 Analysis of the mismatch distribution for the 648 bp dataset under a model of
512 sudden demographic expansion (Schneider and Excoffier 1999), implemented in

513 ARLEQUIN, did not reject the expansion hypothesis (P = 0.678; 20,000 bootstrap
514 replicates). In agreement, the Harpending's (1994) raggedness index was low
515 (0.032; P = 0.552). The model parameter for the expansion time was estimated at
516 1.877 (in mutational time units) (95% CI: 0.715-3.736), which, assuming a
517 nucleotide substitution rate of 1% per Ma, translates into 144,830 years (95%
518 CI: 55,170-288,272). For the 340 bp dataset, the analysis of the mismatch
519 distribution under the sudden expansion model failed to converge, possibly due
520 to the low number of polymorphic sites in this dataset. However, analyses in
521 DNASP showed that the observed mismatch distribution fits better with the
522 expected distribution under a sudden expansion model (Rogers and Harpending
523 1992) than under a model of constant population size (Fig. S1). We also
524 attempted in ARLEQUIN to test the mismatch distribution under a model of spatial
525 expansion (Ray et al 2003; Excoffier 2004), but the least-square procedure to fit
526 the expected and observed mismatch distributions did not converge.

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

531

532 *Demographic simulations*

533 The different MCMC sampling-based skyline plot analyses conducted in BEAST
534 did not identify signs of significant population size change in the Late Quaternary
535 demographic history of the striped hyena. In the BSP analyses using a normally
536 distributed prior for the substitution rate with mean 0.01 and standard deviation

537 0.0025 (in substitutions per site per Ma), truncated to be positive
538 (corresponding to a 95% interquantile range of 0.005-0.015), the estimated
539 mean tree root height was 281 ka (median: 248 ka; 95% HPD interval: 61-571
540 ka) for the piecewise-constant model (ESS > 1500 for all parameters), and 264
541 ka (median: 232 ka; 95% HPD interval: 60-536 ka) for the piecewise-linear
542 model (ESS > 2200 for all parameters). The respective BSPs for the two models
543 are shown in Fig. S5, in which it can be observed that the rise of the population
544 size curve at the left end of the plots never exceeds the 95% HPD limits along the
545 rest of the graph (Ho and Shapiro 2011; Grant 2015), where the curve is
546 essentially flat. The results of analyses using either a uniformly or exponentially
547 distributed prior for the substitution rate exhibited a similar pattern of no
548 significant population size changes (not shown). In the EBSP analyses using the
549 same normally distributed prior for the substitution rate as above, the estimated
550 mean tree root height was 363 ka (median: 331 ka; 95% HPD interval: 85-717
551 ka) for the stepwise model (ESS > 2200 for all parameters), and 341 ka (median:
552 304 ka; 95% HPD interval: 58-688 ka) for the linear model (ESS > 2800 for all
553 parameters). For the 'population size changes' parameter in each model, the
554 estimated mean was 0.867 (median and mode were 1, and the 95% HPD interval
555 was 0-2) and 1.024 (median and mode were 1, and the 95% HPD interval was 0-
556 3), respectively. The respective EBSPs for the two models are presented in Fig.
557 S6. Results of analyses using a uniformly distributed prior for the substitution
558 rate also did not indicate significant population size changes in recent
559 demographic history (not shown). The skyride (ESS > 6500 for all parameters)
560 and skygrid (ESS > 650 for all parameters) analyses also yielded the same result
561 (Fig. S7); their respective mean tree root height estimates were 72 ka (median:

562 66 ka; 95% HPD interval: 24-133 ka) and 277 ka (median: 245 ka; 95% HPD
563 interval: 67-564 ka). In the analyses with the constant population size (uniform
564 population size prior; ESS > 4700 for all parameters) and exponential growth
565 (uniform population size and growth rate priors; ESS > 1500 for all parameters)
566 models, the estimated mean tree root heights were 419 ka (median: 349 ka; 95%
567 HPD interval: 111-809 ka) and 508 ka (median: 349 ka; 95% HPD interval: 103-
568 814 ka). The 95% HPD interval of the posterior distribution of the exponential
569 growth rate parameter ([-1.803, 4.018]) included zero, i.e. did not reject a
570 constant population size. The skygrid model was the one with the highest log
571 marginal likelihood estimate (stepping-stone sampling MLE = -1096; path
572 sampling MLE = -1095), but not significantly different from that of the piecewise-
573 linear version of the BSP (MLE = -1097 by both stepping-stone sampling and
574 path sampling). Still using Bayes factors, the skygrid model had positive support
575 against the piecewise-constant version of the BSP (stepping-stone sampling MLE
576 = -1099; path sampling MLE = -1098) and very strong support (2 ln Bayes factor
577 > 10) against other models. The results of the single-tree skyline plot analyses
578 done in APE suggested population growth during the Late Pleistocene (Figs. S8-
579 S9). The fit of the different demographic models tested in genieR was compared
580 using the Akaike information criterion (AIC). The constant population size model
581 was rejected (ΔAIC = 90), while the exponential growth (AIC = -1575.919) and
582 expansion growth (AIC = -1573.919) models were equally supported (ΔAIC ≤ 2).

583 The implementation of Weiss and Von Haeseler's (1998) method in IPHULA
584 estimated a maximum likelihood model parameter combination (ρ = 10, θ = 2.8,
585 τ = 0.3) that suggests slow exponential growth since the mid-Holocene.
586 Specifically, assuming a nucleotide substitution rate of 1% per Ma and a

587 generation time for the striped hyena of 9.2 years (Pacifici et al 2013), a τ of 0.3
588 translates to about 5 ka. In turn, a θ of 2.8 corresponds to a female N_e of 23484 at
589 that time. However, the 95% confidence set of models, based on the likelihood
590 ratio of each combination of demographic parameters tested and the maximum
591 likelihood model, included scenarios of constant population size.

592 In LAMARC, analyses using the exponential growth or shrinkage model
593 suggested a slow exponential growth scenario. The highest likelihood estimate of
594 *g* was 910 (95% HPD interval: [-83, 4704]; ESS > 700 in all individual replicate
595 runs). Therefore, the 95% HPD interval crossed zero, not ruling out scenarios of
596 constant population size or even contraction. Moreover, LAMARC estimates of *g*
597 tend to be biased upwards when only a few loci are analysed (Kuhner et al.
598 1998). The highest likelihood estimate of current θ per site was 0.008 (95% HPD
599 interval: [0.002, 0.020]; ESS > 350 in all individual replicate runs), which
600 corresponds to a θ of 5.2 for the 648 bp fragment and, again assuming a
601 nucleotide substitution rate of 1% per Ma and a generation time of 9.2 years, can
602 be translated into a female N_e (N_ef) of 43478. However, LAMARC estimates of θ
603 in the varying population size model may also be slightly biased upward, due to
604 correlation with *g*, when only a few loci are analysed (Kuhner et al. 1998). In the
605 analyses with the constant population size model, the highest likelihood estimate
606 of θ per site was 0.004 (95% HPD interval: [0.002, 0.006]; ESS > 850 in all
607 individual replicate runs), which corresponds to a θ of 2.6 for the 648 bp
608 fragment and, assuming a nucleotide substitution rate of 1% per Ma and a
609 generation time of 9.2 years, can be translated into a N_ef of 21739. This θ
610 estimate was very similar to that obtained using Watterson's (1975) estimator
611 based on the number of segregating sites: 2.7.

612 In GENETREE, the analyses using the constant population size model also yielded a
613 ML θ estimate of 2.7 (thus, a $N_e f$ of 22645), while in the exponential growth
614 model the ML θ estimate was 3.1 (i.e. $N_e f$ = 26000); the latter model suggested a
615 slow exponential growth rate. However, a LRT did not detect a significant
616 difference in the fit to the data between the simpler constant population size
617 model and the exponential growth model (P = 0.46).

618

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

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

636 between *Hyaena* and *Parahyaena* in the nucleotide substitution rate estimation).
637 With the exponential growth model and the respective ML estimate of θ of 3.1,
638 the TMRCA estimate in coalescent time units was 1.27 ± 0.35, i.e. 304 ± 84 ka
639 (95% CI: 136-471 ka) (94-748 ka incorporating the aforementioned uncertainty
640 in the nucleotide substitution rate estimate). Analyses of the 340 bp dataset with
641 the constant population size model and a ML estimate of θ of 0.9 yielded a
642 TMRCA estimate in coalescent time units of 1.9 ± 0.8, which, using a nucleotide
643 substitution rate of 0.93% per Ma and a generation time of 9.2 years, translates
644 into 279 ± 118 ka (95% CI: 44-515 ka) (28-939 ka considering the uncertainty in
645 the nucleotide substitution rate estimate). Analyses of the same dataset with the
646 exponential growth model and a ML estimate of θ of 1.1 resulted in a TMRCA
647 estimate in coalescent time units of 1.38 ± 0.45, i.e. 240 ± 78 ka (95% CI: 83-397
648 ka) (52-723 ka considering the uncertainty in the nucleotide substitution rate
649 estimate) (Table 3). The fact that GENETREE TMRCA estimates using two very
650 different demographic models were quite similar suggests a minor effect of the
651 underlying demographic model on the estimates. TMRCAs estimates from the
652 BEAST analyses for different combinations of dataset, calibrations, and clock
653 model, are given in Table 3; ESS values were > 3000 for all parameters in all
654 individual runs under the strict clock, and > 200 under a random local clock.
655 Across the different analyses, the TMRCA point estimate for the observed mtDNA
656 variation in striped hyenas ranged from 330-480 ka, with the clock rate point
657 estimate between 1.1% and 1.6% per Ma, hovering around 1.3% per Ma (Table
658 3), which is quite close to the rate estimates of ≈ 1% that we inferred based on
659 the d_{xy} distance estimate between *Hyaena* and *Parahyaena*.

660

661 **Discussion**

662 Topics such as genetic diversity and structure, phylogeography, and evolutionary
663 history of the striped hyena remain poorly investigated (Wilkinson et al. 2023).
664 To date, our knowledge of the mitochondrial genetic diversity and
665 phylogeographic structure of the striped hyena comes essentially from the
666 analysis of a 340-bp *Cyt b* fragment in 13 individuals from across the historical
667 and present distribution of the species, which was part of a ground-breaking
668 study on the evolutionary history of extant hyenas (Rohland et al. 2005). That
669 analysis found low genetic diversity and no evident geographic structure of this
670 diversity. In our study we wanted to investigate the mitochondrial genetic
671 variation of the striped hyena in Algeria, using for this purpose a *Cyt b* fragment
672 encompassing the one analysed by Rohland et al. (2005). It was also our
673 objective, with the addition of samples from the Arabian Peninsula and Iran
674 (regions not covered in Rohland et al. (2005) and where classic morphological
675 subspecies have been described) and using a diverse set of demographic history
676 and coalescence time inference analyses, to conduct a more exhaustive and
677 better-sampled global analysis of the species than in Rohland et al. (2005). The
678 finding that a sample of 24 Algerian striped hyenas was monomorphic for a 753-
679 bp *Cyt b* fragment was striking, and agrees with the notion of low mitochondrial
680 diversity in the species. The apparent absence of phylogeographic structure in
681 the species seemingly extends also to southern Arabia and Iran (Figs. 2 and S2-
682 S4), with for example haplotype sharing between Oman and Iran for the 648-bp

683 fragment (Fig. 2a), which therefore also calls into question the validity of the
684 *sultana* and *hyaena* subspecies.

