

1 **Genome-wide coancestry reveals details of ancient and recent male-driven reticulation in**
2 **baboons**

3 **Short title: Baboon evolutionary reticulation**

4
5 Erik F. Sørensen^{1†}, R. Alan Harris^{2†}, Liye Zhang^{3†}, Muthuswamy Raveendran^{2†}, Lukas F. K.
6 Kuderna^{4,5}, Jerilyn A. Walker⁶, Jessica M. Storer⁷, Martin Kuhlwilm^{4,8,9}, Claudia Fontseré⁴,
7 Lakshmi Seshadri³, Christina M. Bergey¹⁰, Andrew S. Burrell¹¹, Juraj Bergmann^{1,12}, Jane E.
8 Phillips-Conroy^{13,14}, Fekadu Shiferaw¹⁵, Kenneth L. Chiou^{16,17}, Idrissa S. Chuma¹⁸, Julius D.
9 Keyyu¹⁹, Julia Fischer^{20,21,22}, Marie-Claude Gingras², Sejal Salvi², Harshavardhan Doddapaneni²,
10 Mikkel H. Schierup¹, Mark A. Batzer⁶, Clifford J. Jolly¹¹, Sascha Knauf²³, Dietmar Zinner^{20,21,22},
11 Kyle K.-H. Farh^{5*}, Tomas Marques-Bonet^{4*}, Kasper Munch^{1*}, Christian Roos^{3,24*}, Jeffrey
12 Rogers^{2*}

13
14 †These authors contributed equally.

15 *Co-Corresponding Authors:

16 Kyle K.-H. Farh, kfahr@illumina.com; Tomas Marques-Bonet, tomas.marques@upf.edu; Kasper
17 Munch, kaspermunch@birc.au.dk; Christian Roos, croos@dpz.eu; Jeffrey Rogers,
18 jr13@bcm.edu

19
20 ¹ Bioinformatics Research Centre, Aarhus University; 8000 Aarhus, Denmark

21 ² Human Genome Sequencing Center and Department of Molecular and Human Genetics, Baylor
22 College of Medicine; Houston, TX 77030, USA

- 23 ³ Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research;
24 37077 Göttingen, Germany
- 25 ⁴ Institute of Evolutionary Biology (UPF-CSIC), PRBB; 08003 Barcelona, Spain
- 26 ⁵ Artificial Intelligence Lab, Illumina Inc.; San Diego, CA 92122, USA
- 27 ⁶ Department of Biological Sciences, Louisiana State University; Baton Rouge, LA 70803, USA
- 28 ⁷ Institute for Systems Biology; Seattle, WA 98109, USA
- 29 ⁸ Department of Evolutionary Anthropology, University of Vienna; 1030 Vienna, Austria
- 30 ⁹ Human Evolution and Archaeological Sciences (HEAS), University of Vienna; 1030 Vienna,
31 Austria
- 32 ¹⁰ Department of Genetics, Human Genetics Institute of New Jersey, Rutgers University;
33 Piscataway, NJ 08854, USA
- 34 ¹¹ Department of Anthropology, New York University; New York, NY 10003, USA
- 35 ¹² Section for Ecoinformatics & Biodiversity, Department of Biology, Aarhus University; 8000
36 Aarhus C, Denmark
- 37 ¹³ Department of Neuroscience, Washington University School of Medicine; St. Louis, MO
38 63110, USA
- 39 ¹⁴ Department of Anthropology, Washington University; St. Louis, MO 63130, USA
- 40 ¹⁵ The Guinea Worm Eradication Program, The Carter Center; Addis Ababa, Ethiopia
- 41 ¹⁶ Center for Evolution and Medicine, Arizona State University; Tempe, AZ 85281, USA
- 42 ¹⁷ School of Life Sciences, Arizona State University; Tempe, AZ 85281, USA
- 43 ¹⁸ Tanzania National Parks; Arusha, Tanzania
- 44 ¹⁹ Tanzania Wildlife Research Institute; Arusha, Tanzania

45 ²⁰ Cognitive Ethology Laboratory, German Primate Center, Leibniz Institute for Primate
46 Research; 37077 Göttingen, Germany

47 ²¹ Department of Primate Cognition, Georg-August-Universität Göttingen; 37077 Göttingen,
48 Germany

49 ²² Leibniz ScienceCampus Primate Cognition; 37077 Göttingen, Germany

50 ²³ Institute of International Animal Health/One Health, Friedrich-Loeffler-Institut, Federal
51 Research Institute for Animal Health; 17493 Greifswald - Insel Riems, Germany

52 ²⁴ Gene Bank of Primates, German Primate Center, Leibniz Institute for Primate Research;
53 37077 Göttingen, Germany

54

55 **Abstract:**

56 Baboons (genus *Papio*) are a morphologically and behaviorally diverse clade of catarrhine
57 monkeys that have experienced hybridization between phenotypically and genetically distinct
58 phylogenetic species. We used high coverage whole genome sequences from 225 wild baboons
59 representing 19 geographic localities to investigate population genomics and inter-species gene
60 flow. Our analyses provide an expanded picture of evolutionary reticulation among species and
61 reveal novel patterns of population structure within and among species, including differential
62 admixture among conspecific populations. We describe the first example of a baboon population
63 with a genetic composition that is derived from three distinct lineages. The results reveal
64 processes, both ancient and recent, that produced the observed mismatch between phylogenetic
65 relationships based on matrilineal, patrilineal, and biparental inheritance. We also identified
66 several candidate genes that may contribute to species-specific phenotypes.

67

68 **One-Sentence Summary:**

69 Genomic data for 225 baboons reveal novel sites of inter-species gene flow and local effects due

70 to differences in admixture.

71 **Main Text:**

72 Our understanding of the evolutionary processes involved in the origin of biological
73 diversity has changed significantly over the past two decades. Genetic analyses have
74 demonstrated that hybridization and inter-species gene flow between closely related mammalian
75 species occur more often than previously assumed (1, 2). Traditional studies of natural
76 hybridization among populations and species have relied on phenotypic variation and a few
77 informative genetic markers (3, 4). However, access to large-scale genomic datasets now allows
78 more extensive analyses (5–7) demonstrating that in some cases complex reticulations rather
79 than dichotomously branching phylogenetic trees more accurately represent evolutionary
80 histories.

81 Among primates, including humans, the number of genera found to exhibit complex
82 histories of interspecific reticulation has recently grown considerably (2, 8–12). Baboons (genus
83 *Papio*) have long been recognized as a prime example of inter-species gene flow with several
84 hybrid zones between the six currently recognized parapatric species (Guinea baboons *P. papio*;
85 hamadryas baboons *P. hamadryas*; olive baboons *P. anubis*; yellow baboons *P. cynocephalus*;
86 Kinda baboons *P. kindae*; chacma baboons *P. ursinus*; Fig. 1; for the rationale behind the
87 classification of these major forms as species rather than subspecies see (13)) (14–17). Previous
88 analyses have identified substantial discrepancies in species-level phylogenies based on nuclear
89 DNA, mitochondrial DNA (mtDNA), and phenotypes, indicating para- and polyphyletic
90 relationships, and suggesting a complex history of differentiation and admixture (18–21). Recent
91 comparisons of whole-genome sequence (WGS) data across *Papio* species illustrated the extent
92 of genetic exchange between phenotypically distinct species (22–25). These studies were,

93 however, restricted to one or two populations per species, and therefore unable to analyze wider
94 geographic patterns of genetic diversity or compare the local effects of interspecific contact.

95 This study provides the first detailed WGS-based analysis of coancestry and genomic
96 exchange across all six baboon species, including multiple populations within olive and yellow
97 baboons. We generated deep (>30x; table S1, (13)) WGS data from 225 wild baboons
98 representing 19 localities (Fig. 1, table S2), describing variation within and among localities for
99 autosomes, X- and Y-chromosomes, mtDNA, and other genetic features such as insertions of *Alu*
100 repeats and long interspersed elements (LINEs). In addition to population structure using
101 autosomal single nucleotide variants (SNVs) and repetitive elements, we contrast coancestry
102 inferred from autosomal and X-chromosomal data to reveal sex-biased effects on genetic
103 population structure. Our results provide the most extensive analysis of genetic diversity in
104 baboons to date and reveal processes, both ancient and recent, that produced the observed
105 mismatch between phylogenetic relationships based on matrilineal, patrilineal, and biparental
106 inheritance. The evidence indicates the radiation that produced the six extant species began more
107 than one million years ago. The lineages that diverged around that time have since experienced
108 extensive admixture, as reflected in their current genetic composition. We suggest that these
109 findings inform predictions for similar systems such as hominin and early human evolution, for
110 which baboons have long been recognized as a model (26–29).

111

112 RESULTS

113 WGS analysis across multiple populations of baboons provides a fine-grained picture of
114 present-day population structure and the evolutionary history that generated it. Results of this
115 analysis also document additional locations of ongoing admixture among genetically distinct

116 lineages. Our analyses of SNVs strongly support the existence of differentiated clades including
117 the six recognized species, despite well-known hybrid zones between parapatric species. The
118 initial divergence of evolutionary lineages separates the three northern species (hamadryas, olive,
119 and Guinea baboons) from the three southern species (Kinda, yellow, and chacma baboons).
120 Analyses of population structure (Fig. 2, A to C, figs. S1 to S4) and phylogenomic maximum-
121 likelihood (ML) trees using autosomal, X- and Y-chromosomal, and mtDNA data (figs. S5 to
122 S8) are consistent with the initial north-south split, and with greater overall divergence among
123 southern than northern baboons (see also (23)). Principal component analyses (PCAs) and ML
124 trees of autosomal and X-chromosomal data separate the western Tanzanian yellow baboons
125 located at Mahale and Katavi into their own cluster distinct from eastern Tanzanian yellow
126 baboons from Mikumi, Selous, Ruaha and Udzungwa as well as from Kinda baboons. However,
127 the Y-chromosomal phylogenies, including one based on *Alu* insertions (fig. S9), show six main
128 clusters largely corresponding to the six species and place most western yellow baboons with
129 Kinda baboons. Other western yellow baboons cluster in that analysis with eastern yellow and
130 one olive baboon, providing a clear example of admixture processes not revealed by the whole-
131 genome phylogeny.

132 Across the genome of each individual, we identified the most recent coancestry among all
133 other sampled individuals (ChromoPainter (30)). The corresponding first two principal
134 components (Fig. 2C) show extensive variation among yellow baboons and confirm the primary
135 north-south split. This north-south split is also apparent in the clustering using fineSTRUCTURE
136 (30) (Fig. 2B). ML trees for autosomes and X- and Y-chromosomes (figs. S5 to S7) all support
137 the conclusions based on PCA, with two individuals falling outside their expected species clades
138 (samples PD0266 and PD0662, also anomalous in the PCAs; figs. S1, S2, S10 and S11, (13)). As

139 discussed below, the Y-chromosomal phylogeny places Kinda baboons basal to all others (fig.
140 S7).

141 Unsupervised cluster algorithms group individuals largely by species (see ADMIXTURE
142 analysis; Fig. 2D, fig. S12) with $K = 7$ as the preferred number of clusters. However, in species
143 for which we sampled more than one population (olive and yellow baboons), we find local
144 genetic differences and evidence for a complex evolutionary history (detailed discussion below).
145 These results are also supported by an analysis of LINE-1 (L1) insertions (fig. S13), an
146 independent class of genetic marker that is less prone to parallel mutations. The pelage
147 phenotypes on which taxonomy was traditionally based are generally very consistent within
148 species over wide geographic ranges (31). Yet, we find high genomic variation within and
149 among conspecific populations. Heterozygosity ranges from 0.0006 to 0.0026 (average 0.0018)
150 per base pair across the six species, and from 0.0006 to 0.0029 across the 19 localities, with the
151 lowest values in Guinea baboons (table S3, figs. S14 to S17). The coancestry matrix and its PCA
152 (Fig. 2, B and C) differentiates the various sampling localities and is therefore consistent with the
153 ADMIXTURE analysis (Fig. 2D) showing that the sampled populations within both yellow and
154 olive baboons can be distinguished genetically. The yellow baboons in Mikumi (Fig. 2B, box H)
155 share pelage and morphological phenotypes with those in Ruaha, although they are genetically
156 distinct. Western yellow baboons from Mahale and Katavi (Fig. 2B, box F) exhibit phenotypic
157 traits (somewhat smaller body size than Mikumi baboons, especially cranial metrics; aspects of
158 coat color with some individuals having pink skin around the eyes, and sporadic occurrence of
159 white-furred infants) in which they resemble Kinda baboons (32). The coancestry matrix (Fig.
160 2B) further shows that yellow baboons from Mahale and Katavi (box F) exhibit greater genetic
161 similarity with Kinda (box E) and chacma baboons (box G) than with their supposed

162 conspecifics from eastern Tanzania (box H). Similarly, all olive baboons (with exception of
163 those from Tarangire) share a very consistent pelage and external phenotype. However,
164 ADMIXTURE (Fig. 2D) and ChromoPainter (Fig. 2B) analyses identify clear evidence of
165 genetic differences between the Ethiopian Gog olive baboons and the Tanzanian olive baboons
166 of Lake Manyara and Ngorongoro. Furthermore, the Serengeti population is more similar
167 genetically to both the Gombe and Aberdare populations than to the Ngorongoro or Lake
168 Manyara populations which are geographically much closer.

