

1 **Ancient genomes from southern Africa pushes modern human divergence beyond**  
2 **260,000 years ago**  
3

4 Carina M. Schlebusch<sup>1,4\*</sup>, Helena Malmström<sup>1,4\*</sup>, Torsten Günther<sup>1</sup>, Per Sjödin<sup>1</sup>, Alexandra Coutinho<sup>1</sup>,  
5 Hanna Edlund<sup>1</sup>, Arielle R. Munters<sup>1</sup>, Maryna Steyn<sup>2</sup>, Himla Soodyall<sup>3</sup>, Marlize Lombard<sup>4,5#</sup>, Mattias  
6 Jakobsson<sup>1,4,6#</sup>  
7  
8

9 1 Department of Organismal Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen  
10 18C, SE-752 36 Uppsala, Sweden  
11

12 2 Human Variation and Identification Research Unit, School of Anatomical Sciences, Faculty of Health  
13 Sciences, University of the Witwatersrand, South Africa.  
14

15 3 Division of Human Genetics, School of Pathology, Faculty of Health Sciences, University of the  
16 Witwatersrand and National Health Laboratory Service, Johannesburg, South Africa.  
17

18 4 Centre for Anthropological Research & Department of Anthropology and Development Studies,  
19 University of Johannesburg, P.O. Box 524, Auckland Park, 2006, South Africa  
20

21 5 Stellenbosch Institute for Advanced Study, (STIAS) Wallenberg Research Centre at Stellenbosch  
22 University, Marais Street, Stellenbosch, 7600, South Africa  
23

24 6 SciLife Lab Uppsala, Sweden  
25

26 \* These authors contributed equally  
27

28 # Correspondence: Mattias Jakobsson ([Mattias.jakobsson@ebc.uu.se](mailto:Mattias.jakobsson@ebc.uu.se)) and Marlize Lombard  
29 ([mlombard@uj.ac.za](mailto:mlombard@uj.ac.za))

30 **ABSTRACT**

31 Southern Africa is consistently placed as one of the potential regions for the evolution of *Homo sapiens*.  
32 To examine the region's human prehistory prior to the arrival of migrants from East and West Africa or  
33 Eurasia in the last 1,700 years, we generated and analyzed genome sequence data from seven ancient  
34 individuals from KwaZulu-Natal, South Africa. Three Stone Age hunter-gatherers date to ~2,000 years  
35 ago, and we show that they were related to current-day southern San groups such as the Karretjie People.  
36 Four Iron Age farmers (300-500 years old) have genetic signatures similar to present day Bantu-  
37 speakers. The genome sequence (13x coverage) of a juvenile boy from Ballito Bay, who lived ~2,000  
38 years ago, demonstrates that southern African Stone Age hunter-gatherers were not impacted by recent  
39 admixture; however, we estimate that all modern-day Khoekhoe and San groups have been influenced  
40 by 9-22% genetic admixture from East African/Eurasian pastoralist groups arriving >1,000 years ago,  
41 including the Ju|'hoansi San, previously thought to have very low levels of admixture. Using traditional  
42 and new approaches, we estimate the population divergence time between the Ballito Bay boy and other  
43 groups to beyond 260,000 years ago. These estimates dramatically increases the deepest divergence  
44 amongst modern humans, coincide with the onset of the Middle Stone Age in sub-Saharan Africa, and  
45 coincide with anatomical developments of archaic humans into modern humans as represented in the  
46 local fossil record. Cumulatively, cross-disciplinary records increasingly point to southern Africa as a  
47 potential (not necessarily exclusive) 'hot spot' for the evolution of our species.

## 48 Main text

49 Archaeological, fossil and genetic data consistently place the earliest traces of anatomically modern  
50 humans in sub-Saharan Africa<sup>1-6</sup>. East Africa often features in human origins studies, because the earliest  
51 modern human remains, dating to ~190 kya (kya = thousand years old/ago), originates from Ethiopia<sup>2,3</sup>.  
52 In southern Africa, cross-disciplinary data however converge, indicating that understanding the  
53 population histories of the region can contribute to understanding the origins of our species<sup>4,7-10</sup>.