685 *Demographic history*

686 The observed patterns of moderate haplotype diversity and low nucleotide
687 diversity (Table 2) are compatible with a historical population bottleneck in the
688 species, followed by demographic growth and the appearance/accumulation of
689 new haplotypes (Grant and Bowen 1998). Moreover, the mismatch distributions
690 for both 648 bp and 340 bp datasets showed a significant fit with expected
691 distributions under a sudden expansion model. Nevertheless, it should be taken
692 into account that tests based on the mismatch distribution have been noted as
693 conservative in rejecting the null hypothesis of population expansion (Ramos-
694 Onsins and Rozas 2002; Fahey et al. 2014). On the other hand, none of the
695 mutation-drift equilibrium tests used rejected the hypothesis of constant
696 population size (Table 2). A reduced number of segregating sites, as is the case
697 here, can limit the power of these tests (Ramos-Onsins and Rozas, 2002). In
698 particular, the values of the F_S statistic were clearly negative, a result consistent
699 with expectations in a scenario of population growth, which may suggest a lack
700 of power in the datasets due to a small number of segregating sites. However, a
701 limited number of segregating sites does not necessarily prevent significant
702 results in those tests (Pilkington et al. 2008; Kvist et al. 2011; Cabanne et al.
703 2013; Álvarez-Varas et al. 2015; Hassan-Beigi et al. 2022). These tests also have
704 less power if the demographic expansion is historically very recent (e.g.
705 Holocene), especially if it was not very intense (Fu 1997; Ramos-Onsins and
706 Rozas 2002). Rohland et al. (2005) also did not detect a statistical signal of

707 demographic expansion in their data (the authors did not report which analyses
708 and tests of demographic history were performed), but the observed
709 phylogeographic pattern was inferred to be the result of a recent and rapid range
710 expansion, probably within the last 100 ka and possibly as recently as the
711 Holocene, from a small refuge population, most likely located in Africa as,
712 according to Werdelin and Solounias (1991), no Late Pleistocene fossils of
713 striped hyenas are know outside of Africa. The results of the single-tree skyline
714 plot analyses and demographic model fit comparisons all supported the
715 population growth hypothesis. However, these single-tree methods have the
716 limitation of not taking into account uncertainty in the tree topology (Strimmer
717 and Pybus 2001; Pybus and Rambaut 2002); we attempted to mitigate this
718 problem by obtaining the tree to be used by Bayesian analysis and summarizing
719 the posterior into a MCCT, the tree that maximizes the product of the posterior
720 clade probabilities. All the most sophisticated demographic inference methods
721 used (MCMC sampling-based skyline plots, IPHULA, LAMARC, GENETREE) could
722 not reject with high confidence (> 95%) the hypothesis of constant population
723 size. MCMC sampling-based skyline methods may allow to correctly identify
724 demographic trajectories based on single-locus data (Gill et al. 2013), but such
725 datasets may also lack the power to reject the hypothesis of constant
726 demography (Heled and Drummond 2008). For example, although the sample
727 size we analysed cannot be considered small (n=76), it has been noted in the
728 literature that the BSP method may lack the power to correctly detect
729 demographic expansions using single locus mitochondrial data when the sample
730 size is < 100 (Grant 2015). Again, also for those more complex methods it is
731 obviously preferable that the number of segregating sites is not small (Ho and

732 Shapiro 2011), but they have been applied to datasets with numbers of
733 segregating sites similar to our data, and produced significant results (e.g.
734 Griffiths and Tavaré 1994b, 1999; Kvist et al. 2011; Cabanne et al. 2013; Horreo
735 et al. 2013; Kuo et al. 2014; Álvarez-Varas et al. 2015; Ely et al. 2017). A small
736 number of segregating sites may not be the only challenge for the methods
737 implemented in IPHULA, LAMARC, and GENETREE. Another factor is the assumed
738 population growth model, which in the case of these three methods is the
739 exponential model. It is conceivable that the hypothesis of constant population
740 size may not be rejected with confidence if an exponential growth model is very
741 far from the true demographic history, which could be, for example, one of slow,
742 more or less linear growth, or a complex history alternating recurrent episodes
743 of expansion, stability, and/or decline (e.g. Excoffier and Schneider 1999;
744 Strimmer and Pybus 2001). For example, in the BEAST analysis with the
745 exponential growth coalescent model as tree prior, the marginal posterior
746 distribution of the exponential growth rate parameter included zero, implying
747 that a contrasting model of constant population size could not be rejected.
748 Overall, the observed patterns of haplotype and nucleotide diversity, the
749 mismatch distributions, and the single-tree skyline plot analyses and
750 demographic model fit comparisons supported the hypothesis of population
751 growth, confirming the inference by Rohland et al. (2005). This hypothesis was
752 not statistically significant in the more sophisticated and exacting methods used,
753 probably largely as a result of the limited power provided by a small number of
754 segregating sites.

755 It was surprising to observe only one haplotype in 24 individuals from Algeria,
756 whereas for example 41 animals from Iran showed eight haplotypes (648 bp
757 fragment; Table 1, Fig. 2a). It is therefore possible that the post-bottleneck
758 demographic growth within Africa was not as rapid and intense as that
759 associated with the colonization of Eurasia. If this was the case, then signatures
760 of demographic expansion will perhaps be easier to detect using large Asian
761 population samples. In the same way that in the human species, signatures of
762 demographic expansion are easier to detect in non-African populations than in
763 African ones (Ingman et al. 2000; Atkinson et al. 2008) or in pooled or global
764 datasets (e.g. Atkinson et al. 2009). Given the typical substitution rates in
765 mitochondrial DNA, it is expected that recent demographic expansions (e.g.
766 postglacial), particularly in organisms whose generation time is not short, may
767 be difficult to detect with data from relatively short mtDNA fragments (e.g. <
768 1000 bp), even using efficient methods in extracting the information present in
769 the data (Felsenstein 1992). Data from mitogenomes can be very useful in such
770 cases (Keis et al. 2013; Jensen et al. 2018). Future mitogenomic studies will likely
771 help to clarify open questions regarding the nature and age of the Late
772 Quaternary striped hyena demographic and geographic expansion(s) suggested
773 by the currently available mitochondrial data.

774 *Age of the extant mitochondrial DNA variation*

775 For the estimation of the TMRCA of the sampled genetic variation we also used
776 very different methods, always based on the same rationale that if different
777 analytical approaches yield congruent results, this increases our confidence in
778 them. The Thomson's estimator is simple and does not require the assumption of

779 a model for the demographic history of the population. GENETREE computes
780 TMRCA likelihood distributions conditional on the gene tree mutation structure,
781 a given value of θ, and the assumed model of population size history. For both
782 approaches the TMRCA estimates were translated into absolute time using the
783 estimated substitution rates for each fragment (340 bp and 648 bp). In contrast,
784 in BEAST the time information to calculate the TMRCA in years was derived from
785 fossil calibrations for the divergence between the striped hyena lineage and the
786 brown hyena and spotted hyena lineages. The TMRCA point estimates of the
787 observed genetic variation ranged between about 300 and 420 ka based on the
788 648 bp dataset, and between about 240 and 480 ka based on the 340 bp dataset
789 (Table 3); hence there was relative agreement between the estimates from the
790 two datasets. The variability of the estimates was lower for the dataset with
791 longer sequence length and more segregating sites, as would perhaps be
792 expected. For both datasets, estimates were higher with the Thomson's method
793 (648 bp: 419 ka; 340 bp: 410 ka) and with BEAST (648 bp: 334-387 ka; 340 bp:
794 441-484 ka) than with GENETREE (648 bp: 304-333 ka; 340 bp: 240-279 ka). This
795 pattern is in line with arguments that GENETREE results should be interpreted as
796 estimated lower bounds for the true values (Hammer et al. 1998; Karafet et al.
797 1999). In fact, most estimates from Thomson's method and BEAST tended to
798 hover around 400 ka, coinciding with one of the longest and warmest
799 interglacials of the last 800,000 years (within Marine Isotope Stage 11, MIS 11:
800 374-424 Ka, Lisiecki and Raymo 2005), with environmental conditions similar to
801 the Holocene (Howard 1997; Kleinen et al. 2014; Tzedakis et al. 2022). Rohland
802 et al. (2005), using the phylogenetic program r8s (Sanderson 2003) and a point
803 estimate of 10 Ma for the divergence between spotted and striped/brown

804 hyenas as a calibration date, estimated the TMRCA of striped hyenas at about
805 340 ka, which is quite congruent with our estimates. One of our goals in this
806 study was precisely to provide additional estimates, using different methods, for
807 comparison with that one obtained with r8s. For instance, it has been suggested
808 that BEAST may be more accurate than r8s for clade age estimation, particularly
809 with single locus mitochondrial data (Mulcahy et al. 2012).

810 *Nuclear DNA perspective*

811 A recent whole nuclear genome sequencing study on the evolutionary histories
812 of extant hyenas (Westbury et al. 2021), in which the striped hyena was
813 represented by a single captive individual of unknown geographic origin,
814 indicated a very low genetic diversity also in the nuclear genome of the species,
815 and estimated a demographic growth in the lineage starting ~ 500 ka that was
816 eventually followed by a sharp decline since ~ 100 ka. A priority in the near
817 future should be a genome-wide study with a broad geographic sampling
818 throughout Africa and Asia to allow a better characterization of genetic diversity,
819 and the impact on it of the anthropogenic reduction of the species in the last
820 centuries, and a more powerful and informative assessment of population
821 structure and demographic history, as already attempted for its closest extant
822 species, the brown hyena (Westbury et al. 2018).