169 We used the SNV data to reconstruct the history of population size for each baboon
170 locality (Fig. 3A, figs. S18 to S21). The estimated effective population sizes (N_e) are all
171 essentially the same and on the order of 100,000 until about 1.0-1.2 million years ago, which is
172 consistent with the prior dating of the initial north-south divergence (23). At the separation, the
173 N_e of northern populations fell below that of the southern populations, supporting the idea that
174 the genus arose in southern Africa, and a daughter population from this basal stock spread to the
175 north, then to the west, losing genetic diversity in serial founding events. The suggestion that
176 Guinea baboons represent the descendants of those groups that were at the leading edge of that
177 dispersal for the longest distance and time (33) is supported by the lower heterozygosity in that
178 sample relative to all other baboon species (table S3). Also, whole-genome *Alu* and L1 insertion-
179 based phylogenies place western yellow baboons with Kinda baboons, while Guinea baboons are
180 basal among baboons, and hamadryas baboons are the sister taxon to olive and southern baboons
181 (figs. S22 and S23). These findings may result from Guinea baboons, and to a lesser extent
182 hamadryas baboons, losing polymorphic derived *Alu* and L1 insertions through drift as they
183 dispersed north from the southern geographic origin (34).

184 Earlier studies provided clear evidence for hybridization and gene flow across the contact
185 zones between pairs of parapatric species (15–17, 24, 25, 35). In this study, we present new
186 evidence for additional ancient and recent arenas for gene flow between species pairs. Species
187 tree reconstruction (ASTRAL (36)) using window-based ML trees (50kb and 500kb window
188 size) produced inconsistent branching patterns among datasets and only 58-70% of gene trees fit
189 the species tree at the quartet level (figs. S24 and S25). Both incomplete lineage sorting (ILS)
190 and gene flow are likely contributing to this discordance which is expected to be larger for
191 smaller windows. In addition, a qualitative visualization of these trees (figs. S24 and S25) shows
192 a network-like pattern, again indicating complexity. There is greater shared genetic drift
193 (measured by f_3 outgroup statistics) among eastern yellow baboon localities (Udzungwa, Selous,
194 Mikumi, Ruaha), while western yellow baboons tend to cluster with Kinda baboons (fig. S26). In
195 admixture graphs (Fig. 3B), Kinda baboons are, similarly to the description in (23), represented
196 as a fusion product of populations from southern and ancestral northern clades, while the western
197 yellow baboons share ancestry with both Kinda and olive baboons. More complex graphs (tables
198 S4 and S5, figs. S27 to S29) might be supported, but failed to give replicable results, likely due
199 to complex reticulation and multiple gene flow events at different times and between different
200 local populations which now obscure the processes involved.

201 Taken as a whole, this expanded dataset does not support the previous suggestion that
202 Kinda baboons result from a recent fusion event (23) as shown in Fig. 3B. In PCA plots using
203 genome wide SNVs, Kinda baboons do not fall intermediate between northern and southern
204 clades but in fact are quite distinct (Fig. 2A, figs. S1 and S2). Some ML trees (i.e., Y-
205 chromosome data; fig. S7) place Kinda baboons as sister clade to all other baboons whereas
206 other trees (autosomes and X-chromosome data, figs. S5 and S6) lump them together with

207 yellow and chacma baboons into the southern clade. These results are more consistent with the
208 idea that Kinda baboons show substantial genetic similarity to both northern and southern clade
209 baboons because they are basal and phenotypically resemble the ancestral form from which all
210 extant species are derived. Fossil evidence suggests a southern African origin for baboons (34),
211 and the mtDNA haplotypes of Kinda and western yellow baboons (Fig. 4, fig. S8 (21)) suggest
212 their range in tropical southern Africa may include the area of origin of both northern and
213 southern primary branches. Broader aspects of Y-chromosome data also do not support Kinda
214 baboons as a fusion product; Kinda baboon Y-haplotypes are found in western yellow baboons
215 but not in olive baboons, and no olive baboon mtDNA has been observed in any Kinda baboon to
216 date. Finally, Kinda baboons share more polymorphic *Alu* insertions with geladas than do other
217 *Papio* species, possibly the result of a period of co-existence and hybridization between their
218 ancestors (37).

219 We analyzed the genetic relationships among the eight major regional baboon
220 populations that constitute our samples, i.e., the four single-locality populations of chacma,
221 Kinda, hamadryas and Guinea baboons as well as two groups each of yellow (western and
222 eastern) and olive (Gog and southern) baboons. By modeling the recent ancestry along the
223 chromosomes of individual baboons (Globetrotter (38)) we can represent each group as a
224 mixture of recent ancestry with the remaining seven groups (Fig. 3C). In most of the groups, we
225 can identify a contribution from recent admixture events (the oldest identifiable event estimated
226 at 56 generations; table S6) separate from contributions of older admixture and retention of
227 ancestral polymorphism (bootstrap p-values < 0.01 unless otherwise noted). In Fig. 3, C and D,
228 we distinguish the recent admixture from more ancient shared ancestry by showing the recent
229 admixture estimates as expanded (exploded) wedges.

230 We identified a large amount of shared ancestry between southern olive and eastern
231 yellow baboons not concordant with the overall phylogeny (Fig. 3C). This is also expressed in
232 the coancestry matrix (Fig. 2B, box X) and is additional evidence of persistent admixture
233 between both species (15, 17, 22, 25). Furthermore, western yellow baboons from Mahale and
234 Katavi share substantial ancestry with eastern yellow, Kinda, and southern olive baboons. This
235 cannot be explained as a retention of ancient shared variation present prior to the origin of the six
236 major branches, because there is no equivalent sharing with chacma, hamadryas or Guinea
237 baboons. This is, therefore, the first evidence that a single population (western yellow baboons)
238 contains measurable admixture contributions from more than two distinct lineages. Comparing
239 the ancestry of recently admixing populations (expanded wedges in Fig. 3C) to that of each other
240 group identifies recent admixture from Gog into southern olive baboons, between western and
241 eastern yellow baboons, from southern olive baboons into eastern yellow baboons (p-value 0.04),
242 between Kinda and chacma baboons (p-value 0.02), and between Kinda and western yellow
243 baboons. Repeating the Globetrotter analysis assuming 14 populations representing all major
244 sampling locations differentiates olive and yellow baboon populations (Fig. 3D) and reveals a
245 complex system of recent gene flow (all events < 95 generations) between: i) olive baboon
246 populations, ii) yellow baboon populations, iii) yellow and Tarangire olive baboons, iv) western
247 yellow and Gombe olive baboons, and v) Tarangire olive baboons and Ruaha yellow baboons.
248 These results do not imply direct migration of males (e.g., individual males moving from Gog to
249 Serengeti), but more plausibly the overall consequences of many incremental gene flow events
250 distributing alleles long distances over multiple generations.

251 This is not the first study to suggest that the history of genetic differentiation and
252 reticulation among baboons is complex. Previous studies (10, 18–21, 33, 39, 40) showing

253 widespread phenotype-mitochondrial discordance strongly suggest that nuclear swamping (i.e.
254 the immigration of males into a phenotypically different population, largely or completely
255 displacing the nuclear DNA composition and phenotype of the invaded population, without
256 changing its mtDNA composition) has been a major contributing process. The present study
257 found a similar discordance between the expanded mtDNA phylogeny (Fig. 4, fig. S8) on the one
258 hand and the new autosomal and Y-chromosomal phylogenies on the other (figs. S5 and S7).
259 Thus, our WGS findings strongly support previous suggestions based only on mtDNA and
260 phenotype data that nuclear swamping has been a major factor generating the current pattern of
261 baboon genetic and phenotypic variation.

262 The dense sampling of mtDNA provides important information about matrilineal
263 ancestry. However, as a single locus, mtDNA represents only one of many possible genealogies
264 generated by ILS and admixture. To test the hypothesis that nuclear swamping produced the
265 discord observed between mtDNA phylogenies and relationships based on phenotype, we
266 contrasted ancestry proportions across the X-chromosome and the similar-sized chromosome 8,
267 each contributing thousands of individual genealogies. Admixture by hemizygous males
268 introduces disproportionately more autosomal than X-chromosomal sequence, rendering shared
269 X-chromosome ancestry a better representation of deep species relationships prior to admixture.
270 We found that the X-chromosome of our chacma baboons derives more ancestry from yellow
271 baboons than their chromosome 8 does (0.47 vs 0.62, paired t-test value 0.005; Fig. 5A),
272 suggesting that male-biased admixture from the ancestors of chacma baboons into the southern
273 range of yellow baboons produced northern chacma baboons, including the grayfooted chacma
274 baboons (*P. ursinus grisiepes*) that we analyze here. This observation is consistent with the close
275 relationship between mtDNA found in southern-most yellow and northern chacma baboons

276 (clade B in Fig. 4 (19, 40)). The most compelling evidence of male-biased admixture is the
277 relationship between western yellow and Kinda baboons. The ancestry profile of western yellow
278 baboons (Fig. 5B) is very different from eastern yellow baboons (Fig. 5C). Western yellow
279 baboons share more ancestry with Kinda baboons on the X-chromosome than on chromosome 8
280 (0.27 vs 0.44, paired t-test p-value 0.025) while Kinda baboons contain twice as much western
281 yellow baboon ancestry on the X-chromosome as on chromosome 8 (0.23 vs 0.55, paired t-test p-
282 value $1.8e-13$; Fig. 5D). Furthermore, eastern yellow baboons share more X-chromosomal
283 ancestry with western yellow baboons than chromosome 8 ancestry (0.16 vs 0.20, paired t-test p-
284 value $3.1e-9$; Fig. 5B). Together these observations indicate that western yellow baboons were
285 produced mainly from males carrying haplotypes that originated among eastern yellow and
286 southern olive baboons migrating into the ancestral range of Kinda baboons, replacing Kinda
287 baboon autosomes more than they replaced Kinda baboon X-chromosomes. As a result, western
288 yellow baboons carry genetic input from three distinct lineages.

289 In addition to patterns of shared ancestry among populations and species, we used two
290 strategies to seek preliminary evidence for species-specific genetic adaptations in baboons. First,
291 we used PLINK (41) to identify SNVs enriched in one species relative to all others (table S8).
292 Genes containing possibly functional SNVs enriched in a given taxon were correlated with
293 species phenotypes using Gene Ontology (GO) (42) terms and literature searches. We also used
294 OmegaPlus (43) to test those gene regions for evidence of selective sweeps. Across all species,
295 1,342,371 SNVs met the criteria for being enriched in one particular species, including 4,337
296 missense and 76 stop gained SNVs (table S8). We next searched this list of candidates for genes
297 annotated as influencing known traits of that species. Among them, SNV_1 (Table 1, fig. S32), a
298 missense variant in Serine Protease 8 (*PRSS8*) has a 0.96 allele frequency (AF) in hamadryas

299 baboons and a 0.02 AF in the geographically adjacent Gog olive baboons (absent in other
300 species). *PRSS8* increases epithelial sodium channel activity and mediates sodium reabsorption
301 through the kidneys (44). *PRSS8* is under positive selection in the desert-adapted canyon mouse
302 (*Peromyscus crinitus*) (45), and hamadryas baboons inhabit the most arid environment of all
303 baboons (46). SNV_2 (Table 1, fig. S33) has a 1.0 AF in both hamadryas and Guinea baboons
304 and is absent from other species. This is a missense variant in Neurexin 1 (*NRXN1*) which is
305 associated with the GO term “social behavior”. *NRXN1* knockout mice exhibit changes in male
306 aggression (47). Guinea and hamadryas baboons differ from others in the genus in exhibiting a
307 multi-level male-philopatric social organization with substantial male-male tolerance (29, 48).
308 This contrasts with the matrilineal, male-dispersing social organization typical and likely
309 ancestral for the genus. This observation is compatible with the speculation that until "swamped"
310 by males from olive and yellow baboon populations, male-philopatric "pre-Guinea" and "pre-
311 hamadryas" baboon populations occupied the northern savanna-woodland belt and much of the
312 East African savanna-woodland corridor (33). SNV_3 (Table 1, fig. S34) has a 1.0 AF in Kinda
313 baboons and a 0.05 AF in yellow baboons (western yellow baboons and Ruaha) and one
314 Serengeti olive baboon. This is a missense variant in the pigmentation-associated Agouti
315 Signaling Protein (*ASIP*). In mice, this gene affects melanin synthesis, shifting eumelanin
316 production (black/brown hair) to pheomelanin (red/yellow hair) (49). Kinda baboons display
317 several unique coat color traits, including a substantial proportion of infants with white natal
318 coats (16).