54  
55 Recent syntheses indicate the occupation of the southern African landscape by the genus *Homo* from  
56 about 2 Mya<sup>11</sup>, with a major transitional phase between 600 kya and 200 kya (from the Earlier Stone  
57 Age into the Middle Stone Age)<sup>12</sup>. Current interpretation of the fossil record indicates the presence of  
58 archaic *H. sapiens* at >200 kya, during the earlier phases of the Middle Stone Age, and anatomically  
59 modern humans from ~120 kya<sup>11</sup>. After ~120 kya we also see some of the earliest archaeological  
60 evidence for modern human behavior and thinking in sub-Saharan Africa<sup>13</sup>. Genetic studies  
61 (mitochondria, Y-chromosome, and autosomes) constantly identify southern African Khoe-San  
62 populations as carrying more unique variants and more divergent lineages than any other living groups<sup>1</sup>  
63 <sup>8-10, 14-17</sup>. The deepest population split among modern humans – between Khoe-San and other groups –  
64 is estimated to ~160–100 kya, based on short sequence fragments<sup>15, 16</sup>, and genome-wide SNP data<sup>8</sup>.  
65 Some of these patterns were used to argue for a southern African origin of humans<sup>10</sup>, although others  
66 suggested several regions<sup>8, 18</sup>.

67  
68 Middle Stone Age sites in KwaZulu-Natal, South Africa, e.g., Sibudu Cave<sup>13</sup>, Umhlatuzana Rock  
69 Shelter<sup>19</sup> and Border Cave<sup>20</sup>, demonstrate human occupation since >100 kya. Here we report the  
70 sequencing and analyses of the genomes of seven ancient individuals from KwaZulu-Natal (Table 1,  
71 Figure 1), directly dated to the last 2 kya. The genomes of individuals from Ballito Bay and Doonside  
72 represent the first genomic data for Stone Age hunter-gathers prior to the arrival of migrants from within  
73 and outside of Africa <2 kya. These data reveal previously unknown admixture patterns for southern  
74 African indigenous groups, and push the emergence of modern humans back to >260 kya.

75  
76 We sequenced the genomes of three Stone Age hunter-gatherers and four Iron Age farmers, directly  
77 radiocarbon dated to ~2 kya and 0.5-0.3 kya respectively, to between 0.01x and 13.2x genome coverage  
78 (Figure 1, Table 1, SI 1 for archaeological contexts, and SI 2-3 for sampling and laboratory procedures).  
79 The DNA-sequence data display all features characteristic of ancient DNA (e.g.<sup>21</sup>, short DNA fragments  
80 and consistently higher cytosine deamination at fragment ends, SI 4). Five individuals were  
81 morphologically sex-assigned (SI 1), and have now been genetically confirmed (SI 5). The previously  
82 unknown sex of the Ballito Bay A juvenile has been genetically determined as male (Table 1, SI 1, 5).

83  
84 The three Stone Age individuals, Ballito Bay A Ballito Bay B, and Doonside, and an Iron Age individual  
85 from Champagne Castle, carry mitochondrial sub-haplogroups belonging to haplogroup L0d (SI 5.2),  
86 common in current-day Khoe-San populations<sup>14</sup>. The remaining three Iron Age individuals, from  
87 Newcastle, Eland Cave, and Mfongosi, have mtDNA haplogroups that fall within L3e, common to  
88 current-day Bantu-speaking groups<sup>14</sup>. Both males from Ballito Bay carry the Y chromosome A1b1b2

89 haplotype (SI Section 5.1), common among modern-day Khoe-San<sup>17</sup>.

90  
91 To assess population affinities among the ancient individuals and their relations to modern-day groups,  
92 we merged the ancient southern African genome data with published genotype datasets from southern  
93 Africa<sup>8, 9</sup>, Africa as a whole, and from across the globe (SI Section 6, Table S6.1). We further merged  
94 and investigated the ancient genome data with a set of complete genomes of 11 individuals from across  
95 the world<sup>22</sup>, including individuals from southern, eastern and western Africa (SI Section 6). Principal  
96 Component Analysis (PCA) and admixture analyses show that the three Stone Age individuals are  
97 related to present-day Khoe-San groups, specifically to southern Khoe-San, such as the Karretjie People<sup>8</sup>  
98 and the Lake Chrissie San<sup>23</sup> (Figs. 1B-C, Extended Data Figs. 1, 7, 8, Figs. S6.1-S6.5, S9.6-S9.9). The  
99 four Iron Age individuals all group with populations of West African origin/descent, specifically with  
100 southeast Bantu speakers from South Africa (Figs. 1B-C, S6.1-S6.6). This observation is consistent with  
101 archaeological evidence for the arrival of migrant Iron Age farmers of West African descent to the  
102 eastern parts of southern Africa at ~1.7 kya<sup>24</sup>.