823

824

825

Figures and Tables

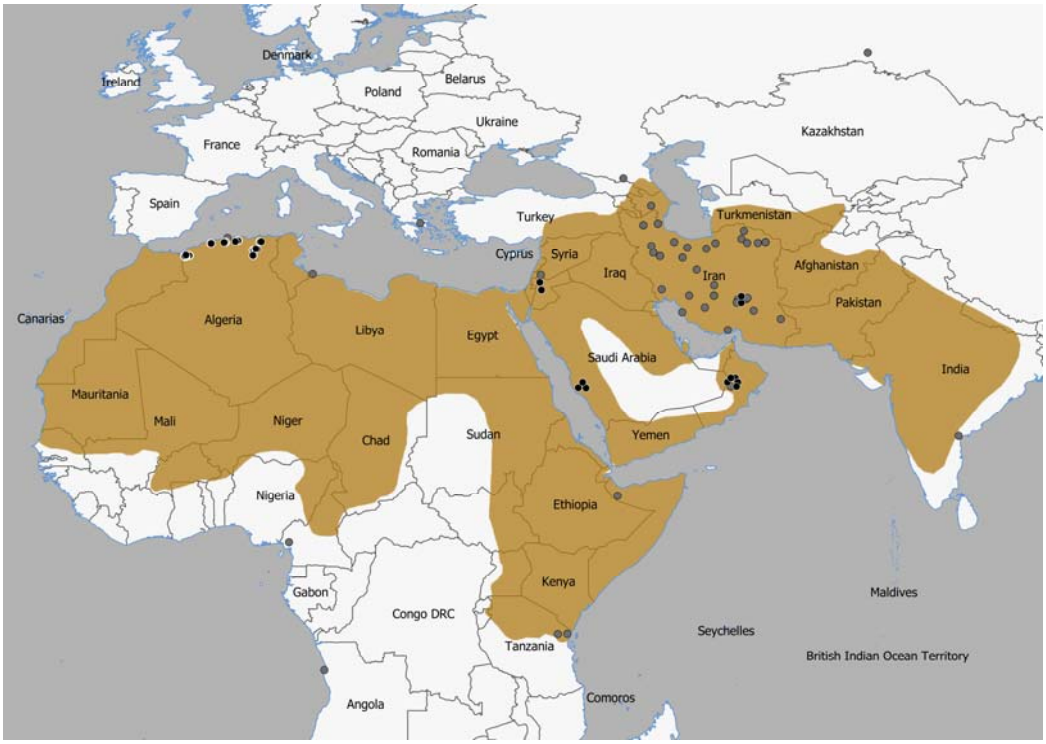


Fig. 1. Map showing the geographic range of the striped hyena according to Abi-Said and Dloniak (2015). The plotted dots indicate the origin of the samples and sequences analysed; black dots in the case of our samples and dark grey dots for Genbank sequences (see Table 1 for details).

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

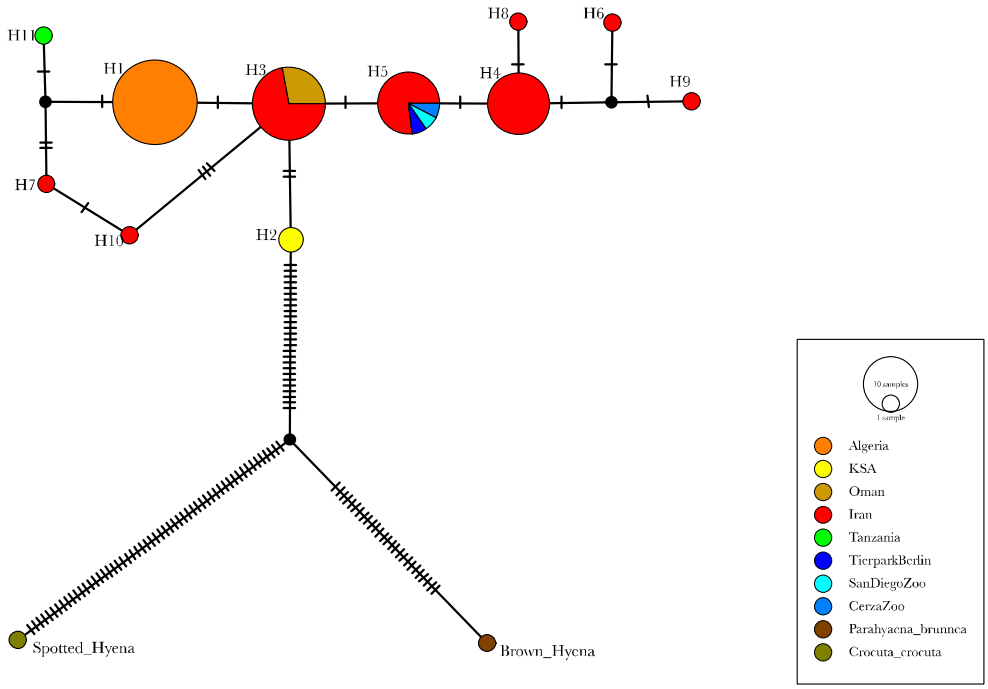
26

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 = 0.02)	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)


H11
H1
H3
H5
H4
H8
H6
H9
H7
H10
H2
Spotted_Hyena
Brown_Hyena
10 samples
1 sample
Algeria
KSA
Oman
Iran
Tanzania
TierparkBerlin
SanDiegoZoo
CerzaZoo
Parahyaena_brunnea
Crocuta_crocuta

b)

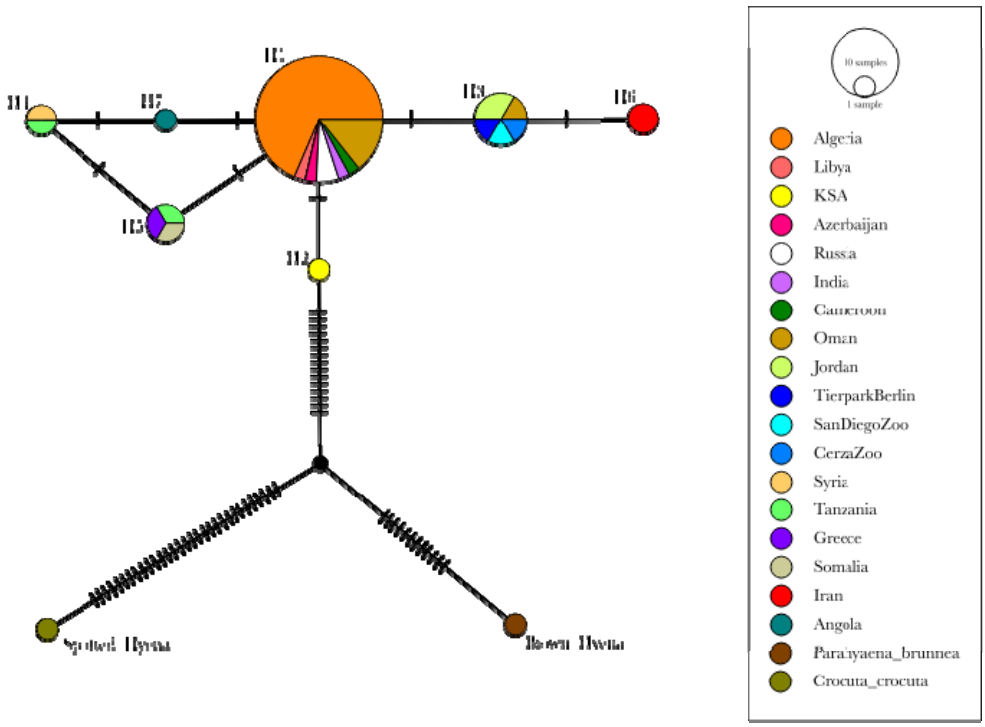

10 samples
1 sample
Algeria
Libya
KSA
Azerbaijan
Russia
India
Cameroon
Oman
Jordan
TierparkBerlin
SanDiegoZoo
CerzaZoo
Syria
Tanzania
Greece
Somalia
Iran
Angola
Parahyaena_brunnea
Crocuta_crocuta

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

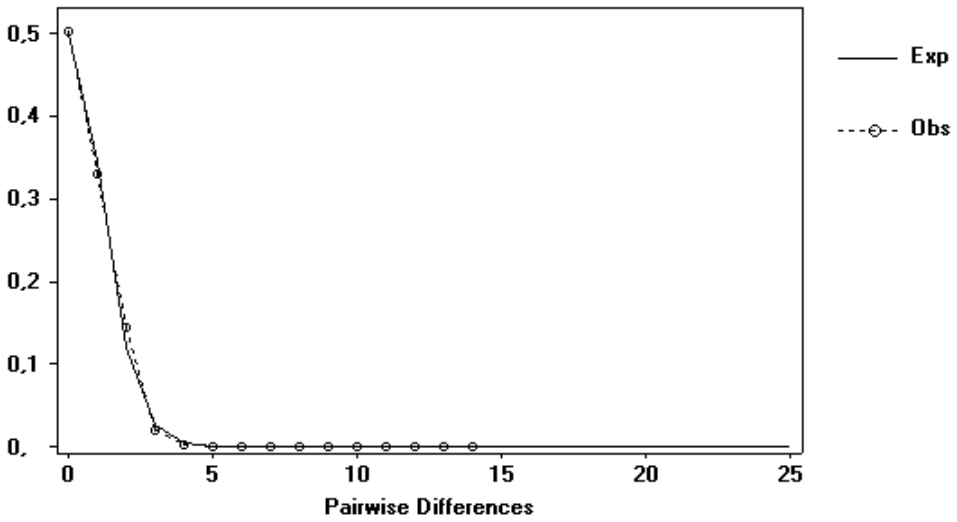
[c] For each of the analyses carried out in BEAST we also present the respective estimated clock rate in % per Ma

[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 Material

a)



b)