319 In our second approach to functional variation, we searched for genomic regions of
320 elevated differentiation between pairs of closely related species (for details see (13)). We asked
321 whether regions with the strongest evidence of differentiation (windows in the top 0.1%) were

322 enriched for genes with particular GO terms. Genomic regions most distinct between Kinda and
323 yellow baboons were enriched for genes linked to skeletal development and morphogenesis (p-
324 value adjusted for false discovery rate, $p = 1.77 \times 10^{-4}$); tables S9 to S11; fig. S35) including limb
325 development (e.g., embryonic forelimb morphogenesis, adj. $p = 0.02$). This enrichment was
326 driven by one region on chromosome 3 containing a HOXA gene cluster (fig. S36) and may
327 influence the distinctively small size and gracile, long-limbed build of Kinda baboons (16).
328 Genes linked to male sexual differentiation were also increased in regions highly differentiated
329 between Kinda and yellow baboons (adj. $p = 0.0484$), possibly related to the reduced sexual
330 dimorphism in Kinda baboons (50).

331

332 **DISCUSSION**

333 Our expanded whole-genome dataset provides several novel insights into genetic
334 reticulation and the evolutionary history of multiple local populations of baboons. Previous work
335 showed that gene flow occurs among phenotypically and genetically distinct baboon species and
336 pointed to nuclear swamping as a major contributing process. Our study extends and adds higher
337 resolution to this picture, using genetic data to confirm hybrid zones that were previously
338 suspected from field observation of phenotypic variation alone. We also identify the first local
339 population (western Tanzanian yellow baboons) that has clear evidence for genetic contributions
340 from three genetically distinct lineages.

341 While our results significantly extend our knowledge of baboon evolutionary history,
342 some gaps remain. The richness of evolutionary detail to be derived from denser sampling is
343 indicated by our results from East African populations. More materials are needed to document
344 other regions with complex biogeographic and evolutionary history, including the olive-Guinea

345 baboon interface in West Africa (21), and regions of southern Africa where chacma baboons
346 have experienced both ancient and recent periods of genetic divergence and reticulation (39, 40).
347 Other geographic regions, e.g., the northern savanna-woodland belt west of our Gog population
348 have not been studied and would likely provide further information, especially regarding the
349 origins and history of olive and Guinea baboons. Nevertheless, our dense sampling in East Africa
350 clearly identifies new arenas of gene flow and documents the complexity of the evolutionary
351 history of baboons in this region.

352 Our results lead to several substantive conclusions. With regard to methods, we find that
353 while comparison of mtDNA and phenotypic variation are effective in detecting nuclear
354 swamping, analyses comparing levels of shared ancestry across the X-chromosome to that across
355 autosomes provides a more quantitative assessment of demographic processes and genetic
356 history. Second, we conclude that Kinda baboons are not the product of a recent fusion event.
357 Instead, they are more likely close to the basal ancestor of all extant baboons. Next, we find
358 additional support for the prior observation that the primary separation of northern and southern
359 baboon species is the result of dispersal from the south to the north, with Guinea baboons
360 recognized as the most recent occupants of the leading edge of that dispersal. Despite the sharp
361 gradient of phenotypes that is characteristic of baboon inter-species contact zones, gene flow
362 distributes the introgressed alleles far from the regions of obvious hybridization. Last, we report
363 that extant western yellow baboons carry genetic contributions from three genetically different
364 baboon lineages.

365 The patterns of local, regional and species-level genetic structure in baboons are likely a
366 valuable model for population structure in other primate clades that consist of multiple closely
367 related species, such as African green monkeys (genus *Chlorocebus* (51)) and macaques (genus

368 *Macaca* (52)). Clades in other mammalian orders are also revealing complex, often reticulated,
369 evolutionary histories like those of baboons (e.g., polar bears (53, 54), giraffes (7), and deer
370 (55)). The results for baboons also provide informative parallels and contrasts to the evolutionary
371 differentiation and relationships among early human ancestors that arose, differentiated and
372 admixed over a timespan remarkably similar to that of baboon cladogenesis (56).

373

374 **MATERIALS AND METHODS SUMMARY**

375 Extended materials and methods are available in the supplementary materials.

376 **Samples and DNA Sequencing**

377 Blood samples from 225 baboons and two geladas were gathered in accordance with local
378 regulations. Genomic DNA was extracted from blood and libraries were prepared for sequencing
379 on the NovaSeq 6000 platform (Illumina).

380 **Variant Calling and Phasing**

381 We used BWA-MEM to map reads to the Panu_3.0 baboon and the Mmul_10 rhesus assemblies.
382 GATK was used to call variants following best practices. Panu_3.0 SNVs were phased using
383 WhatsHap and SHAPEIT.

384 **Population Structure and Phylogenetic Analyses**

385 Population structure based on SNVs was examined using PCA, ADMIXTURE, and
386 fastSTRUCTURE. Phylogenetic trees based on autosomal and sex chromosome SNVs and
387 Geneious assembled mitochondrial genomes were generated using IQ-TREE and visualized with
388 FigTree. Polymorphic mobile elements were identified using DELLY and MELT. STRUCTURE
389 and MELT were used to analyze population structure of L1 and *Alu* elements. PAUP was used to
390 generate maximum parsimony trees from *Alu* and L1 elements. We used MSMC2 to infer

391 baboon demographic history and population structure through time. Admixture graphs and F3
392 outgroup statistics were generated using ADMIXTOOLS 2.

393 **Inference of most recent coancestry along each chromosome**

394 ChromoPainter was used to infer the most recent coancestry along chromosomes and
395 fineSTRUCTURE was used to identify relationships between individuals based on their most
396 recent coancestry. We used Globetrotter to compute p-values for a coancestry contribution from
397 recent admixture.

398 **Functional Variation**

399 Functional variation was examined using PLINK for association analyses, OmegaPlus for
400 selective sweep identification, and differentiation-based scans for selection using windowed F_{ST}
401 values.

402

403 **REFERENCES AND NOTES**

- 404 1. M. L. Arnold, Y. Sapir, N. H. Martin, Genetic exchange and the origin of adaptations:
405 prokaryotes to primates. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **363**, 2813-2820 (2008).
406 doi: 10.1098/rstb.2008.0021; pmid: 18522920
- 407 2. R. R. Ackermann, M. L. Arnold, M. D. Baiz, J. Cahill, L. Cortes-Ortiz, B. J. Evans, B. R.
408 Grant, P. R. Grant, B. Hallgrimsson, R. Humphreys, C. J. Jolly, J. Malukiewicz, C. J.
409 Percival, T. Ritzman, C. Roos, C. C. Roseman, L. Schroeder, F. Smith, K. Warren, R.
410 Wayne, D. Zinner, Hybridization in human evolution: insights from other organisms.
411 *Evol. Anthropol.* **28**, 189-209 (2019). doi: 10.1002/evan.21787; pmid: 31222847
- 412 3. N. H. Barton, G. M. Hewitt, Analysis of hybrid zones. *Ann. Rev. Ecol. Syst.* **16**, 113-148
413 (1985). doi: 10.1146/annurev.es.16.110185.000553

- 414 4. T. E. Dowling, C. L. Secor, The role of hybridization and introgression in the
415 diversification of animals. *Ann. Rev. Ecol. Syst.* **28**, 598-619 (1997). doi:
416 10.1146/annurev.ecolsys.28.1.593
- 417 5. M. de Manuel, R. Barnett, M. Sandoval-Velasco, N. Yamaguchi, F. Garrett Vieira, M. L.
418 Zepeda Mendoza, S. Liu, M. D. Martin, M. H. S. Sinding, S. S. T. Mak, C. Carøe, S. Liu,
419 C. Guo, J. Zheng, G. Zazula, G. Baryshnikov, E. Eizirik, K. P. Koepfli, W. E. Johnson,
420 A. Antunes, T. Sicheritz-Ponten, S. Gopalakrishnan, G. Larson, H. Yang, S. J. O'Brien,
421 A. J. Hansen, G. Zhang, T. Marques-Bonet, M. Gilbert, The evolutionary history of
422 extinct and living lions. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 10927-10934 (2020). doi:
423 10.1073/pnas.1919423117; pmid: 32366643
- 424 6. S. Lamichhaney, F. Han, M. T. Webster, B. R. Grant, P. R. Grant, L. Andersson, Female-
425 biased gene flow between two species of Darwin's finches. *Nat. Ecol. Evol.* **4**, 979-986
426 (2020). doi: 10.1038/s41559-020-1183-9; pmid: 32367030
- 427 7. R. T. F. Coimbra, S. Winter, V. Kumar, K. P. Koepfli, R. M. Gooley, P. Dobrynin, J.
428 Fennessy, A. Janke, Whole-genome analysis of giraffe supports four distinct species.
429 *Curr. Biol.* **31**, 2929-2938.e5 (2021). doi: 10.1016/j.cub.2021.04.033; pmid: 33957077
- 430 8. M. L. Arnold, A. Meyer, Natural hybridization in primates: one evolutionary mechanism.
431 *Zoology* **109**, 261-276 (2006). doi: 10.1016/j.zool.2006.03.006; pmid: 16945512
- 432 9. R. E. Green, J. Krause, A. W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson,
433 H. Li, W. Zhai, M. H.-Y. Fritz, N. F. Hansen, E. Y. Durand, A.-S. Malaspinas, J. D.
434 Jensen, T. Marques-Bonet, C. Alkan, K. Prüfer, M. Meyer, H. A. Burbano, J. M. Good,
435 R. Schultz, A. Aximu-Petri, A. Butthof, B. Höber, B. Höffner, M. Siegemund, A.
436 Weihmann, C. Nusbaum, E. S. Lander, C. Russ, N. Novod, J. Affourtit, M. Egholm, C.

- 437 Verna, P. Rudan, D. Brajkovic, Ž. Kucan, I. Gušić, V. B. Doronichev, L. V Golovanova,
438 C. Lalueza-Fox, M. de la Rasilla, J. Fortea, A. Rosas, R. W. Schmitz, P. L. F. Johnson, E.
439 E. Eichler, D. Falush, E. Birney, J. C. Mullikin, M. Slatkin, R. Nielsen, J. Kelso, M.
440 Lachmann, D. Reich, S. Pääbo, A draft sequence of the Neandertal genome. *Science* **328**,
441 710-722 (2010). doi: 10.1126/science.1188021; pmid: 20448178
- 442 10. D. Zinner, M. L. Arnold, C. Roos, The strange blood: natural hybridization in primates.
443 *Evol. Anthropol.* **20**, 96-103 (2011). doi: 10.1002/evan.20301; pmid: 22034167
- 444 11. J. Tung, L. B. Barreiro, The contribution of admixture to primate evolution. *Curr. Opin.*
445 *Genet. Dev.* **47**, 61-68 (2017). doi: 10.1016/j.gde.2017.08.010; pmid: 28923540
- 446 12. C. Fontserè, M. de Manuel, T. Marques-Bonet, M. Kuhlwilm, Admixture in mammals
447 and how to understand its functional implications: on the abundance of gene flow in
448 mammalian species, its impact on the genome, and roads into a functional understanding.
449 *BioEssays* **41**, 1900123 (2019). doi: 10.1002/bies.201900123; pmid: 31664727
- 450 13. Detailed information is provided in the supplementary materials.
- 451 14. U. Nagel, A comparison of anubis baboons, hamadryas baboons and their hybrids at a
452 species border in Ethiopia. *Folia Primatol.* **19**, 104-165 (1973). doi: 10.1159/000155536;
453 pmid: 4201907
- 454 15. J. Tung, M. J. E. Charpentier, D. A. Garfield, J. Altmann, S. C. Alberts, Genetic evidence
455 reveals temporal change in hybridization patterns in a wild baboon population. *Mol. Ecol.*
456 **17**, 1998-2011 (2008). doi: 10.1111/j.1365-294X.2008.03723.x; pmid: 18363664
- 457 16. C. J. Jolly, A. S. Burrell, J. E. Phillips-Conroy, C. M. Bergey, J. Rogers, Kinda baboons
458 (*Papio kindae*) and grayfoot chacma baboons (*P. ursinus griseipes*) hybridize in the

- 459 Kafue river valley, Zambia. *Am. J. Primatol.* **73**, 291-303 (2011). doi: 10.1002/ajp.20896;
460 pmid: 21274900
- 461 17. M. J. E. Charpentier, M. C. Fontaine, E. Cherel, J. P. Renoult, T. Jenkins, L. Benoit, N.
462 Barthès, S. C. Alberts, J. Tung, Genetic structure in a dynamic baboon hybrid zone
463 corroborates behavioural observations in a hybrid population. *Mol. Ecol.* **21**, 715-731
464 (2012). doi: 10.1111/j.1365-294X.2011.05302.x; pmid: 21988698
- 465 18. D. E. Wildman, T. J. Bergman, A. al-Aghbari, K. N. Sterner, T. K. Newman, J. E.
466 Phillips-Conroy, C. J. Jolly, T. R. Disotell, Mitochondrial evidence for the origin of
467 hamadryas baboons. *Mol. Phylogenet. Evol.* **32**, 287-296 (2004). doi:
468 10.1016/j.ympev.2003.12.014; pmid: 15186814
- 469 19. D. Zinner, L. F. Groeneveld, C. Keller, C. Roos, Mitochondrial phylogeography of
470 baboons (*Papio* spp.) - indication for introgressive hybridization? *BMC Evol. Biol.* **9**, 83
471 (2009). doi: 10.1186/1471-2148-9-83; pmid: 19389236
- 472 20. D. Zinner, J. Wertheimer, R. Liedigk L. F. Groeneveld, C. Roos, Baboon phylogeny as
473 inferred from complete mitochondrial genomes. *Am. J. Phys. Anthropol.* **150**, 133-140
474 (2013). doi: 10.1002/ajpa.22185; pmid: 23180628
- 475 21. C. Roos, S. Knauf, I. S. Chuma, A. Maille, C. Callou, R. Sabin, R. Portela Miguez, D.
476 Zinner, New mitogenomic lineages in *Papio* baboons and their phylogeographic
477 implications. *Am. J. Phys. Anthropol.* **174**, 407-417 (2021). doi: 10.1002/ajpa.24186;
478 pmid: 33244782
- 479 22. J. D. Wall, S. S. Schlebusch, S. C. Alberts, L. A. Cox, N. Snyder-Mackler, K. Nevenon,
480 L. Carbone, J. Tung, Genome-wide ancestry and divergence patterns from low-coverage