103  
104 The Stone Age individuals form one extreme in the PCA (that separates Khoe-San individuals from all  
105 other Africans and non-Africans, Fig. 1B and S6.2). All modern-day Khoe-San are drawn towards other  
106 Africans and non-Africans compared to the ancient individuals from Ballito Bay, including Ju|'hoansi  
107 San, thus far thought to be the least affected by recent admixture<sup>8, 9</sup>. A more detailed PCA points towards  
108 Eurasian and/or east African admixture into all modern-day Khoe-San (Fig. 1B, S6.3, Extended Data  
109 Fig. 1). We tested various admixture-scenarios<sup>25</sup> into Khoe-San groups using the good-coverage, high  
110 quality, genome of Ballito Bay A (SI 6.4-6.5) and conclude that the admixture source was an (already  
111 admixed) Eurasian/East African group (Extended Data Fig. 2, Extended Data Table 1 and 2), comparable  
112 to the Amhara, today living in Ethiopia (SI Section 6.4-6.5). We estimate that Ju|'hoansi individuals  
113 received 9-14% (Extended Data Fig. 2, Extended Data Table 2, SI 6.4-6.5) admixture from this mixed  
114 East African/Eurasian group (69%/31%, Extended Data Fig. 2, and SI Section 6.5), and all Khoe-San  
115 groups show 9-22% of this admixture (Fig. 2, Extended Data Table 2, SI 6.4-6.5).

116  
117 We dated this admixture event into the Khoe-San using admixture LD decay patterns<sup>25</sup> to between 50  
118 ( $\pm 3$ ) generations ago for the Ju|'hoansi (San) and 44 ( $\pm 4$ ) generations for the Nama (Khoekhoe)  
119 corresponding to 1.5-1.3 kya (assuming 30 years/generation) (SI 6.6). The East African/Eurasian source  
120 of the admixture is particularly pronounced in herding Khoe groups such as the Nama (Extended Data  
121 Table 2, SI 6.4-6.5). Based on these results, we suggest a migration from East Africa into southern  
122 Africa, resulting in admixture with local hunter-gatherers  $\geq 1.5$  kya. This scenario is consistent with a  
123 model of herding practices being introduced from the northeast by migrating pastoralists<sup>8, 26, 27</sup>. The  
124 migration had a pronounced impact on all current Khoe-San groups, not only on the descendants of Stone  
125 Age herders, such as the Khoekhoe (Fig. 2B, Extended Data Table 2, SI 6.4-6.5). This has been an elusive  
126 result since all modern-day Khoe-San individuals display  $\geq 9\%$  recent admixture. The East  
127 African/Eurasian admixture into San and Khoekhoe groups resulted in elevated diversity in present-day  
128 Khoe-San groups (Fig 3, Extended Data Fig. 3, SI Section S7.1-S7.2). The admixture also inflates  
129 inference of past effective population sizes<sup>28, 29</sup>, where modern-day San individuals display elevated

130 effective population sizes compared to other African groups (including Ballito Bay A, Extended Data  
131 Fig. 4, SI 7.3).

132

133 To decipher early human history, we used several complementary approaches<sup>15, 25, 28</sup>, and developed a  
134 novel coalescent-based approach (the various demographic events discussed are depicted in Figure 3).  
135 We focus on the good coverage, high-quality ancient genomes, in particular the Stone Age hunter-  
136 gatherer boy, Ballito Bay A. In contrast with modern-day southern African individuals, he was  
137 unaffected by admixture with herders from East Africa<sup>8, 26, 27</sup>, Bantu-speaking farmers from West Africa,  
138 or Eurasian immigrants. Specifically, we estimated divergences between various sets of individuals  
139 based on diploid called sites of Ballito Bay A (SI Section 8) and 12 previously published high coverage  
140 modern and archaic human genomes<sup>22, 30</sup> using a coalescent-based approach (G-PhoCS,<sup>15</sup> SI 8),  
141 assuming  $1.5 \times 10^{-8}$  mutations/generation<sup>31</sup> and 30 years/generation. We estimate the split times between  
142 Ballito Bay A and modern-day individuals (excluding Ju|'hoansi) to 285-356 kya (Fig 3, Extended Data  
143 Fig. 5, SI 8), and the deepest split time to  $356 \pm 7$  kya ago for the comparison with the Mandenka of West  
144 Africa (Extended Data Fig. 5). Population split times using the admixed Ju|'hoansi instead of Ballito  
145 Bay A are on average 55,000 years younger (average: 268 kya based on Ju|'hoansi *vs* average: 323 kya  
146 based on Ballito Bay A, Extended Data Fig. 5, Table S8.1). This difference is likely caused by Ju|'hoansi  
147 carrying 9-14% admixture from East Africans/Euradians, and may also be impacted by intrinsic  
148 properties of ancient DNA<sup>21</sup>(SI section 4).