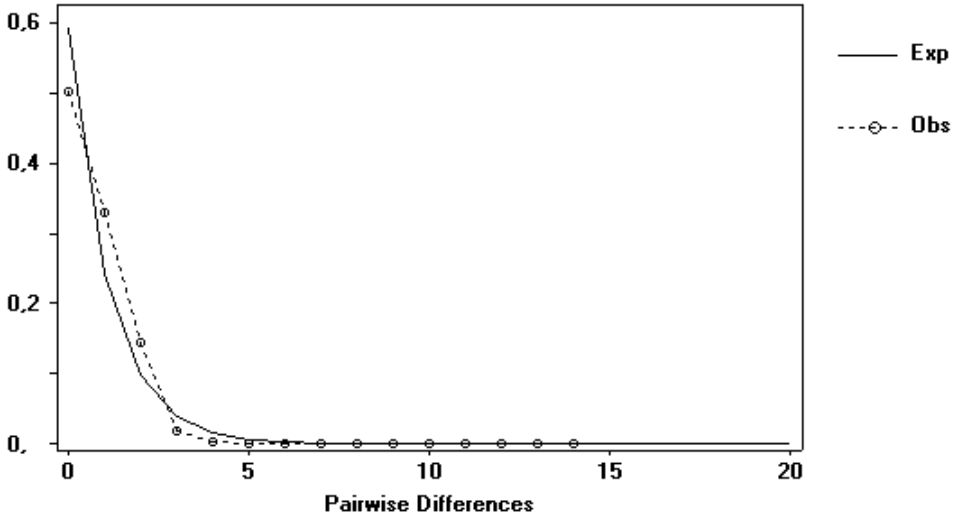
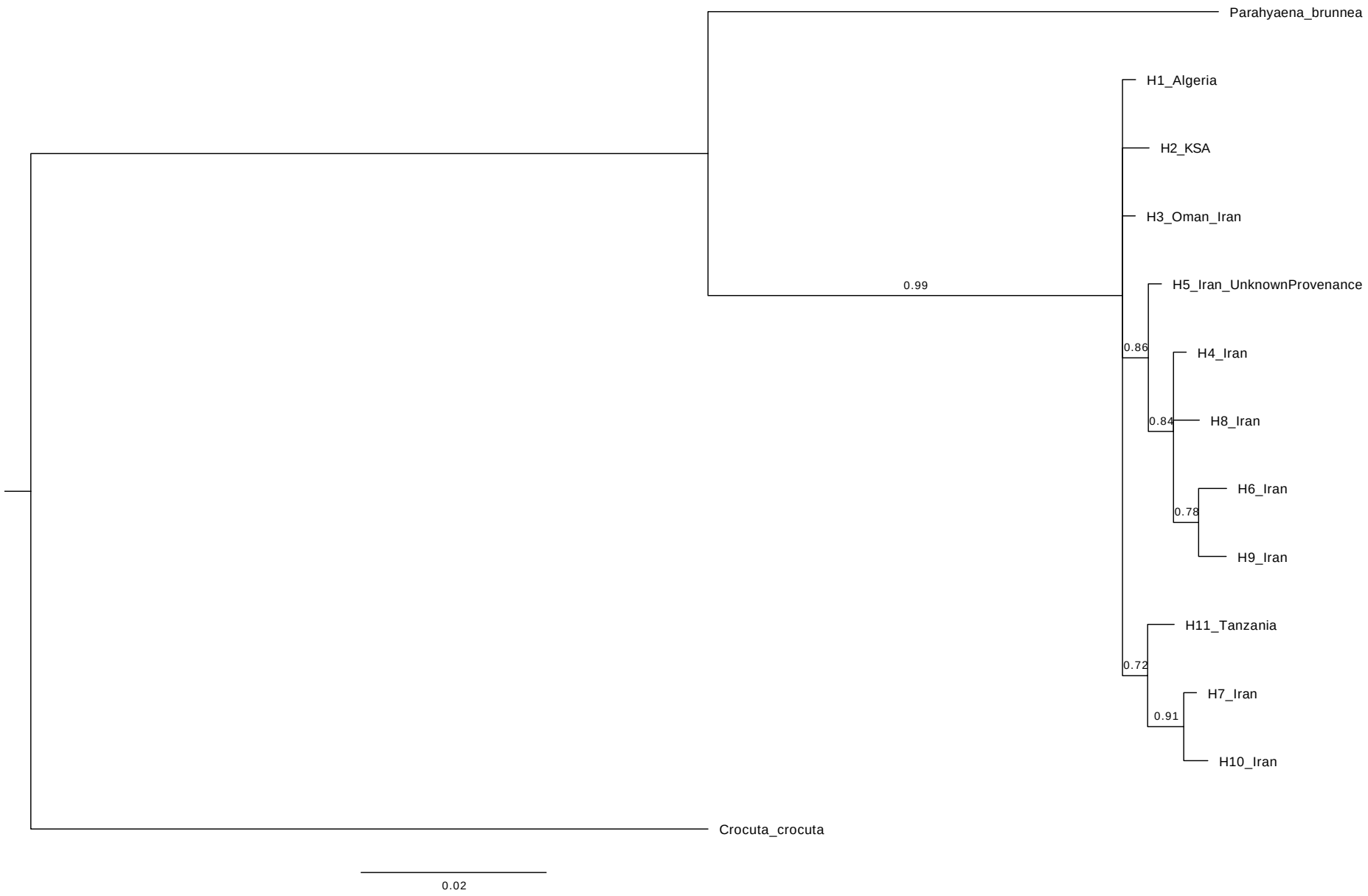


Fig. S1. DNASP graphs of mismatch distributions for the striped hyena 340 bp dataset comparing the observed distribution with (a) the expected values under a model of population growth or decline (Rogers and Harpending 1992), and (b) the expected values under a model of constant population size

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)

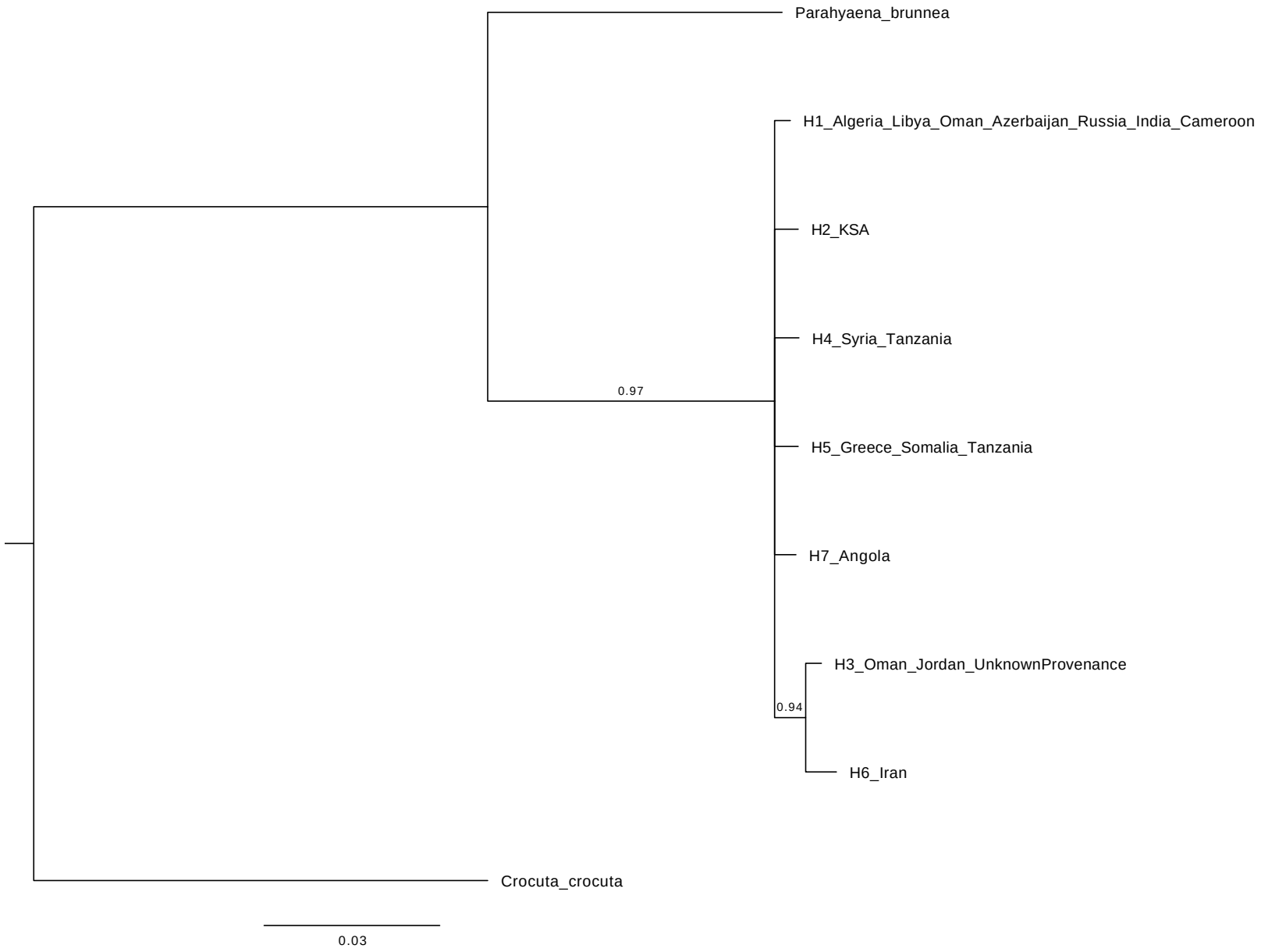

Parahyaena_brunnea
H1_Algeria_Libya_Oman_Azerbaijan_Russia_India_Cameroon
H2_KSA
H4_Syria_Tanzania
0.97
H5_Greece_Somalia_Tanzania
H7_Angola
H3_Oman_Jordan_UnknownProvenance
0.94
H6_Iran
Crocuta_crocuta
0.03

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.

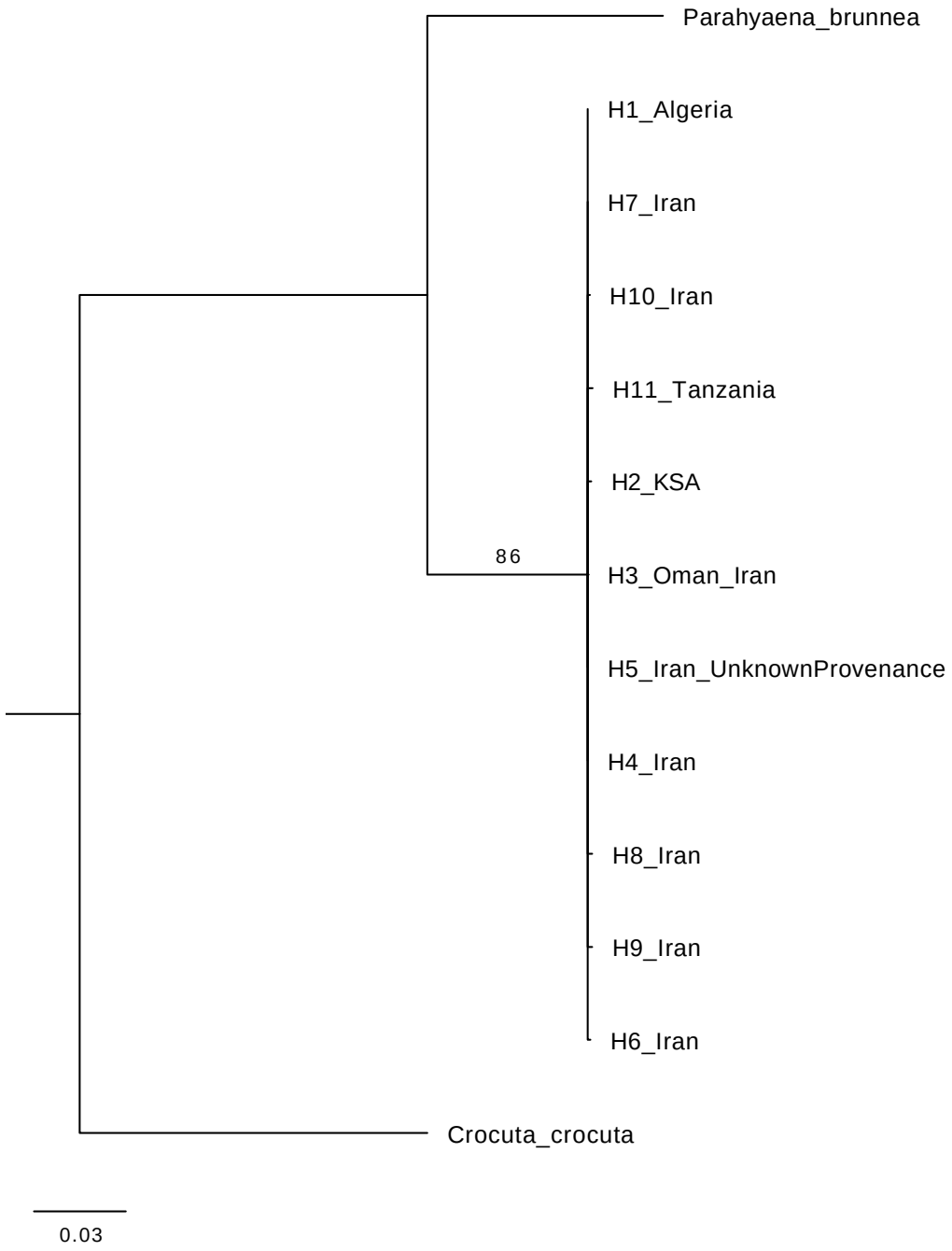


Fig. S3. Majority-rule consensus phylogram (cut-off 70%) of the maximum likelihood analysis of the 648 bp dataset with two outgroups (*Parahyaena brunnea* and *Crocuta crocuta*). See Table 1 for haplotype information. Numbers above branches are bootstrap values.