- 481 sequencing data reveal a complex history of admixture in wild baboons. *Mol. Ecol.* **25**,
482 3469-3483 (2016). doi: 10.1111/mec.13684; pmid: 27145036
- 483 23. J. Rogers, M. Raveendran, R. A. Harris, T. Mailund, K. Leppälä, G. Athanasiadis, M. H.
484 Schierup, J. Cheng, K. Munch, J. A. Walker, M. K. Konkel, V. E. Jordan, C. J. Steely, T.
485 O. Beckstrom, C. Bergey, A. Burrell, D. Schrempf, A. Noll, M. Kothe, G. H. Kopp, S.
486 Murali, K. Billis, F. J. Martin, M. Muffato, L. A. Cox, J. Else, T. Disotell, D. M. Muzny,
487 J. Phillips-Conroy, B. Aken, E. E. Eichler, M. W. Hahn, T. Marques-Bonet, C. Kosiol,
488 M. A. Batzer, J. Tung, D. Zinner, C. Roos, C. J. Jolly, R. A. Gibbs, K. C. Worley,
489 Baboon Genome Analysis Consortium, The comparative genomics and complex
490 population history of *Papio* baboons. *Sci. Adv.* **5**, eaau6947 (2019). doi:
491 10.1126/sciadv.aau6947; pmid: 30854422
- 492 24. K. L. Chiou, C. M. Bergey, A. S. Burrell, T. R. Disotell, J. Rogers, C. J. Jolly, J. E.
493 Phillips-Conroy, Genome-wide ancestry and introgression in a Zambian baboon hybrid
494 zone. *Mol. Ecol.* **30**, 1907-1920 (2021). doi: 10.1111/mec.15858; pmid: 33624366
- 495 25. T. P. Vilgalys, A. S. Fogel, J. A. Anderson, R. S. Mututua, J. K. Warutere, I. L. Siodi, S.
496 Y. Kim, T. N. Voyles, J. A. Robinson, J. D. Wall, E. A. Archie, S. C. Alberts, J. Tung,
497 Selection against admixture and gene regulatory divergence in a long-term primate field
498 study. *Science* **377**, 635-641 (2022). doi: 10.1126/science.abm.4917; pmid: 35926022
- 499 26. C. J. Jolly, A proper study for mankind: analogies from the papionin monkeys and their
500 implications for human evolution. *Am. J. Phys. Anthropol.* **Suppl. 33**, 177-204 (2001).
501 doi: 10.1002/ajpa.10021; pmid: 11786995

- 502 27. S. Elton, Forty years on and still going strong: The use of hominin-cercopithecoid
503 comparisons in palaeoanthropology. *J. Roy. Anthropol. Inst.* **12**, 19-38 (2006). doi:
504 10.1111/j.1467-9655.2006.00279.x
- 505 28. C. J. Jolly, “Analogies and models in the study of the early hominins” in *Early Hominin*
506 *Paleoecology*, M. Sponheimer, J. A. Lee-Thorp, K. E. Reed, P. S. Ungar, Eds. (Colorado
507 Univ. Press, 2013), pp. 437-455.
- 508 29. J. Fischer, J. P. Higham, S. C. Alberts, L. Barrett, J. C. Beehner, T. J. Bergman, A. J.
509 Carter, A. Collins, S. Elton, J. Fagot, M. J. Ferreira da Silva, K. Hammerschmidt, P.
510 Henzi, C. J. Jolly, S. Knauf, G. H. Kopp, J. Rogers, C. Roos, C. Ross, R. M. Seyfarth, J.
511 B. Silk, N. Snyder-Mackler, V. Staedele, L. Swedell, M. L. Wilson, D. Zinner, Insights
512 into the evolution of social systems and species from baboon studies. *eLife* **8**, e50989
513 (2019). doi: 10.7554/eLife.50989; pmid: 31711570
- 514 30. D. J. Lawson, G. Hellenthal, S. Myers, D. Falush, Inference of population structure using
515 dense haplotype data. *PLoS Genet.* **8**, e1002453 (2012). doi:
516 10.1371/journal.pgen.1002453; pmid: 22291602
- 517 31. C. J. Jolly, “Species, subspecies, and baboon systematics” in *Species, Species Concepts,*
518 *and Primate Evolution*, W. H. Kimbel, L. B. Martin, Eds. (Plenum, 1993), pp. 67-101.
- 519 32. M. V. Anandam, E. L. Bennett, T. R. B. Davenport, N. J. Davies, K. M. Detwiler, A.
520 Engelhardt, A. A. Eudey, E. L. Gadsby, C. P. Groves, A. Healy, K. P. Karanth, S. Molur,
521 T. Nadler, M. C. Richardson, E. P. Riley, C. Roos, A. B. Rylands, L. K. Sheeran, N.
522 Ting, J. Wallis, S. S. Waters, D. J. Whittaker, D. Zinner, “Family Cercopithecidae (Old
523 World monkeys) – species accounts of Cercopithecidae” in *Handbook of the Mammals of*

- 524 *the World, Vol. 3 Primates*, R. A. Mittermeier, A. B. Rylands, D. E. Wilson, Eds. (Lynx,
525 2013), pp. 628-753.
- 526 33. C. J. Jolly, Philopatry at the frontier: a demographically driven scenario for the evolution
527 of multilevel societies in baboons (*Papio*). *J. Hum. Evol.* **146**, 102819 (2020). doi:
528 10.1016/j.hevol.2020.102819; pmid: 32736063
- 529 34. C. Gilbert, S. Frost, K. Pugh, M. Anderson, E. Delson, Evolution of the modern baboon
530 (*Papio hamadryas*): a reassessment of the African Plio-Pleistocene record. *J. Hum. Evol.*
531 **122**, 38-69. (2018). doi: 10.1016/j.hevol.2018.04.012; pmid: 29954592
- 532 35. T. J. Bergman, J. E. Phillips-Conroy, C. J. Jolly, Behavioral variation and reproductive
533 success of male baboons (*Papio anubis* × *Papio hamadryas*) in a hybrid social group.
534 *Am. J. Primatol.* **70**, 136-147 (2008). doi: 10.1002/ajp.20467; pmid: 17724672
- 535 36. C. Zhang, M. Rabiee, E. Sayyari, S. Mirarab, ASTRAL-III: polynomial time species tree
536 reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19** (suppl. 6), 153
537 (2018). doi: 10.1186/s12859-018-2129-y; pmid: 29745866
- 538 37. J. A. Walker, V. E. Jordan, J. M. Storer, C. J. Steely, P. Gonzalez-Quiroga, T. O.
539 Beckstrom, L. C. Rewerts, C. P. St. Romain, C. E. Rockwell, J. Rogers, C. J. Jolly, M. K.
540 Konkel, The Baboon Genome Analysis Consortium, M. A. Batzer, *Alu* insertion
541 polymorphisms shared by *Papio* baboons and *Theropithecus gelada* reveal an intertwined
542 common ancestry. *Mobile DNA* **10**, 46 (2019). doi: 10.1186/s13100-019-0187-y; pmid:
543 31788036
- 544 38. G. Hellenthal, G. B. J. Busby, G. Band, J. F. Wilson, C. Capelli, D. Falush, S. Myers, A
545 genetic atlas of human admixture history. *Science* **343**, 747-751 (2014). doi:
546 10.1126/science.1243518; pmid: 24531965

- 547 39. R. Sithaldeen, J. M. Bishop, R. R. Ackermann, Mitochondrial DNA analysis reveals Plio-
548 Pleistocene diversification within the chacma baboon. *Mol. Phylogenet. Evol.* **53**, 1042-
549 1049 (2009). doi: 10.1016/j.ympev.2009.07.038; pmid: 19665055
- 550 40. C. Keller, C. Roos, L. F. Groeneveld, J. Fischer, D. Zinner, Introgressive hybridization in
551 southern African baboons shapes patterns of mtDNA variation. *Am. J. Phys. Anthropol.*
552 **142**, 125-136 (2010). doi: 10.1002/ajpa.21209; pmid: 19918986
- 553 41. S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bendler, J.
554 Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, P. C. Sham, PLINK: a tool set for whole-
555 genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559-
556 575 (2007). doi: 10.1086/519795; pmid: 17701901
- 557 42. Gene Ontology Consortium, The Gene Ontology resource: enriching a GOld mine.
558 *Nucleic Acids Res.* **49**, D325-D334 (2021). doi: 10.1093/nar/gkaa1113; pmid: 33290552
- 559 43. N. Alachiotis, A. Stamatakis, P. Pavlidis, OmegaPlus: a scalable tool for rapid detection
560 of selective sweeps in whole-genome datasets. *Bioinformatics* **28**, 2274-2275 (2012). doi:
561 10.1093/bioinformatics/bts419; pmid: 22760304
- 562 44. T. Narikiyo, K. Kitamura, M. Adachi, T. Miyoski, K. Iwashita, N. Shiraishi, H.
563 Nonoguchi, L.-M. Chen, K. X. Chai, J. Chao, K. Tomita, Regulation of prostaticin by
564 aldosterone in the kidney. *J. Clin. Invest.* **109**, 401-408 (2002). doi: 10.1172/JCI13229;
565 pmid: 11828000
- 566 45. J. P. Colella, A. Tigano, O. Dudchenko, A. D. Omer, R. Khan, I. D. Bochkov, E. L.
567 Aiden, M. D. MacManes, Limited evidence for parallel evolution among desert-adapted
568 *Peromyscus* deer mice. *J. Hered.* **112**, 286-302 (2021). doi: 10.1093/jhered/esab009;
569 pmid: 33686424

- 570 46. D. Zinner, M. Klapproth, A. Schell, L. Ohrndorf, D. Chala, J. U. Ganzhorn, J. Fischer,
571 Comparative ecology of Guinea baboons (*Papio papio*). *Primate Biol.* **8**, 19-35 (2021).
572 doi: 10.5194/pb-8-19-2021; pmid: 34109265
- 573 47. H. M. Grayton, M. Missler, D. A. Collier, C. Fernandes, Altered social behaviours in
574 neurexin 1 α knockout mice resemble core symptoms in neurodevelopmental disorders.
575 *PLoS One* **8**, e67114 (2013). doi: 10.1371/journal.pone.0067114; pmid: 23840597
- 576 48. L. Swedell, J. Saunders, A. Schreier, B. Davis, T. Tesfaye, M. Pines, Female “dispersal”
577 in hamadryas baboons: transfer among social units in a multilevel society. *Am. J. Phys.*
578 *Anthropol.* **145**, 360-370 (2011). doi: 10.1002/ajpa.21504; pmid: 21469076
- 579 49. J. Voisey, A. van Daal, Agouti: from mouse to man, from skin to fat. *Pigment Cell Res.*
580 **15**, 10-18 (2002). doi: 10.1034/j.1600-0749.2002.00039.x; pmid: 11837451
- 581 50. M. Petersdorf, A. H. Weyher, J. M. Kamilar, C. Dubuc, J. P. Higham, Sexual selection in
582 the Kinda baboon. *J. Hum. Evol.* **135**, 102635 (2019). doi: 10.1016/j.jhevol.2019.06.006;
583 pmid: 31421317
- 584 51. H. Svardal, A. J. Jasinska, C. Apetrei, G. Coppola, Y. Huang, C. A. Schmitt, B.
585 Jacquelin, V. Ramensky, M. Müller-Trutwin, M. Antonio, G. Weinstock, J. P. Grobler,
586 K. Dewar, R. K. Wilson, T. R. Turner, W. C. Warren, N. B. Freimer, M. Nordborg,
587 Ancient hybridization and strong adaptation to viruses across African vervet monkey
588 populations. *Nat. Genet.* **49**, 1705-1713 (2017). doi: 10.1038/ng.3980; pmid: 29083404
- 589 52. Y. Song, C. Jiang, K. H. Li, J. Li, H. Qiu, M. Price, Z.-X. Fan, J. Li, Genome-wide
590 analysis reveals signatures of complex introgressive gene flow in macaques (genus
591 *Macaca*). *Zool. Res.* **42**, 433-449 (2021). doi: 10.24272/j.issn.2095-8137.2021.038; pmid:
592 34114757