149

150 We developed a novel method (SI section 9) based on the 'concordance' approach<sup>8, 32</sup> that alleviates  
151 assumptions about past population sizes, and is robust to low levels of admixture. Briefly, assuming a  
152 general split model without migration, and picking 2 chromosomes (from two different individuals or  
153 from one individual with diploid data) from each of two subpopulations, it is possible to analytically  
154 derive the parameters of the model based on the frequencies of the 8 possible polymorphic sample  
155 configurations (assuming an infinite sites model and a known ancestral variant state). The method  
156 estimates the population split-time separately for each branch in a two-population model, resulting in  
157 two estimates of the same split. It provides the possibility to estimate, independently, the split between  
158 the Stone Age Ballito Bay A boy and other groups, using genetic data from modern-day individuals  
159 avoiding bias caused by properties of ancient DNA<sup>21</sup>, and the need for phased data. We evaluate this  
160 approach, demonstrating that split-time estimates are accurate, little affected by low levels of  
161 admixture/migration (SI section 9.1), and improve with genome coverage (SI section 9.1). For the Ballito  
162 Bay A *vs* Dinka split, the two branches are estimated to  $301 \pm 5$  kya (Ballito Bay A) and  $265 \pm 5$  kya  
163 (Dinka, Fig 3b, Extended Data Fig. 6, SI 9.1), demonstrating that the deepest split among humans is  
164  $>260$  kya; even if we base the analysis on the genetic variation in the Dinka to avoid possible impact of  
165 ancient DNA properties (SI section 9). The Ju|'hoansi *vs* Dinka split has similar split times for the two  
166 estimated branches ( $258 \pm 5$  kya, Ju|'hoansi and  $255 \pm 5$  kya, Dinka, Fig 3b Extended Data Fig. 6 and SI  
167 section 9.1), thus some 43-10 kya less than the Ballito Bay A *vs* Dinka split. This difference is likely due  
168 to the additional admixture from East Africans/Euradians into present-day Khoe-San reported above (SI  
169 6).

170

171 Cumulatively, our data show that the deepest split among modern humans occurred at >260 kya (Figure  
172 3), pushing the emergence of *H. sapiens* to beyond 260 kya. Potential additional gene-flow between  
173 southern African hunter-gatherers and other groups >2 kya would only lead to the estimates provided  
174 here being underestimates of the true population split time. This deep divergence at >260 kya is halfway  
175 to the human vs Neandertal/Denisovan split (Fig 3, SI 8-9, Table S9.1,<sup>30</sup>), and as deep as the split between  
176 the Denisovans and Neandertals<sup>4, 30</sup>. The deepest split among humans is also 2.5-3.5 times as deep as  
177 human migration out of Africa, and predates the next oldest split in human population history by ~100  
178 kya (African rainforest hunter-gatherers vs western/eastern Africans). We acknowledge that mutation-  
179 rate estimates are debated, varying based on methodology. Recently, a consensus rate of  $1.5 \times 10^{-8}$  per  
180 base pair per generation has emerged<sup>31</sup>, but this estimate might also be revised, affecting the  
181 chronological dating of events inferred by genomic data. Thus, although our assumptions on mutation  
182 rates and generation times influence the exact chronological estimates, our results notably increase the  
183 time depth for the deepest split for modern humans on a relative scale.

184  
185 Several studies point to the possibility of deep population structure among sub-Saharan groups from  
186 central and West Africa<sup>4, 33-35</sup>, but not for the Khoe-San. It is possible that some fraction of the deep split  
187 times between Ballito Bay A (and Jul'hoansi if restricted to modern-day individuals) and other modern-  
188 day sub-Saharan individuals/groups can be explained by low levels of deep structure/admixture. But,  
189 unless it is common to all non-San groups, it is unlikely to have a substantial effect on the split time  
190 estimates in this study. Interestingly, West African populations partly capture this deep population  
191 structure (Extended Data Fig 2), but it is not seen in East African groups (SI 6.5).

192  
193 The San, often represented by Jul'hoansi, has consistently been included in influential investigations on  
194 human evolutionary history as a 'non-admixed' population<sup>4, 6, 8, 10, 16, 22, 30</sup> to date the deepest splits among  
195 modern humans, to infer admixture patterns of archaic humans outside of Africa, and to study the  
196 population history of sub-Saharan Africa. Many of these inferences may be biased by the recent  
197 admixture into all San groups.