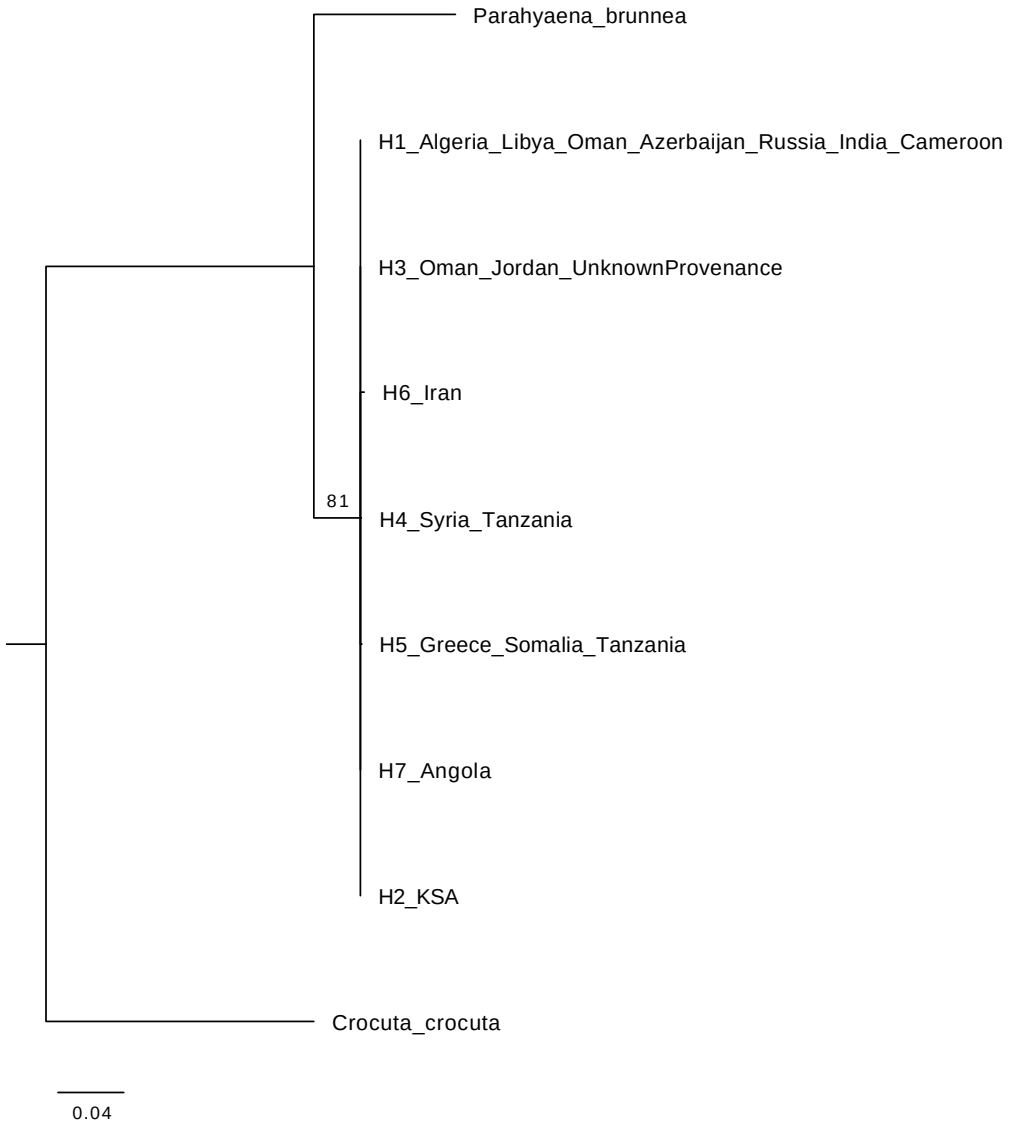
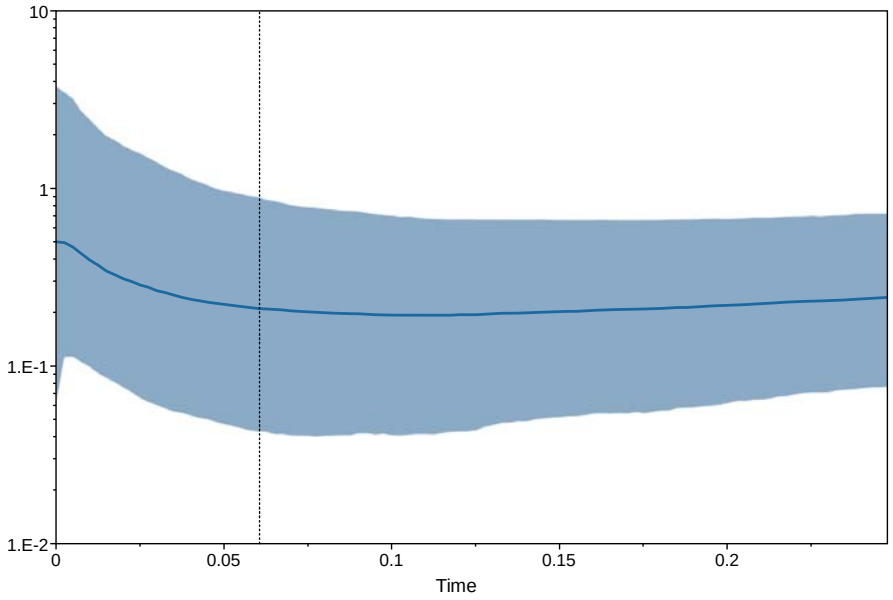


Fig. S4. Majority-rule consensus phylogram (cut-off 70%) of the maximum likelihood analysis of the 340 bp dataset with two outgroups (*Parahyaena brunnea* and *Crocuta crocuta*). See Table 1 for haplotype information. Numbers above branches are bootstrap values.

a)



b)

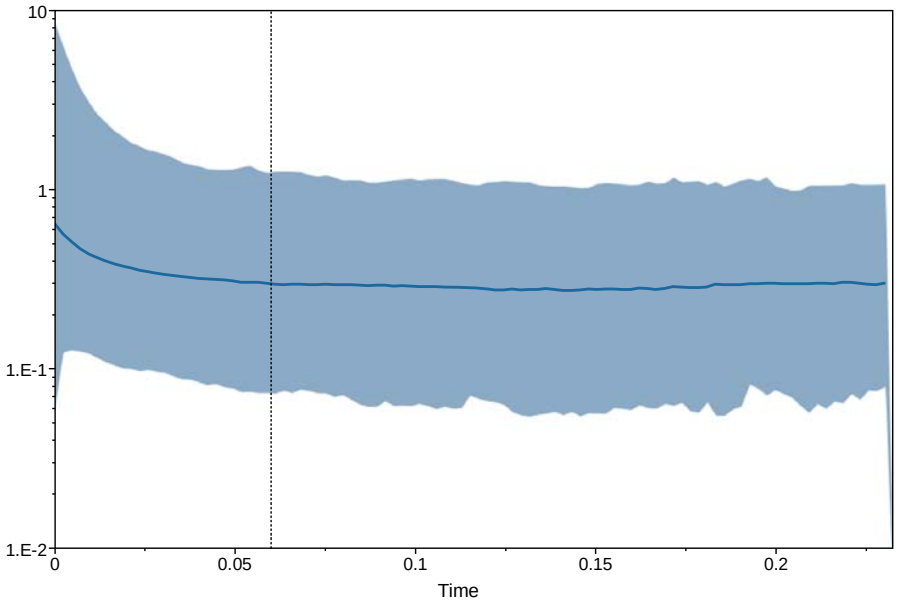


Fig. S5. BSPs for the sample of 76 striped hyena 648-bp *Cyt b* sequences, using the a) piecewise-constant and b) piecewise-linear models, and a normally distributed prior for the substitution rate with mean 0.01 and standard deviation 0.0025 (in substitutions per site per Ma). The thicker solid dark 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 Y-axis is in logarithmic scale and the X-axis in linear scale.