- 593 53. J. A. Cahill, R. E. Green, T. L. Fulton, M. Stiller, F. Jay, N. Ovsyanikov, R. Salamzade,
594 J. St. John, I. Stirling, M. Slatkin, B. Shapiro, Genomic evidence for island population
595 conversion resolves conflicting theories of polar bear evolution. *PLoS Genet.* **9**,
596 e1003345 (2013). doi: 10.1371/journal.pgen.1003345; pmid: 23516372
- 597 54. J. A. Cahill, P. D. Heintzman, K. Harris, M. D. Teasdale, J. Kapp, A. E. R. Soares, I.
598 Stirling, D. Bradley, C. J. Edwards, K. Graim, A. A. Kisleika, A. V. Maley, N.
599 Monaghan, R. E. Green, B. Shapiro, Genomic evidence of widespread admixture from
600 polar bears into brown bears during the last ice age. *Mol. Biol. Evol.* **35**, 1120-1129
601 (2018). doi: 10.1093/molbev/msy018; pmid: 29471451
- 602 55. F. J. Combe, L. Jaster, A. Ricketts, D. Haukos, A. G. Hope, Population genomics of free-
603 ranging Great Plains white-tailed and mule deer reflects a long history of inter-specific
604 hybridization. *Evol. Appl.* **15**, 111-131 (2022). doi:10.1111/eva.13330; pmid: 35126651
- 605 56. Y. Liu, X. Mao, J. Krause, Q. Fu, Insights into human history from the first decade of
606 ancient human genomics. *Science* **373**, 1479-1484 (2021). doi: 10.1126/science.abi8208;
607 pmid: 34554811
- 608 57. I. S. Chuma, E. K. Batamuzi, D. A. Collins, R. D. Fyumagwa, L. K. Hallmaier-Wacker,
609 R. R. Kazwala, J. D. Keyyu, I. A. Lejora, I. F. Lipende, S. Lüert, F. M. D. Paciência, A.
610 Piel, F. A. Stewart, D. Zinner, C. Roos, S. Knauf, Widespread *Treponema pallidum*
611 infection in nonhuman primates, Tanzania. *Emerg. Infect. Dis.* **24**, 1002-1009 (2018).
612 doi: 10.3201/eid2406.180037; pmid: 29774840
- 613 58. S. Knauf, J. F. Gogarten, V. J. Schuenemann, H. M. de Nys, A. Dux, M. Strouhal, L.
614 Mikalová, K. I. Bos, R. Armstrong, E. K. Batamuzi, I. S. Chuma, B. Davoust, G. Diatta,
615 R. D. Fyumagwa, R. R. Kazwala, J. D. Keyyu, I. A. V. Lejora, A. Levasseur, H. Liu, M.

- 616 A. Mayhew, O. Medianniko, D. Raoult, R. M. Wittig, C. Roos, F. H. Leendertz, D.
617 Šmajš, K. Nieselt, J. Krause, S. Calvignac-Spencer, African nonhuman primates are
618 infected with the yaws bacterium *Treponema pallidum* subsp. *pertenue*. *Emerg. Microbes*
619 *Infect.* **7**, 157 (2018). doi: 10.1038/s41426-018-0156-4; pmid: 30228266
- 620 59. I. S. Chuma, C. Roos C, A. Atickem, T. Bohm, D. A. Collins, L. Grillová, L. K.
621 Hallmaier-Wacker, R. R. Kazwala, J. D. Keyyu, S. Lüert, U. Maloueki, J. Oppelt, K. J.
622 Petrželková, A. Piel, F. A. Stewart, D. Šmajš, S. Knauf, Evidence for rare interspecies
623 transmission of the yaws bacterium *Treponema pallidum* subsp. *pertenue* in African
624 primates. *Sci. Rep.* **9**, 14243 (2019). doi:10.1038/41598-019-50779-9; pmid: 31578447
- 625 60. J. Rogers, K. K. Kidd, Nuclear DNA polymorphisms in a wild population of yellow
626 baboons (*Papio hamadryas cynocephalus*) from Mikumi National Park, Tanzania. *Am. J.*
627 *Phys. Anthropol.* **90**, 477-486 (1993). doi: 10.1002/ajpa.1330900407; pmid: 8097371
- 628 61. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler
629 transform. *Bioinformatics* **25**, 1754-1760 (2009). doi: 10.1093/bioinformatics/btp; pmid:
630 19451168
- 631 62. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytzky, K.
632 Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo MA, The Genome Analysis
633 Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data.
634 *Genome Res.* **20**, 1297-1303 (2010). doi: 10.1101/gr.107524.110; pmid: 20644199
- 635 63. M. Patterson, T. Marschall, N. Pisanti, L. van Iersel, L. Stougie, G. W. Klau, A.
636 Schönhuth, WhatsHap: weighted haplotype assembly for future-generation sequencing
637 reads. *J. Comput. Biol.* **22**, 498-509 (2015). doi: 10.1089/cmb.2014.0157; pmid:
638 25658651

- 639 64. O. Delaneau, J.-F. Zagury, M. R. Robinson, J. L. Marchini, E. T. Dermitzakis, Accurate,
640 scalable and integrative haplotype estimation. *Nat. Commun.* **10**, 5436 (2019). doi:
641 10.1038/s41467-019-13225-y; pmid: 31780650
- 642 65. N. Patterson, A. L. Price, D. Reich, Population structure and Eigenanalysis. *PLoS Genet.*
643 **2**, e190 (2006). doi: 10.1371/journal.pgen.0020190; pmid: 17194218
- 644 66. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in
645 unrelated individuals. *Genome Res.* **19**, 1655-1664 (2009). doi: 10.1101/gr.094052.109;
646 pmid: 19648217
- 647 67. L. T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh. IQ-TREE: a fast and
648 effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol.*
649 *Evol.* **32**, 268-274 (2015). doi:10.1093/molbev/msu300; pmid: 25371430
- 650 68. D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2:
651 improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518-522 (2018).
652 doi:10.1093/molbev/msx281; pmid: 29077904
- 653 69. S. Kalyaanamoorthy, B. Q., Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermiin,
654 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**,
655 587-589 (2017). doi: 10.1038/nmeth.4285; pmid: 28481363
- 656 70. T. H. To, M. Jung, S. Lycett, O. Gascuel, Fast dating using least-squares criteria and
657 algorithms. *Syst. Biol.* **65**, 82-97 (2016). doi: 10.1093/sysbiosyv068; pmid: 26424727
- 658 71. N. G. Jablonski, S. Frost, "Cercopithecoidea" in *Cenozoic Mammals of Africa*, L.
659 Werdelin, W. J. Sanders, Eds. (California Univ. Press, 2010), pp. 393-428.
- 660 72. T. Harrison, "Catarrhine origins" in *A Companion to Paleoanthropology*, D. R. Begun,
661 Ed. (Blackwell, 2013), pp. 376-396.

- 662 73. R. R. Bouckaert, J. Heled, D. Kuehnert, T. G. Vaughan, C.-H. Wu, D. Xie, M. A.
663 Suchard, A. Rambaut, A. J. Drummond, BEAST 2: a software platform for Bayesian
664 evolutionary analysis. *PLoS Comput. Biol.* **10**, e1003537 (2014). doi:
665 10.1371/journal.pcbi.1003537; pmid: 24722319
- 666 74. R. C. Edgar, Quality measures for protein alignment benchmarks. *Nucleic Acids Res.* **38**,
667 2145-2153 (2010). doi: 10.1093/nar/gkp1196; pmid: 20047958
- 668 75. A. Larsson, AliView: a fast and lightweight alignment viewer and editor for large
669 datasets. *Bioinformatics* **30**, 3276-3278 (2014). doi: 10.1093/bioinformatics/btu531;
670 pmid: 25095880
- 671 76. T. Rausch, T. Zichner, A. Schlattl, A. M. Stütz, V. Benes, J. O. Korbel, DELLY:
672 structural variant discovery by integrated paired-end and split-read analysis.
673 *Bioinformatics* **28**, i333-i339 (2012). doi: 10.1093/bioinformatics/bts378; pmid:
674 22962449
- 675 77. E. J. Gardner, V. K. Lam, D. N. Harris, N. T. Chuang, E. C. Scott, W. S. Pittard, R. E.
676 Mills, 1000 Genomes Project Consortium, S.E. Devine, The Mobile Element Locator
677 Tool (MELT): population-scale mobile element discovery and biology. *Genome Res.* **27**,
678 1916-1929 (2017). doi: 10.1101/gr.218032.116; pmid: 28855259
- 679 78. D. Falush, M. Stephens, J. K. Pritchard, Inference of population structure using
680 multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**,
681 1567-1587 (2003). doi: 10.1093/genetics/164.4.1567; pmid: 12930761
- 682 79. G. Evanno, S. Regnaut, J. Goudet, Detecting the number of clusters of individuals using
683 the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**, 2611-2620 (2005). doi:
684 10.1111/j.1365-294X.2005.02553.x; pmid: 15969739

- 685 80. D. A. Earl, B. M. vonHoldt, STRUCTURE HARVESTER: a website and program for
686 visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet.*
687 *Resour.* **4**, 359-361 (2012). doi: 10.1007/s12686-011-9548-7
- 688 81. J. M. Storer, R. Hubley, J. Rosen, A. F. A. Smith, Curation guidelines for *de novo*
689 generated transposable element families. *Curr. Protoc.* **1**, e154 (2021). doi:
690 10.1002/cpz1.154; pmid: 34138525
- 691 82. M. K. Konkel, J. A. Walker, M. A. Batzer, LINEs and SINEs of primate evolution. *Evol.*
692 *Anthropol.* **19**, 236-249 (2010). doi: 10.1002/evan.20283; pmid: 25147443
- 693 83. W. C. Warren, R. A. Harris, M. Haukness, I. T. Fiddes, S. C. Murali, J. Fernandes, P. C.
694 Dishuck, J. M. Storer, M. Raveendran, L. W. Hillier, D. Porubsky, Y. Mao, D. Gordon,
695 M. R. Vollger, A. P. Lewis, K. M. Munson, E. DeVogelaere, J. Armstrong, M. Diekhans,
696 J. A. Walker, C. Tomlinson, T. A. Graves-Lindsay, M. Kremitzki, S. R. Salama, P. A.
697 Audano, M. Escalona, N. W. Maurer, F. Antonacci, L. Mercuri, F. A. M. Maggiolini, C.
698 R. Catacchio, J. G. Underwood, D. H. O'Connor, A. D. Sanders, J. O. Korbel, B.
699 Ferguson, H. M. Kubisch, L. Picker, N. H. Kalin, D. Rosene, J. Levine, D. H. Abbott, S.
700 B. Gray, M. M. Sanchez, Z. A. Kovacs-Balint, J. W. Kemnitz, S. M. Thomasy, J. A.
701 Roberts, E. L. Kinnally, J. P. Capitanio, J. H. P. Skene, M. Platt, S. A. Cole, R. E. Green,
702 M. Ventura, R. W. Wiseman, B. Paten, M. A. Batzer, J. Rogers, E. E. Eichler, Sequence
703 diversity analyses of an improved rhesus macaque genome enhance its biomedical utility.
704 *Science* **370**, eabc6617 (2020). doi: 10.1126/science.abc6617; pmid: 33335035
- 705 84. D. L. Swofford, *PAUP*: Phylogenetic Analysis using Parsimony (*and other*
706 *Methods)*, *Version 4.0a169* (Sinauer, 2011).

- 707 85. S. Schiffels, K. Wang, MSMC and MSMC2: the multiple sequentially Markovian
708 coalescent. *Methods Mol. Biol.* **2090**, 147-166 (2020). doi: 10.1007/978-1-0716-0199-
709 0_7; pmid: 31975167
- 710 86. N. Patterson, P. Moorjani, Y. Luo, S. Mallick, N. Rohland, Y. Zhan, T. Genschoreck, T.
711 Webster, D. Reich, Ancient admixture in human history. *Genetics* **192**, 1065-1093
712 (2012). doi: 10.1534/genetics.112.145037; pmid: 22960212
- 713 87. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis,
714 R. Durbin, 1000 Genome Project Data Processing Subgroup, The
715 sequence/alignment/map (SAM) format and Samtools. *Bioinformatics* **25**, 2078-2079
716 (2009). doi: 10.1093/bioinformatics/btp352; pmid: 19505943
- 717 88. K. Leppälä, S. V. Nielsen, T. Mailund, admixturegraph: an R package for admixture
718 graph manipulation and fitting. *Bioinformatics* **33**, 1738-1740 (2017). doi:
719 10.1093/bioinformatics/btx048; pmid: 28158333
- 720 89. M. Petr, B. Vernot, J. Kelso, admixr-R package for reproducible analyses using
721 ADMIXTOOLS. *Bioinformatics* **35**, 3194-3195 (2019). doi:
722 10.1093/bioinformatics/btz030; pmid: 30668635
- 723 90. P. Wangkumhang, M. Greenfield, G. Hellenthal, An efficient method to identify, date,
724 and describe admixture events using haplotype information. *Genome Res.* in press. doi:
725 10.1101/gr.275994.121; pmid: 35794007
- 726 91. J. Fadista, A. K. Manning, J. C. Florez, L. Groop, The (in)famous GWAS P-value
727 threshold revisited and updated for low-frequency variants. *Eur. J. Hum Genet.* **24**, 1202-
728 1205 (2016). doi: 10.1038/ejhg.2015.269; pmid: 26733288