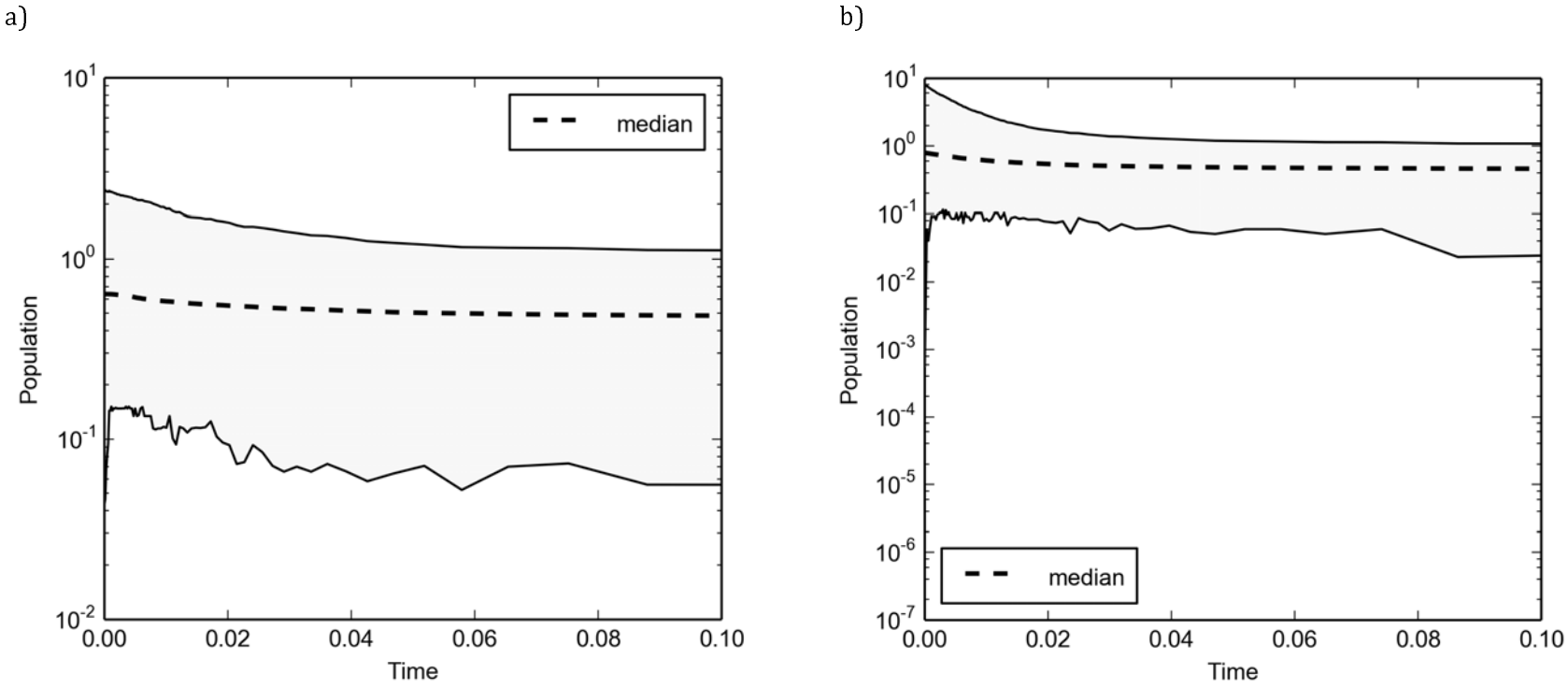


Fig. S6. EBSPs for the sample of 76 striped hyena 648-bp *Cyt b* sequences, using the a) stepwise and b) linear models, and a normally distributed prior for the substitution rate with mean 0.01 and standard deviation 0.0025 (in substitutions per site per Ma). The thicker dashed line is the median estimate of population size, and the shaded region represents the 95% HPD envelope. The Y-axis is in logarithmic scale and the X-axis in linear scale.

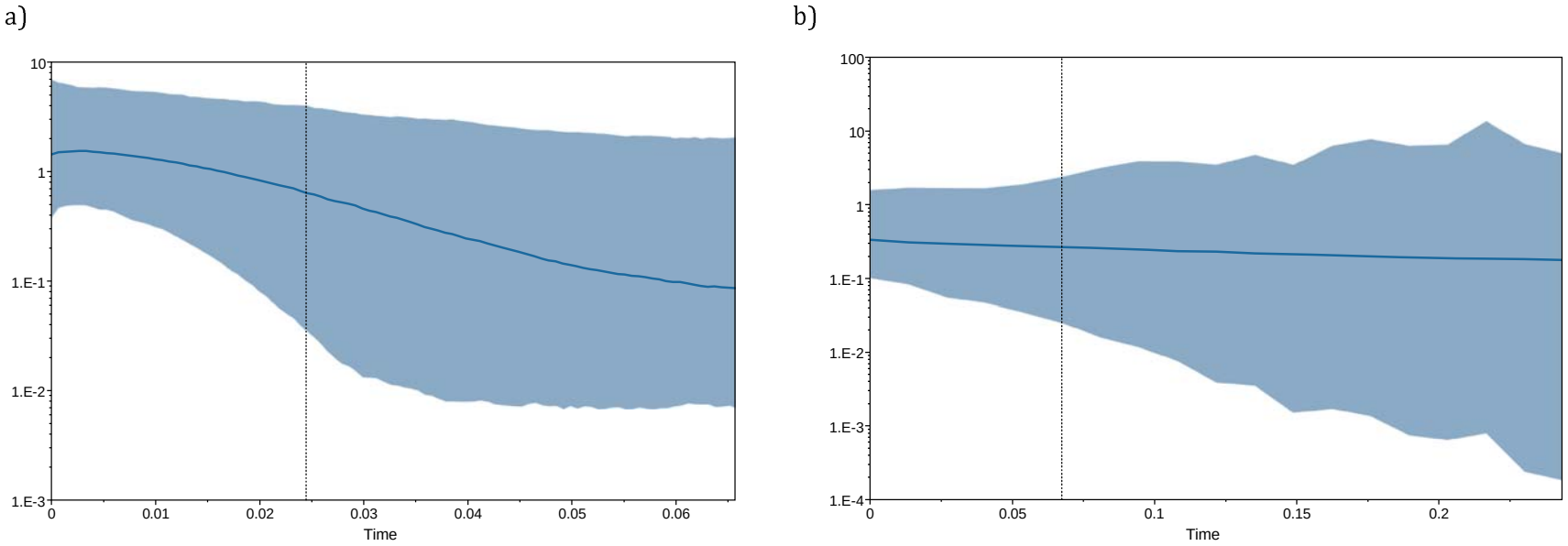


Fig. S7. Skyline plots for the sample of 76 striped hyena 648-bp *Cyt b* sequences, using the a) skyride and b) skygrid methods, and a normally distributed prior for the substitution rate with mean 0.01 and standard deviation 0.0025 (in substitutions per site per Ma). The thicker solid dark blue lines 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 Y-axes are in logarithmic scale and the X-axes in linear scale.

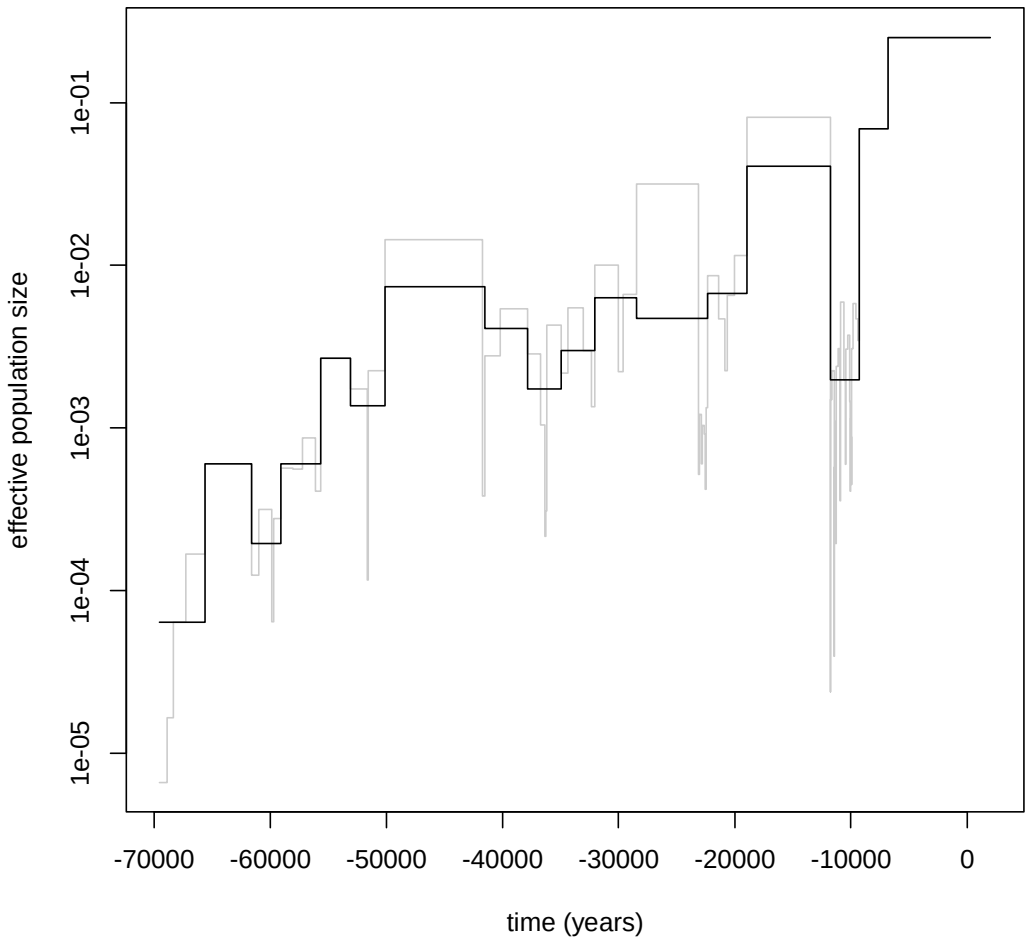


Fig. S8. Generalized skyline plot (black line), and classic skyline plot (grey line) in the background. Time is zero at the present.