- 729 92. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic
730 features. *Bioinformatics* **26**, 841-842 (2010). doi: 10.1093/bioinformatics/btq033; pmid:
731 20110278
- 732 93. P. Rentzsch, D. Witten, G. M. Cooper, J. Shendure, M. Kircher, CADD: predicting the
733 deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* **47**, D886-
734 D894 (2019). doi: 10.1093/nar/gky1016; pmid: 30371827
- 735 94. N. M. Ioannidis, J. H. Rothstein, V. Pejaver, S. Middha, S. K. McDonnell, S. Baheti, A.
736 Musolf, Q. Li, E. Holzinger, D. Karyadi, L. A. Cannon-Albright, C. C. Teerlink, J. L.
737 Stanford, W. B. Isaacs, J. Xu, K. A. Cooney, E. M. Lange, J. Schleutker, J. D. Carpten, I.
738 J. Powell, O. Cussenot, G. Cancel-Tassin, G. G. Giles, R. J. MacInnis, C. Maier, C.-L.
739 Hsieh, F. Wiklund, W. J. Catalona, W. D. Foulkes, D. Mandal, R. A. Eeles, Z. Kote-Jarai,
740 C. D. Bustamante, D. J. Schaid, T. Hastie, E. A. Olander, J. E. Bailey-Wilson, P.
741 Radivojac, S. N. Thibodeau, A. S. Whittemore, W. Sinh, REVEL: an Ensemble method
742 for predicting pathogenicity of rare missense variants. *Am. J. Hum. Genet.* **99**, 877-885
743 (2016). doi: 10.1016/j.ajhg.2016.08.016; pmid: 27666373
- 744 95. D. W. Huang, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of
745 large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44-57 (2009). doi:
746 10.1038/nprot.2008.211; pmid: 19131956
- 747 96. B. S. Weir, C. C. Cockerham, Estimating F -statistics for the analysis of population
748 structure. *Evolution* **38**, 1358-1370 (1984). doi: 10.1111/j.1558-5646.1984.tb05657.x;
749 pmid: 28563791
- 750 97. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E.
751 Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, 1000 Genomes

- 752 Project Analysis Group, The variant call format and VCFtools. *Bioinformatics* **27**, 2156-
753 2158 (2011). doi: 10.1093/bioinformatics/btr330; pmid: 21653522
- 754 98. J. Reimand, M. Kull, H. Peterson, J. Hansen, J. Vilo, g:Profiler--a web-based toolset for
755 functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res.* **35**,
756 W193-W200 (2007). doi: 10.1093/nar/gkm226; pmid: 17478515
- 757 99. D. Zinner, S. Buba, S. Nash, C. Roos “Pan-African voyagers: the phylogeography of
758 baboons” in *Primates of Gashaka*, V. Sommer, C. Ross, Eds. (Springer, 2011), pp. 267-
759 306.
- 760 100. R. R. Ackermann, L. Schroeder, J. Rogers, J. M. Cheverud, Further evidence for
761 phenotypic signatures of hybridization in descendant baboon populations. *J. Hum. Evol.*
762 **76**, 54-62 (2014). doi: 10.1016/j.jhevol.2014.05.004; pmid: 24935168
- 763 101. C. Groves, *Primate Taxonomy* (Smithsonian, 2001).
- 764 102. L. Swedell, “African papionins: diversity of social organization and ecological
765 flexibility” in *Primates in Perspective*, C. J. Campbell, A. Fuentes, K. C. MacKinnon, S.
766 K. Bearder, R. M. Stumpf, Eds. (Oxford Univ. Press, 2. Edition, 2011), pp. 241-277.
- 767 103. S. Boissinot, L. Alvarez, J. Giraldo-Ramirez, M. Tollis, Neutral nuclear variation in
768 baboons (genus *Papio*) provides insights into their evolutionary and demographic
769 histories. *Am. J. Phys. Anthropol.* **155**, 621-634 (2014). doi: 10.1002/ajpa.22618; pmid:
770 25234435
- 771 104. J. Cracraft, “Species concepts and speciation analysis” in *Current Ornithology, Vol. 1*, R.
772 F. Johnston, Ed. (Springer, 1983), pp. 159-187.
- 773 105. J. Cracraft, “Speciation and its ontology” in *Speciation and Its Consequences*, D. Otte, J.
774 A. Endler, Eds. (Sinauer, 1989), pp. 28-59.

- 775 106. M. de Manuel, M. Kuhlwilm, P. Frandsen, V. C. Sousa, T. Desai, J. Prado-Martinez, J.
776 Hernandez-Rodriguez, I. Dupanloup, O. Lao, P. Hallast, J. M. Schmidt, J. M. Heredia-
777 Genestar, A. Benazzo, G. Barbujani, B. M. Peter, L. F. Kuderna, F. Casals, S. Angedakin,
778 M. Arandjelovic, C. Boesch, H. Kühl, L. Vigilant, K. Langergraber, J. Novembre, M.
779 Gut, I. Gut, A. Navarro, F. Carlsen, A. M. Andrés, H. R. Siegismund, A. Scally, L.
780 Excoffier, C. Tyler-Smith, S. Castellano, Y. Xue, C. Hvilsom, T. Marques-Bonet,
781 Chimpanzee genomic diversity reveals ancient admixture with bonobos. *Science* **354**,
782 477-481 (2016). doi: 10.1126/science.aag2602; pmid: 27789843
- 783 107. R. C. Williams, M. B. Blanco, J. W. Poelstra, K. E. Hunnicutt, A. A. Comeault, A. D.
784 Yoder, Conservation genomic analysis reveals ancient introgression and declining levels
785 of genetic diversity in Madagascar's hibernating dwarf lemurs. *Heredity* **124**, 236-251
786 (2020). doi: 10.1038/s41437-019-0260-9; pmid: 31435007
- 787 108. H. V. Figueiro, G. Li, F. J. Trindade, J. Assis, F. Pais, G. Fernandes, S. H. D. Santos, G.
788 M. Hughes, A. Komissarov, A. Antunes, C. S. Trinca, M. R. Rodrigues, T. Linderoth, K.
789 Bi, L. Silveira, F. C. C. Azevedo, D. Kantek, E. Ramalho, R. A. Brassaloti, P. M. S.
790 Villela, A. L. V. Nunes, R. H. F. Teixeira, R. G. Morato, D. Loska, P. Saragüeta, T.
791 Gabaldón, E. C. Teeling, S. J. O'Brien, R. Nielsen, L. L. Coutinho, G. Oliveira, W. J.
792 Murphy, E. Eizirik, Genome-wide signatures of complex introgression and adaptive
793 evolution in the big cats. *Sci. Adv.* **3**, e1700299 (2017). doi: 10.1126/sciadv.1700299;
794 pmid: 28776029
- 795 109. E. Mayr, *Systematics and the Origin of Species* (Columbia Univ. Press, 1942).

- 796 110. R. W. Thorington, C. Groves, “An annotated classification of the Cercopithecoidea” in
797 *Old World Monkeys: Evolution, Systematics and Behavior*, J. R. Napier, P. H. Napier,
798 Eds. (Academic, 1970), pp. 629-647.
- 799 111. J. F. Fleage, Identifying primate species. *Evol. Anthropol.* **23**, 1 (2014).
800 doi:10.1002/evan.21398; pmid: 24591130
- 801 112. C. Groves, The species in primatology. *Evol. Anthropol.* **23**, 2-4 (2014).
802 doi:10.1002/evan.21395; pmid: 24591131
- 803 113. C. J. Jolly, A Darwinian species definition and its implications. *Evol. Anthropol.* **23**, 36-
804 38 (2014). doi:10.1002/evan.21396; pmid: 24591142
- 805 114. D. Zinner, C. Roos, So what is a species? A primatological perspective. *Evol. Anthropol.*
806 **23**, 21-23 (2014). doi:10.1002/evan.21390; pmid: 24591137

807

808 **ACKNOWLEDGEMENTS**

809 We thank all countries and their respective governmental and non-governmental institutions that
810 supported sampling and sample analysis. Specifically, we thank the Government of the United
811 Republic of Tanzania, Ministry for Education and Vocational Training, Commission for Science
812 and Technology, Ministry for Natural Resources and Tourism, Ministry for Agriculture, Natural
813 Resources, Livestock and Fisheries, Department of Forestry and Non-renewable Natural
814 Resources, Tanzania Wildlife Authority, Tanzania Wildlife Research Institute, Tanzania
815 National Parks (Yustina A. Kiwango, Rehema Kaitila, Inyasi A. V. Lejora), Ngorongoro
816 Conservation Area Authority, Sokoine University of Agriculture (Rudovick R. Kazwala),
817 National Institute for Medical Research (Clara C. Lubinza, Sayoki G. M. Mfinanga), Jane
818 Goodall Institute (Iddi F. Lippende, D. Anthony Collins), and the Greater Mahale Ecosystem

819 Research and Conservation Project (Alexander Piel, Fiona A. Stewart). For Zambia, we thank
820 the Government of Zambia, the Zambia Wildlife Authority (Jack Chulu, Edwin Matokwani;
821 Chilanga), the staff of Kafue National Park, and the Department of Veterinary and Livestock
822 Development (Yona Sinkala, Lusaka). For Ethiopia, we thank the Government of Ethiopia, the
823 Ethiopian Public Health Institute (Ebba Abate, Addis Ababa), the Guinea Work Eradication
824 Program of The Carter Center (Jim Zingeser, Ernesto Ruiz-Tiben, Zerihun Tadesse, Addis
825 Ababa) as well as James Else, Harry Marshall and the logistics team in Gog Woreda. For
826 Senegal, we thank the *Diréction des Parcs Nationaux* and *Ministère de l'Environnement et de la*
827 *Protéction de la Nature de la République du Sénégal* for permission to work in the Niokolo Koba
828 National Park. We particularly thank the former conservateurs of the park Colonel Ousmane
829 Kane and Commandant Mallé Gueye for their cooperation and logistical support during the study
830 period, and all the staff and field assistants of the CRP Simenti, in particular Mustapha Faye,
831 Armél Louis Nyafouna, Elhadji Dansokho, Lamine Diedhiou, Moustapha Dieng and Touradou
832 Sonko for their support in the field.

833 **Funding:** “la Caixa” Foundation (ID 100010434), fellowship code LCF/BQ/PR19/11700002
834 (MK), the Vienna Science and Technology Fund (WWTF) and the City of Vienna project
835 VRG20-001 (MK), German Research Foundation grants FI707/9-1, KN0197/3-1, KN1097/4-1,
836 ZI548/5-1 and RO3055/2-1 (JF, SK, DZ, CR), Novo Nordisk Foundation grant 0058553 (EFS,
837 KM), R01 GM59290 (MAB), Internal funding from Baylor College of Medicine (JR)

838 **Author contributions:** Conceptualization: KK-HF, TM-B, CR, JR; Data curation: EFS, RAH,
839 LZ, MR, LFKK; Formal analysis: EFS, RAH, LZ, MR, JAW, JMS, MK, CF, LS, CMB, ASB,
840 JB, KM, CR; Funding acquisition: KK-HF, TM-B, KM, CR, JR; Investigation: EFS, RAH, LZ,
841 MR, LFKK, JAW, JMS, MK, CF, LS, CMB, ASB, JB, MHS, MAB, CJJ, KM, CR;

842 Methodology: MK, CF, CMB, M-CG, SS, HD, MAB, SK, DZ, TM-B, KM, CR, JR; Project
843 administration: KM, CR, JR; Resources/Sample acquisition: CMB, ASB, JEP-C, FS, KLC, ISC,
844 JDK, JF, CJJ, SK, DZ, CR, JR; Supervision: TM-B, KM, CR, JR; Visualization: EFS, RAH, LZ,
845 JAW, JMS, MK, CF, CMB, DZ, KM, CR; Writing – original draft: EFS, RAH, LZ, JAW, MK,
846 CF, CMB, CJJ, DZ, TM-B, KM, CR, JR; Writing – review & editing: all authors

847 **Competing interests:** LFKK and KK-HF are employees of Illumina Inc.; all other authors
848 declare that they have no competing interests.

849 **Data and materials availability:** The sequencing data used in these analyses are available
850 through the Short Read Archive under BioProject accession PRJEB49549. Additional data are
851 available in the Supplementary Materials.

852

853 **SUPPLEMENTARY MATERIALS**

854 Materials and Methods

855 Supplementary Text S1 to S5

856 Figs. S1 to S36

857 Tables S1 to S11

858 References and Notes (57–114)

859 **Tables**

860 **Table 1. Species enriched SNV statistics.** Cluster and OmegaPlus statistics for the hamadryas
861 and Guinea baboon shared SNV_2 are shown for hamadryas baboons. CADD and REVEL
862 scores from human annotations predict functional impact of mutations (see Supplementary
863 Materials).

SNV ID	SNV	PLINK p value	Cluster Length (bp)	SNVs in Cluster	OmegaPlus (percentile)	CADD PHRED	REVEL
SNV_1	20:27347531:G:T	1.40×10^{-78}	64,284	24	5.59 (1.8%)	0.001	0.351
SNV_2	13:49896439:G:C	9.72×10^{-101}	126,701	96	11.01 (0.3%)	7.266	NA
SNV_3	10:30107617:T:C	2.52×10^{-69}	39,912	58	4.92 (0.7%)	19.140	0.080

864

865 **Figure Legends**

866 **Fig. 1. Distribution of the six baboon species and sampling sites.** Species distributions are
867 modified from (20). The insert map shows sampling sites in Tanzania. Drawings of male
868 baboons by Stephen Nash, used with permission. Numbers of samples per species are given in
869 parentheses.

870

871 **Fig. 2. Population structure and coancestry of the six baboon species.** (A) PCA of autosomal
872 SNVs. (B) ChromoPainter coancestry matrix with fineSTRUCTURE dendrogram. Each row in
873 the coancestry matrix represents an individual and illustrates how its most recent common
874 ancestry is distributed across all other sampled individuals. The ordering of individuals is the
875 same for rows and columns. The row color labels are the same as in A and correspond to clusters
876 shown for eight populations labeled with boxes: A: Gog olive (Ethiopia), B: hamadryas, C:
877 Guinea, D: southern olive (Kenya and Tanzania), E: Kinda, F: western yellow, G: chacma, H:
878 eastern yellow; X: olive coancestry in western yellows suggesting admixture (see alternate
879 fineSTRUCTURE figure fig. S4). Colors labels below the dendrogram represent the 14 groups
880 named in the figure legend. (C) PCA of the coancestry matrix. (D) ADMIXTURE plot with the
881 preferred grouping of baboons into seven clusters ($K = 7$; for $K = 2-10$ see fig. S12).

882

883 **Fig. 3. Population history and complex reticulation between baboon populations.** (A)
884 MSMC2 plots using a mutation rate of 0.9×10^{-8} and a generation time of 11 years (23). (B)
885 Admixture graph of the populations used in this study, based on 48,730,011 SNPs with data for
886 all individuals, and a predefined number of two admixture events. Numbers on solid branches
887 correspond to the estimated drift in f_2 units of squared frequency difference; labels on dotted

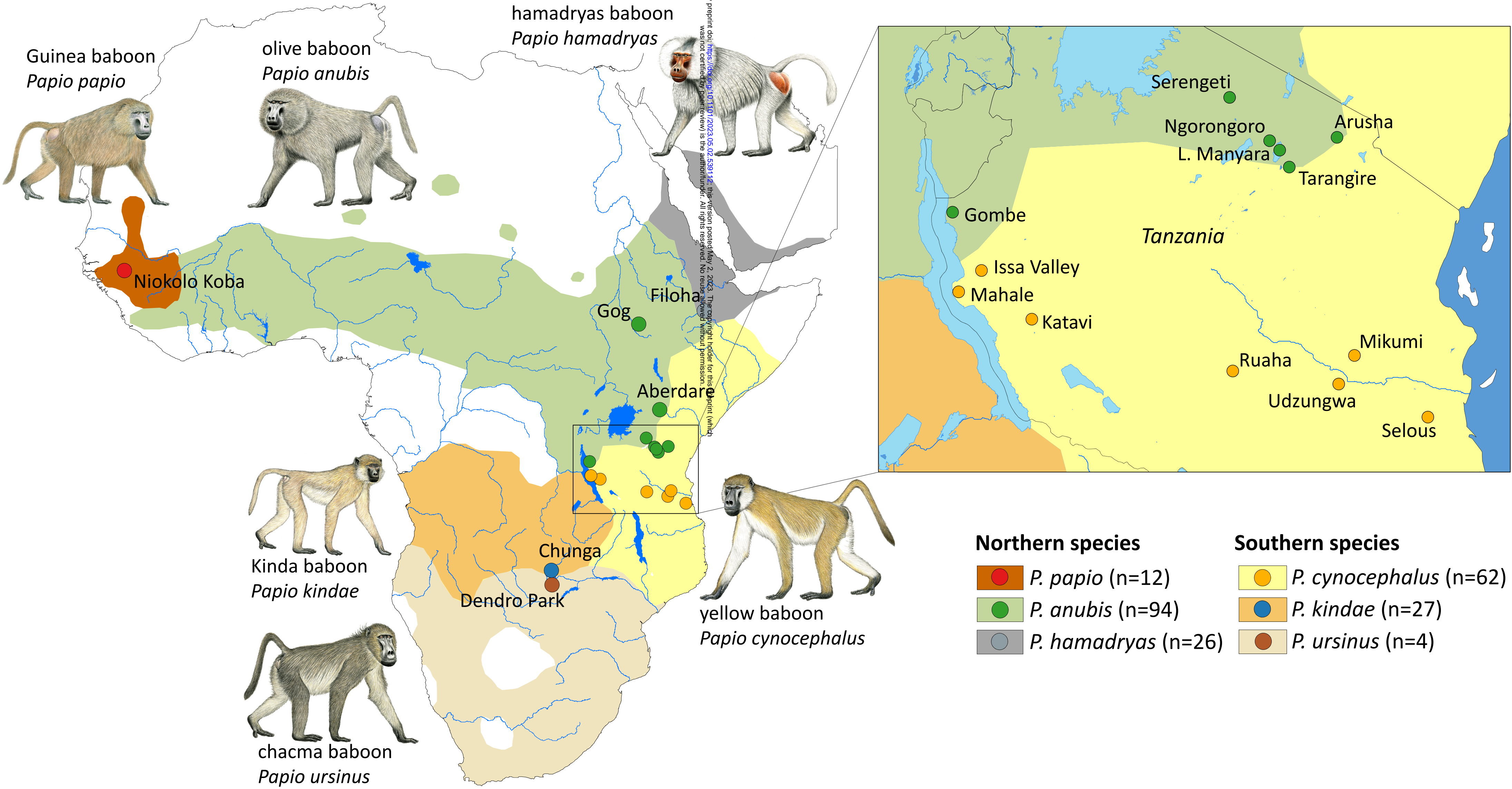
888 edges give admixture proportions. (C) Globetrotter analysis of the eight major regional
889 populations. The pie chart for each cluster shows ancestry contributions from other clusters.
890 Expanded wedges represent ancestry that can be attributed to recent admixture (< 56 generations,
891 bootstrap p-values < 0.05). (D) Same as C but for 14 populations separating each major sampling
892 location (here, expanded wedges represent ancestry that can be attributed to admixture more
893 recent than 95 generations, bootstrap p-values < 0.05).

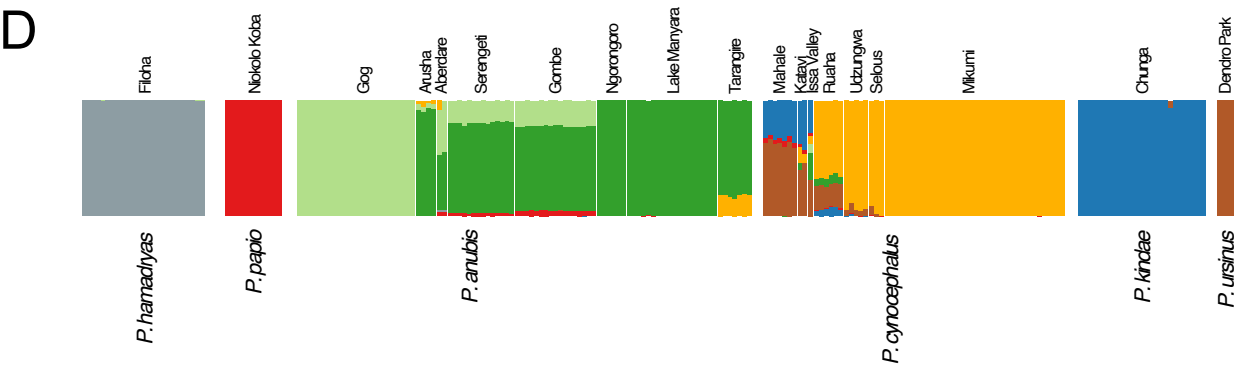
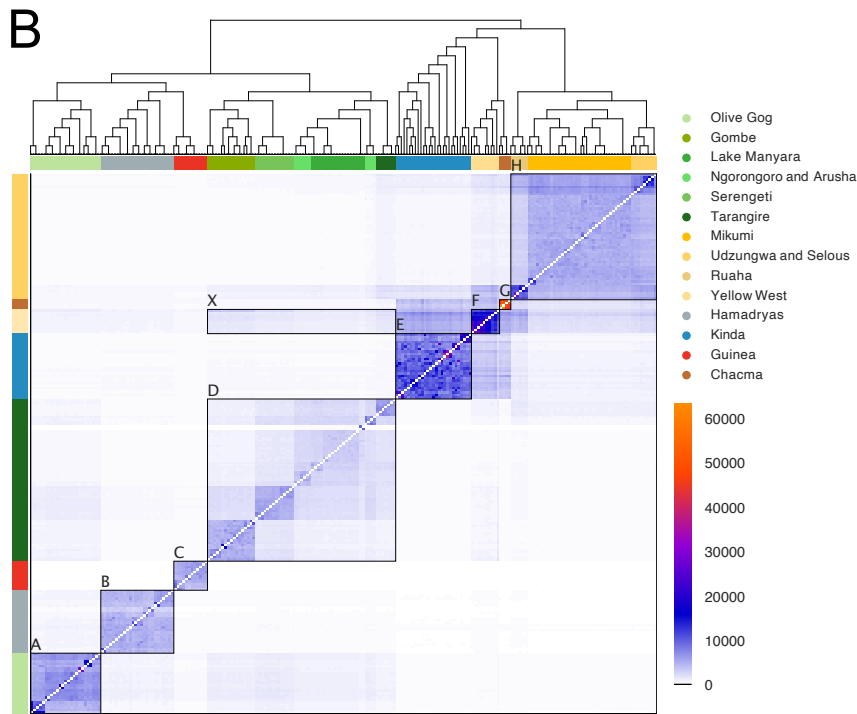
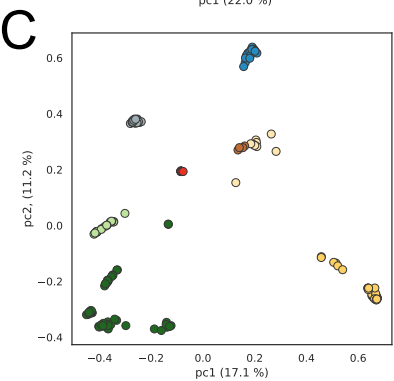
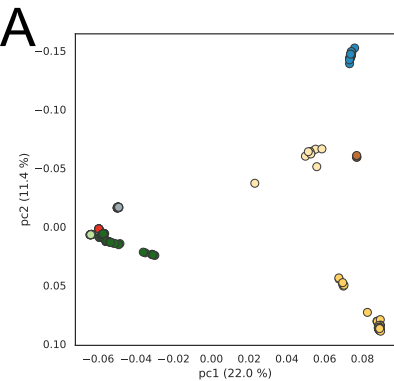
894