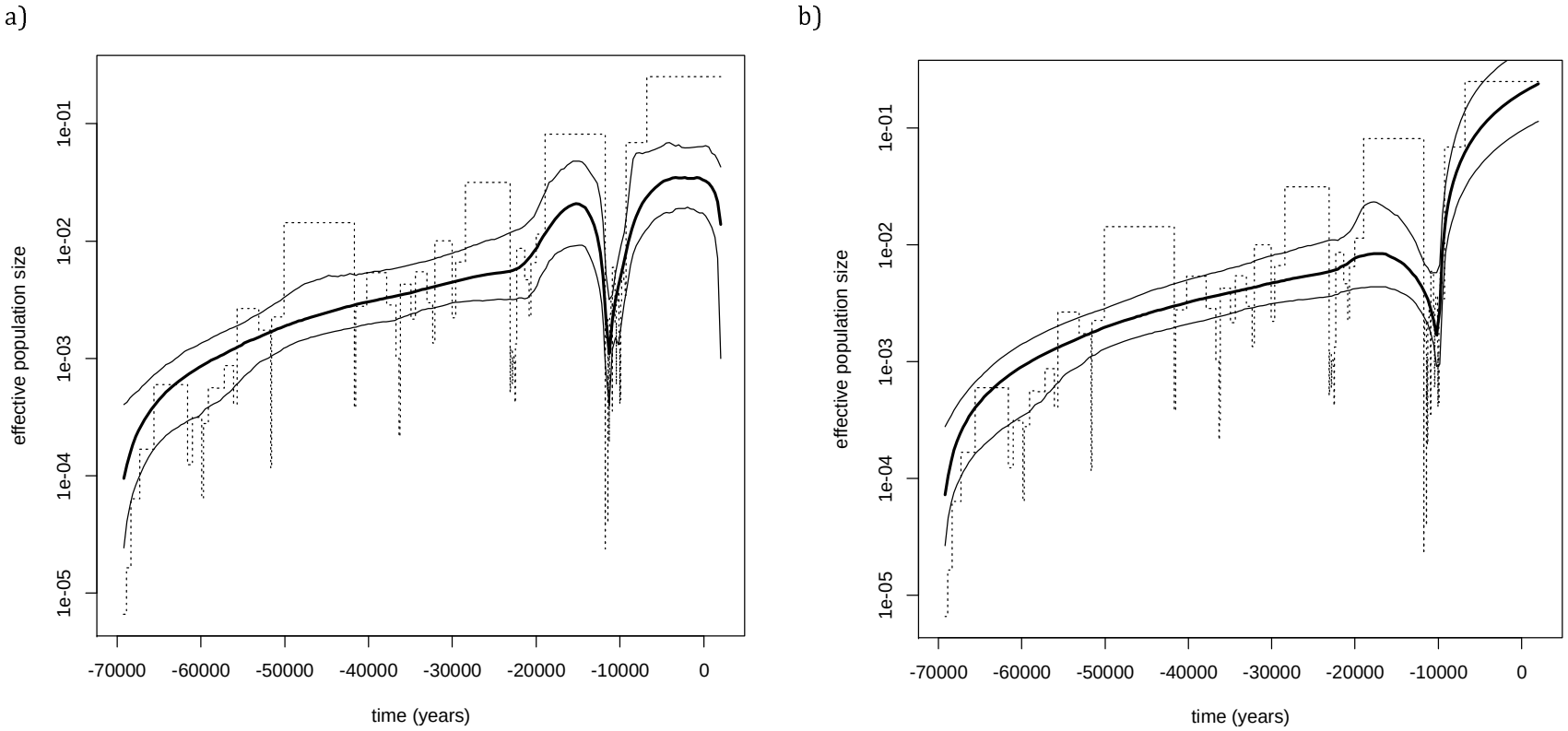


Fig. S9. Bayesian MCP plots estimated using a) a constant population size prior and b) a skyline plot prior. The thicker black lines are the median of the posterior distribution of population size, and the thinner lines represent the 95% confidence intervals. In both figures, for comparison, the classic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

Competing Interests and Funding

On behalf of all authors, the corresponding author states that there is no conflict of interest. This work was supported by Portuguese National Funds provided by FCT (Fundação para a Ciência e a Tecnologia) through strategic projects granted to CE3C: UID/BIA/00329/2019, UIDB/00329/2020 (https://doi.org/10.54499/UIDB/00329/2020), and UIDP/00329/2020 (https://doi.org/10.54499/UIDP/00329/2020).

Acknowledgments

The costs associated with the sampling were supported by the private resources of Louiza Derouiche. Carlos Rodríguez Fernandes thanks the support of CE3C through an assistant researcher contract (FCiência.ID contract #366) and FPUL for a contract of invited assistant professor. 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), and UIDP/00329/2020 (https://doi.org/10.54499/UIDP/00329/2020). We thank Yassine Righi for a striped hyena sample from Belezma National Park, Algeria. We are extremely grateful to Dr. Helmut Simon for all his advice on TMRCA estimation, and for the NumPy code to compute Thomson's estimator. We thank Professor Robert Griffiths for his advice on running GENETREE and analysing the results. We thank Dr. Emmanuel Paradis for his advice on how to compute the

unfolded site frequency spectrum in his R package PEGAS. Lastly, we would like to thank Dr. Heiko A. Schmidt for his advice in installing the IPHULA software.

References

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

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

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

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

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

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

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

Atkinson, Q.D., Gray, R.D., Drummond, A.J., 2009. Bayesian coalescent inference of major human mitochondrial DNA haplogroup expansions in Africa. Proc. Biol. Sci. 276, 367–373. https://doi.org/10.1098/rspb.2008.0785

Atkinson, Q.D., Gray, R.D., Drummond, A.J., 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, D.L., Cummings, M.P., Baele, G., Darling, A.E., Lewis, P.O., Swofford, D.L., Huelsenbeck, J.P., Lemey, P., Rambaut, A., Suchard, M.A., 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, F.U., Filipski, A., Hedges, S.B., 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, D.L., Hewitt, G.M., 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, D.R., 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, A.M., Duchêne, D.A., Ho, S.Y.W., 2018. Bayesian molecular dating: opening up the black box. Biol. Rev. 93, 1165–1191. https://doi.org/10.1111/brv.12390

Brown, J.W., Smith, S.A., 2018. The past sure is tense: on interpreting phylogenetic divergence time estimates. Syst. Biol. 67, 340–353. https://doi.org/10.1093/sysbio/syx074

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

Brown, W.M., George, M., Wilson, A.C., 1979. Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. U.S.A. 76, 1967–1971. https://doi.org/10.1073/pnas.76.4.1967

Cabanne, G.S., Sari, E.H.R., Meyer, D., Santos, F.R., Miyaki, C.Y., 2013. Matrilineal evidence for demographic expansion, low diversity and lack of phylogeographic structure in the Atlantic forest endemic Greenish Schiffornis *Schiffornis virescens* (Aves: Tityridae). J Ornithol. 154, 371–384. https://doi.org/10.1007/s10336-012-0901-8

Cassens, I., Van Waerebeek, K., Best, P.B., Crespo, E.A., Reyes, J., Milinkovitch, M.C., 2003. The phylogeography of dusky dolphins (*Lagenorhynchus obscurus*): a critical examination of network methods and rooting procedures. Mol. Ecol. 12, 1781–1792. https://doi.org/10.1046/j.1365-294X.2003.01876.x

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

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772. https://doi.org/10.1038/nmeth.2109

Derouiche, L., Irzagh, A., Rahmouni, R., Tahri, R., Hadjeloum, M., Bouhadad, R., Fernandes, C., 2020. Deep mitochondrial DNA phylogeographic divergence in the threatened aoudad *Ammotragus lervia* (Bovidae, Caprini). Gene 739, 144510. https://doi.org/10.1016/j.gene.2020.144510

Drummond, A.J., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 22, 1185–1192. https://doi.org/10.1093/molbev/msi103

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88. https://doi.org/10.1371/journal.pbio.0040088

Drummond, A.J., Suchard, M.A., 2010. Bayesian random local clocks, or one rate to rule them all. BMC Biol. 8, 114. https://doi.org/10.1186/1741-7007-8-114

Dubach, J.M., Briggs, M.B., White, P.A., Ament, B.A., Patterson, B.D., 2013. Genetic perspectives on "Lion Conservation Units" in Eastern and Southern Africa. Conserv. Genet. 14, 741–755. https://doi.org/10.1007/s10592-013-0453-3

Ely, C.R., Wilson, R.E., Talbot, S.L., 2017. Genetic structure among greater white-fronted goose populations of the Pacific Flyway. Ecol. Evol. 7, 2956–2968. https://doi.org/10.1002/ece3.2934

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

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

Excoffier, L., Schneider, S., 1999. Why hunter-gatherer populations do not show signs of pleistocene demographic expansions. Proc. Natl. Acad. Sci. U.S.A 96, 10597–10602. https://doi.org/10.1073/pnas.96.19.10597

Fahey, A.L., Ricklefs, R.E., Dewoody, J.A., 2014. DNA-based approaches for evaluating historical demography in terrestrial vertebrates. Biol. J. Linn. Soc. Lond. 112, 367–386. https://doi.org/10.1111/bij.12259

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

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

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

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

Ferreira, M.A.R., Suchard, M.A., 2008. Bayesian analysis of elapsed times in continuous-time Markov chains. Can. J. Statistics 36, 355–368. https://doi.org/10.1002/cjs.5550360302

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

Fu, Y.X., Li, W.H., 1993. Statistical tests of neutrality of mutations. Genetics 133, 693–709. https://doi.org/10.1093/genetics/133.3.693

Gill, M.S., Lemey, P., Faria, N.R., Rambaut, A., Shapiro, B., Suchard, M.A., 2013. Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. Mol. Biol. Evol. 30, 713–724. https://doi.org/10.1093/molbev/mss265

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

Grant, W.A.S., Bowen, B.W., 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. J. Hered. 89, 415–426. https://doi.org/10.1093/jhered/89.5.415

Grant, W.S., 2015. Problems and cautions with sequence mismatch analysis and Bayesian skyline plots to infer historical demography. J. Hered. 106, 333–346. https://doi.org/10.1093/jhered/esv020

Griffiths, R.C., Tavaré, S., 1994a. Simulating probability distributions in the coalescent. Theor. Popul. Biol. 46, 131–159. https://doi.org/10.1006/tpbi.1994.1023

Griffiths, R.C., Tavaré, S., 1994b. Sampling theory for neutral alleles in a varying environment. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 344, 403–410. https://doi.org/10.1098/rstb.1994.0079

Griffiths, R.C., Tavaré, S., 1994c. Ancestral inference in population genetics. Stat. Sci. 9, 307–319. https://doi.org/10.1214/ss/1177010378

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

Griffiths, R.C., Tavaré, S., 1999. The ages of mutations in gene trees. Ann. Appl. Prob. 9, 567–590. https://doi.org/10.1214/aoap/1029962804

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

Hammer, M.F., Karafet, T., Rasanayagam, A., Wood, E.T., Altheide, T.K., Jenkins, T., Griffiths, R.C., Templeton, A.R., Zegura, S.L., 1998. Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. Mol. Biol. Evol. 15, 427–441. https://doi.org/10.1093/oxfordjournals.molbev.a025939

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

Harding, R.M., Fullerton, S.M., Griffiths, R.C., Bond, J., Cox, M.J., Schneider, J.A., Moulin, D.S., Clegg, J.B., 1997. Archaic African and Asian lineages in the genetic ancestry of modern humans. Am. J. Hum. Genet. 60, 772–789.

Harpending, H.C., 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Hum. Biol. 66, 591–600.

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

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

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

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

Ho, S.Y.W., Phillips, M.J., Drummond, A.J., Cooper, A., 2005. Accuracy of rate estimation using relaxed-clock models with a critical focus on the early metazoan radiation. Mol. Biol. Evol. 22, 1355–1363. https://doi.org/10.1093/molbev/msi125

Ho, S.Y.W., Shapiro, B., 2011. Skyline-plot methods for estimating demographic history from nucleotide sequences. Mol. Ecol. Resour. 11, 423–434. https://doi.org/10.1111/j.1755-0998.2011.02988.x

Holder, M.T., Sukumaran, J., Lewis, P.O., 2008. A justification for reporting the majority-rule consensus tree in Bayesian phylogenetics. Syst. Biol. 57, 814–821. https://doi.org/10.1080/10635150802422308

Holekamp, K.E. and Kolowski, J.M., 2009. Hyaenidae. In: Wilson, D.E., Mittermeier R.A. (Eds.), Handbook of the mammals of the world. Volume 1: Carnivores. Lynx Edicions, Madrid, Spain, pp. 234–260.