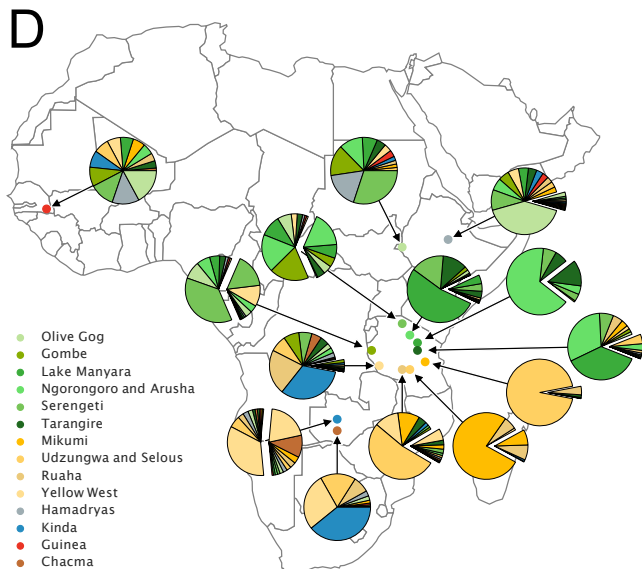
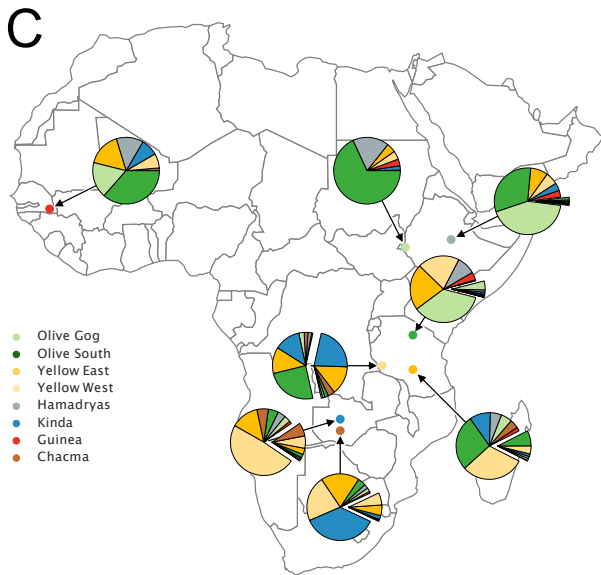
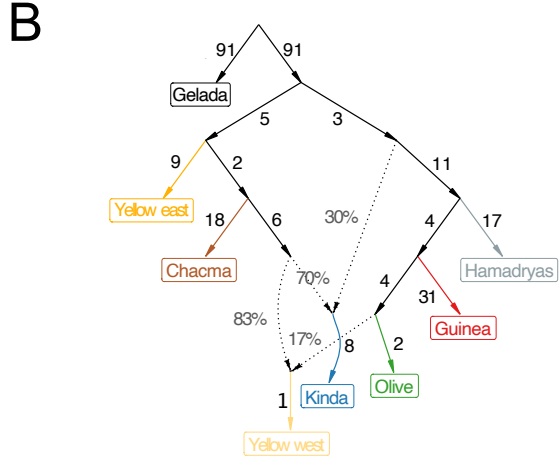
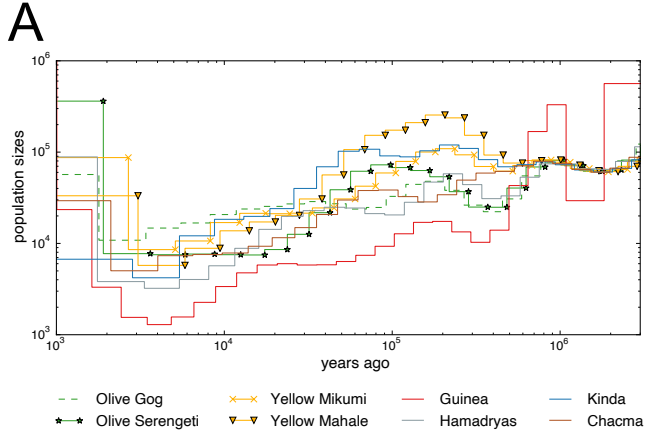
895 **Fig. 4. Geographic distribution of mtDNA clades and mtDNA phylogeny.** (A) Distribution
896 ranges of baboon species and the four main mtDNA clades (south, south-east, north-east, north-
897 west, dashed lines) including major mitochondrial lineages (A-R). (B) Phylogeny based on
898 complete mtDNA genomes (see also fig. S8). Clade designation follows (20, 21), asterisks
899 indicate lineages from which mtDNA genomes have been generated in this study. For identical
900 haplotypes see table S7.

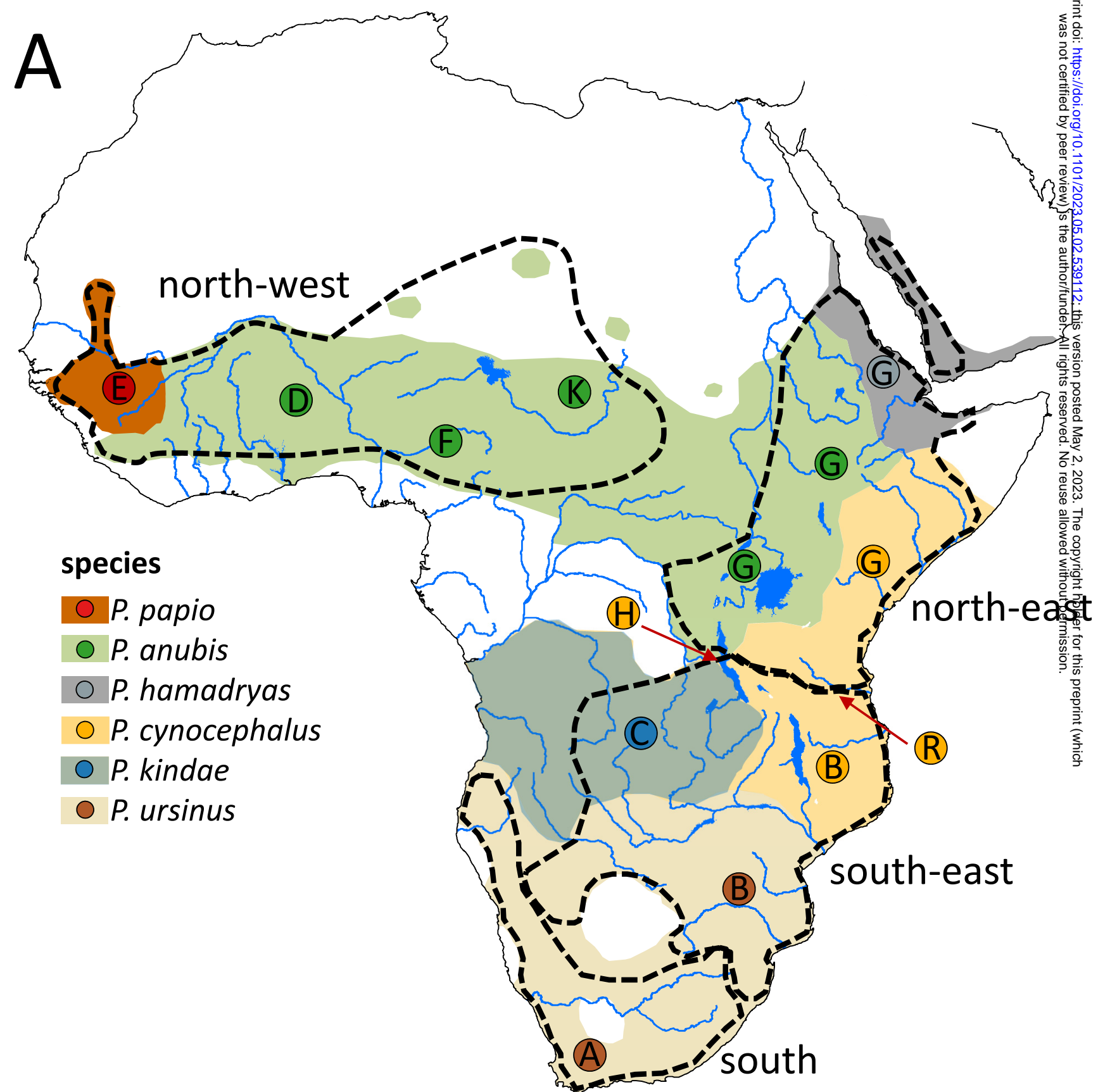
901

902 **Fig. 5. Differential ancestry profiles on the X-chromosome and an autosome.** (A) Ancestry
903 proportions of female chacma baboons. Each marker represents the fraction of total chromosome
904 ancestry of one individual that is assigned to each of the remaining donor populations. Black dots
905 and grey crosses represent ancestry proportions of chromosomes 8 and X, respectively. (B) Same
906 as A but for female western yellow baboons. (C) Same as A but for female eastern yellow
907 baboons. (D) Same as A but for female Kinda baboons. For additional profiles see figs. S10, S30
908 and S31.

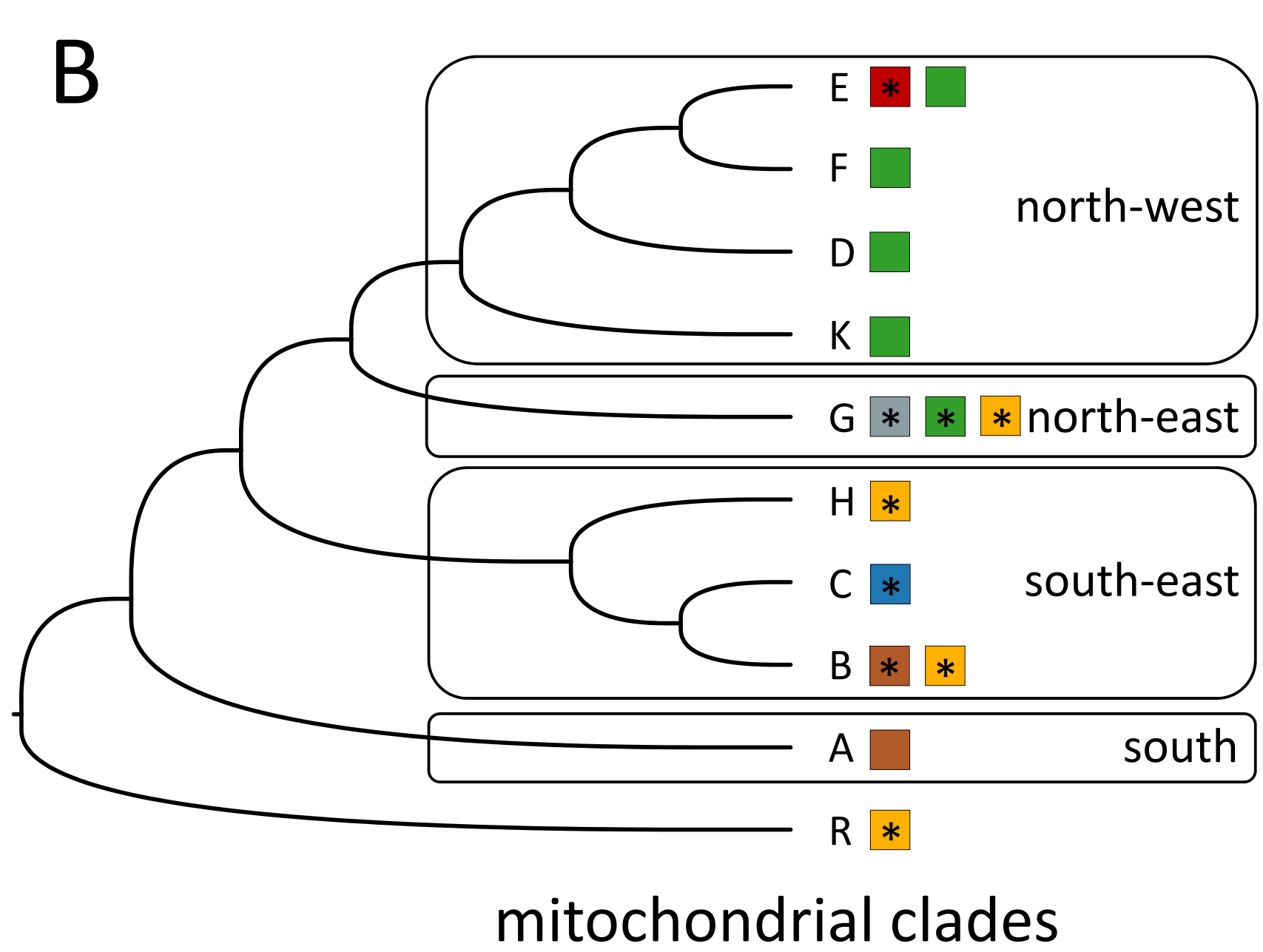


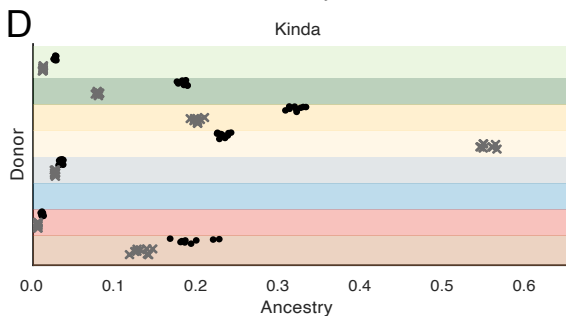
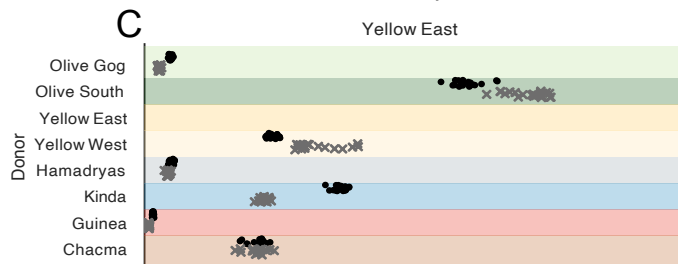
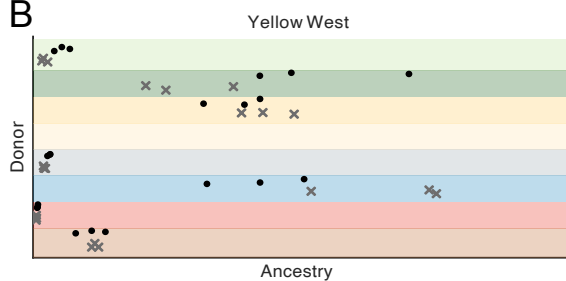
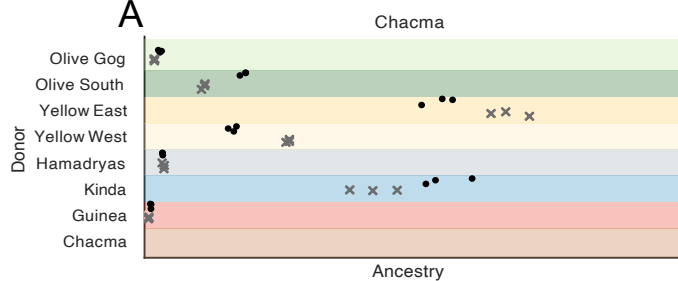






print doi: <https://doi.org/10.1101/2023.05.02.539112>; this version posted May 2, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.





Chromosome

- chr8
- × chrX