Horreo, J.L., Palacín, C., Alonso, J.C., Milá, B., 2013. A link between historical population decline in the threatened great bustard and human expansion in Iberia: evidence from genetic and demographic data. Biol. J. Linn. Soc. Lond. 110, 518–527. https://doi.org/10.1111/bij.12152

Howard, W.R., 1997. A warm future in the past. Nature 388, 418–419. https://doi.org/10.1038/41201

Hu, J., Westbury, M.V., Yuan, J., Zhang, Z., Chen, S., Xiao, B., Hou, X., Ji, H., Lai, X., Hofreiter, M., Sheng, G., 2021. Ancient mitochondrial genomes from Chinese cave hyenas provide

insights into the evolutionary history of the genus *Crocuta*. Proc. R. Soc. B. 288, 20202934. https://doi.org/10.1098/rspb.2020.2934

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

Huelsenbeck, J.P., Larget, B., Alfaro, M.E., 2004. Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. Mol. Biol. Evol. 21, 1123–1133. https://doi.org/10.1093/molbev/msh123

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

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

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

Irwin, D.M., Kocher, T.D., Wilson, A.C., 1991. Evolution of the cytochrome *b* gene of mammals. J. Mol. Evol. 32, 128–144. https://doi.org/10.1007/BF02515385

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

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

Jenks, S.M., Werdelin, L. 1998. Taxonomy and systematics of living hyaenas (Family Hyaenidae). In: Mills, M.G.L., Hofer, H. (Compilers), Hyaenas: Status Survey and

Conservation Action Plan. IUCN/SSC Hyaena Specialist Group, IUCN, Gland, Switzerland, pp. 8–17.

Jensen, E.L., Miller, J.M., Edwards, D.L., Garrick, R.C., Tapia, W., Caccone, A., Russello, M.A., 2018. Temporal mitogenomics of the Galapagos giant tortoise from Pinzón reveals potential biases in population genetic inference. J. Hered. 109, 631–640. https://doi.org/10.1093/jhered/esy016

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., Von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285

Karafet, T.M., Zegura, S.L., Posukh, O., Osipova, L., Bergen, A., Long, J., Goldman, D., Klitz, W., Harihara, S., de Knijff, P., Wiebe, V., Griffiths, R.C., Templeton, A.R., Hammer, M.F., 1999. Ancestral Asian source(s) of new world Y-chromosome founder haplotypes. Am. J. Hum. Genet. 64, 817–831. https://doi.org/10.1086/302282

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

Kass, R.E., Raftery, A.E., 1995. Bayes Factors. J. Am. Stat. Assoc. 90, 773–795. https://doi.org/10.1080/01621459.1995.10476572

Keis, M., Remm, J., Ho, S.Y.W., Davison, J., Tammeleht, E., Tumanov, I.L., Saveljev, A.P., Männil, P., Kojola, I., Abramov, A.V., Margus, T., Saarma, U., 2013. Complete mitochondrial genomes and a novel spatial genetic method reveal cryptic phylogeographical structure and migration patterns among brown bears in north-western Eurasia. J. Biogeogr. 40, 915–927. https://doi.org/10.1111/jbi.12043

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

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

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

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

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

Koepfli, K.-P., Jenks, S.M., Eizirik, E., Zahirpour, T., Valkenburgh, B.V., Wayne, R.K., 2006. Molecular systematics of the Hyaenidae: Relationships of a relictual lineage resolved by a molecular supermatrix. Mol. Phylogenet, Evol. 38, 603–620. https://doi.org/10.1016/j.ympev.2005.10.017

Kowalski, K., Rzebik-Kowalska, B., 1991. Mammals of Algeria. Polish Academy of Science, Institute of Systematics and Evolution of Animals, Wroclaw.

Kuhner, M.K., 2006. LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics 22, 768–770. https://doi.org/10.1093/bioinformatics/btk051

Kuhner, M.K., Smith, L.P., 2007. Comparing likelihood and Bayesian coalescent estimation of population parameters. Genetics 175, 155–165. https://doi.org/10.1534/genetics.106.056457

Kuhner, M.K., Yamato, J., Felsenstein, J., 1998. Maximum likelihood estimation of population growth rates based on the coalescent. Genetics 149, 429–434. https://doi.org/10.1093/genetics/149.1.429

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

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

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

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34, 772–773. https://doi.org/10.1093/molbev/msw260

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

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

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

Lisiecki, L.E., Raymo, M.E., 2005. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. Paleoceanography 20, PA1003. https://doi.org/10.1029/2004PA001071

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

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

Mills, M.G.L., Hofer, H., 1998. Hyaenas: status survey and conservation action plan. IUCN/SSC Hyaena Specialist Group, IUCN, Gland, Switzerland.

Minin, V.N., Bloomquist, E.W., Suchard, M.A., 2008. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. Mol. Biol. Evol. 25, 1459–1471. https://doi.org/10.1093/molbev/msn090

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

Morrison, D.A., 2008. How to summarize estimates of ancestral divergence times. Evol Bioinform 4, 75-95. https://doi.org/10.4137/EBO.S545

Mulcahy, D.G., Noonan, B.P., Moss, T., Townsend, T.M., Reeder, T.W., Sites, J.W., Wiens, J.J., 2012. Estimating divergence dates and evaluating dating methods using phylogenomic and mitochondrial data in squamate reptiles. Mol. Phylogenet. Evol. 65, 974–991. https://doi.org/10.1016/j.ympev.2012.08.018

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

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

Nguyen, L.-T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. https://doi.org/10.1093/molbev/msu300

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

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

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

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

Pilkington, M.M., Wilder, J.A., Mendez, F.L., Cox, M.P., Woerner, A., Angui, T., Kingan, S., Mobasher, Z., Batini, C., Destro-Bisol, G., Soodyall, H., Strassmann, B.I., Hammer, M.F.,

2008. Contrasting signatures of population growth for mitochondrial DNA and Y chromosomes among human populations in Africa. Mol. Biol. Evol. 25, 517–525. https://doi.org/10.1093/molbev/msm279

Pocock, R.I., 1934. The races of the striped and brown hyaenas. P. Zool. Soc. Lond. 104, 799–825. https://doi.org/10.1111/j.1096-3642.1934.tb01667.x

Pybus, O.G., Rambaut, A., 2002. GENIE: estimating demographic history from molecular phylogenies. Bioinformatics 18, 1404–1405. https://doi.org/10.1093/bioinformatics/18.10.1404

Pybus, O.G., Rambaut, A., Harvey, P.H., 2000. An integrated framework for the inference of viral population history from reconstructed genealogies. Genetics 155, 1429–1437. https://doi.org/10.1093/genetics/155.3.1429

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

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

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

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

Rieger, I., 1981. *Hyaena hyaena*. Mammalian Species 150, 1–5. https://doi.org/10.2307/41353899

Ritchie, A.M., Lo, N., Ho, S.Y.W., 2017. The impact of the tree prior on molecular dating of data sets containing a mixture of inter- and intraspecies sampling. Syst. Biol. 66, 413–425. https://doi.org/10.1093/sysbio/syw095

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

Robinson-Rechavi, M., Huchon, D., 2000. RRTree: Relative-Rate Tests between groups of sequences on a phylogenetic tree. Bioinformatics 16, 296–297. https://doi.org/10.1093/bioinformatics/16.3.296

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

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

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542. https://doi.org/10.1093/sysbio/sys029

Sanderson, M.J., 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19, 301–302. https://doi.org/10.1093/bioinformatics/19.2.301

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

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

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

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.

Smith, M.F., Thomas, W.K., Patton, J.L., 1992. Mitochondrial DNA-like sequence in the nuclear genome of an akodontine rodent. Mol. Biol. Evol. 9, 204–215. https://doi.org/10.1093/oxfordjournals.molbev.a040714

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

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

Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol. 4, vey016. https://doi.org/10.1093/ve/vey016

Sugiura, N., 1978. Further analysis of the data by Akaike's information criterion and the finite corrections. Commun. Stat. – Theor. M. 7, 13–26. https://doi.org/10.1080/03610927808827599

Sukumaran, J., Holder, M.T., 2010. DendroPy: a Python library for phylogenetic computing. Bioinformatics 26, 1569–1571. https://doi.org/10.1093/bioinformatics/btq228

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

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

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

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

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

Thomson, R., Pritchard, J.K., Shen, P., Oefner, P.J., Feldman, M.W., 2000. Recent common ancestry of human Y chromosomes: Evidence from DNA sequence data. Proc. Natl. Acad. Sci. U.S.A. 97, 7360–7365. https://doi.org/10.1073/pnas.97.13.7360

Tourani, M., Moqanaki, E.M., Kiabi, B.H., 2012. Vulnerability of Striped Hyaenas, *Hyaena hyaena* , in a human-dominated landscape of Central Iran. Zool. Middle East 56, 133–136. https://doi.org/10.1080/09397140.2012.10648948

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

Tzedakis, P.C., Hodell, D.A., Nehrbass-Ahles, C., Mitsui, T., Wolff, E.W., 2022. Marine Isotope Stage 11c: An unusual interglacial. Quaternary Sci. Rev. 284, 107493. https://doi.org/10.1016/j.quascirev.2022.107493

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

Villanea, Kitchen, Kemp, 2019. Applications of Bayesian skyline plots and approximate Bayesian computation for human demography. Human Biology 91, 279. https://doi.org/10.13110/humanbiology.91.4.04

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

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

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

Werdelin, L., 2003. Carnivora from the Kanapoi hominid site, Turkana Basin, northern Kenya, in: Harris, J.M., Leakey, M.G. (Eds.), Geology and Vertebrate Paleontology of the Early Pliocene Site of Kanapoi, northern Kenya. Contrib. Sci. 498, 115–132.

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

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

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

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

Westbury, M.V., Cahsan, B.D., Dalerum, F., Norén, K., Hofreiter, M., 2019. Aardwolf population diversity and phylogenetic positioning inferred using complete mitochondrial genomes. Afr. J. Wildl. Res. 49, 27-33. https://doi.org/10.3957/056.049.0027

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

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

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

Wilkinson, C.E., Dheer, A., Zett, T., Torrents-Ticó, M., Yarnell, R.W., Bar Ziv, E., Bhandari, S., Jacobson, A., Dloniak, S.M., 2023. Review of the global research on Hyaenidae and implications for conservation and management. Mammal Rev. mam.12337. https://doi.org/10.1111/mam.12337

Wilson, A.C., Cann, R.L., Carr, S.M., George, M., Gyllensten, U.B., Helm-Bychowski, K.M., Higuchi, R.G., Palumbi, S.R., Prager, E.M., Sage, R.D., 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, W.C., 2005. Order Carnivora. In: Wilson, D.E., Reeder, D.M. (Eds.), Mammal Species of the World: A Taxonomic and Geographic Reference (3rd edition). 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. JOSS 4, 634. https://doi.org/10.21105/joss.00634

Xie, W., Lewis, P.O., 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, D.X., Hewitt, G.M., 1996. Nuclear integrations: challenges for mitochondrial DNA markers. Trends Ecol. Evol. 11, 247–251. https://doi.org/10.1016/0169-5347(96)10031-8