198  
199 For the reconstruction of robust population histories and origins, evidence from the archaeological, fossil  
200 and genetic records should ideally converge<sup>7</sup>. The successful genetic sequencing of the seven ancient  
201 individuals from KwaZulu-Natal contributes to the better understanding of two relatively recent events  
202 that are clear in the archaeological record. One is the migration of East African herders introducing  
203 pastoralism  $\geq 1.5$  kya to southern Africa, and their mixing with local hunter-gatherers. The other is the  
204 large-scale migration of Iron Age Bantu-speaking farmers from West Africa introducing agriculture in  
205 southern Africa. We have established, for the first time, a new pre-admixture genetic baseline for hunter-  
206 gatherers from southern Africa, based on the three oldest individuals. The genetic data for these ~2 kya  
207 individuals from KwaZulu-Natal provide a record of the people on the landscape at the time, and inform  
208 on the deep human history represented by hunter-gatherers of this region.

209  
210 The outcome of the deep divergence time for *H. sapiens*, evidenced through the Ballito Bay A boy,  
211 requires integration with the local archaeological and fossil records. The deepest split-time calculation

212 of >260 kya is consistent with the archaeological estimate for the onset of the Middle Stone Age across  
213 sub-Saharan Africa<sup>12</sup>. Southern African archaeological sites with absolute dates >160 kya are still rare,  
214 e.g., Cunene in Namibia, Border Cave, Florisbad, Kathu Pan 1, Pinnacle Point Cave 13B and  
215 Wonderwerk Cave in South Africa<sup>13</sup>. Dated Middle Stone Age contexts of ~300-250 kya are thus far  
216 limited to three layers at Florisbad<sup>36</sup>, and one at Kathu Pan 1<sup>37</sup>. At the Florisbad site in central South  
217 Africa, Middle Stone Age artefacts were found together with human remains dating to 259±35 kya<sup>36</sup>.  
218 Amongst the remains is a partial cranium, with a cranial capacity similar to that of modern humans,  
219 interpreted as representing a combination of archaic and modern characteristics<sup>36, 38</sup>. Human remains  
220 from Hoedjiespunt, South Africa of ~300-200 kya were ascribed to *H. heidelbergensis*, because although  
221 morphologically modern, they seemed larger than modern-day Africans<sup>39</sup>. These records attest to the  
222 presence of humans on the southern African landscape at the time of the earliest modern human  
223 divergence, predating 260 kya, and the fossils deserve closer morphological scrutiny. Whether the  
224 Florisbad skull represents a modern human ancestor, or an archaic form of human who contributed little  
225 or no genetic material to modern humans is an open question, as is our potential relationship with *Homo*  
226 *naledi* who also roamed the South African landscape between about 236 kya and 335 kya<sup>40</sup>. Although  
227 we do not rule out that the ancestors of KwaZulu-Natal Stone Age hunter-gatherers might have  
228 originated elsewhere in sub-Saharan Africa, or might have mixed with other groups >2 kya, we suggest  
229 that archaeological, fossil and genetic records increasingly point towards a modern human development  
230 that includes southern Africa.

231

## 232 **Materials and methods**

233 See supplementary information for full description of methods and analyses.

234

## 235 **Acknowledgements**

236 We thank Dr. Carolyn Thorpe, Dr. Gavin Whitelaw and Mudzunga Munzhedzi at the Kwa-Zulu Natal  
237 Museum for access to the museum and Luciana Simões and Mário Vicente for technical assistance. This  
238 project was supported by grants from Knut and Alice Wallenberg foundation (to MJ), the Swedish  
239 Research council (no. 642-2013-8019 MJ, and no. 621-2014-5211 to CS), the Göran Gustafsson  
240 foundation (to MJ), a Wenner-Gren foundations postdoctoral fellowship (to TG), and African Origins  
241 Platform grant from the South African National Research Foundation (to ML). Sequencing was  
242 performed at the National Genomics Infrastructure (NGI), Uppsala, and computations were performed  
243 at Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX). Data is available  
244 for download through the European Nucleotide Archive (EBI-ENA) under project xxx.

245

## 246 **Author contributions**

247 C.M.S., H.M., M.L. and M.J. conceived the study, M.L. coordinated local stakeholders and permitting,  
248 M.S. conducted morphological analysis. H.M., A.C. and H.E. sampled the material and performed DNA  
249 laboratory work, C.M.S., H.M., T.G., P.S., A.C., A.R.M. and M.J. analyzed the genomic data, M.L.  
250 provided fossil and archaeological interpretations, C.M.S., H.S. and M.L. provided historical and  
251 ethnographic interpretations, and C.M.S., M.L. and M.J. wrote the paper with input from all authors.

252

253 **Main Tables and Figures**

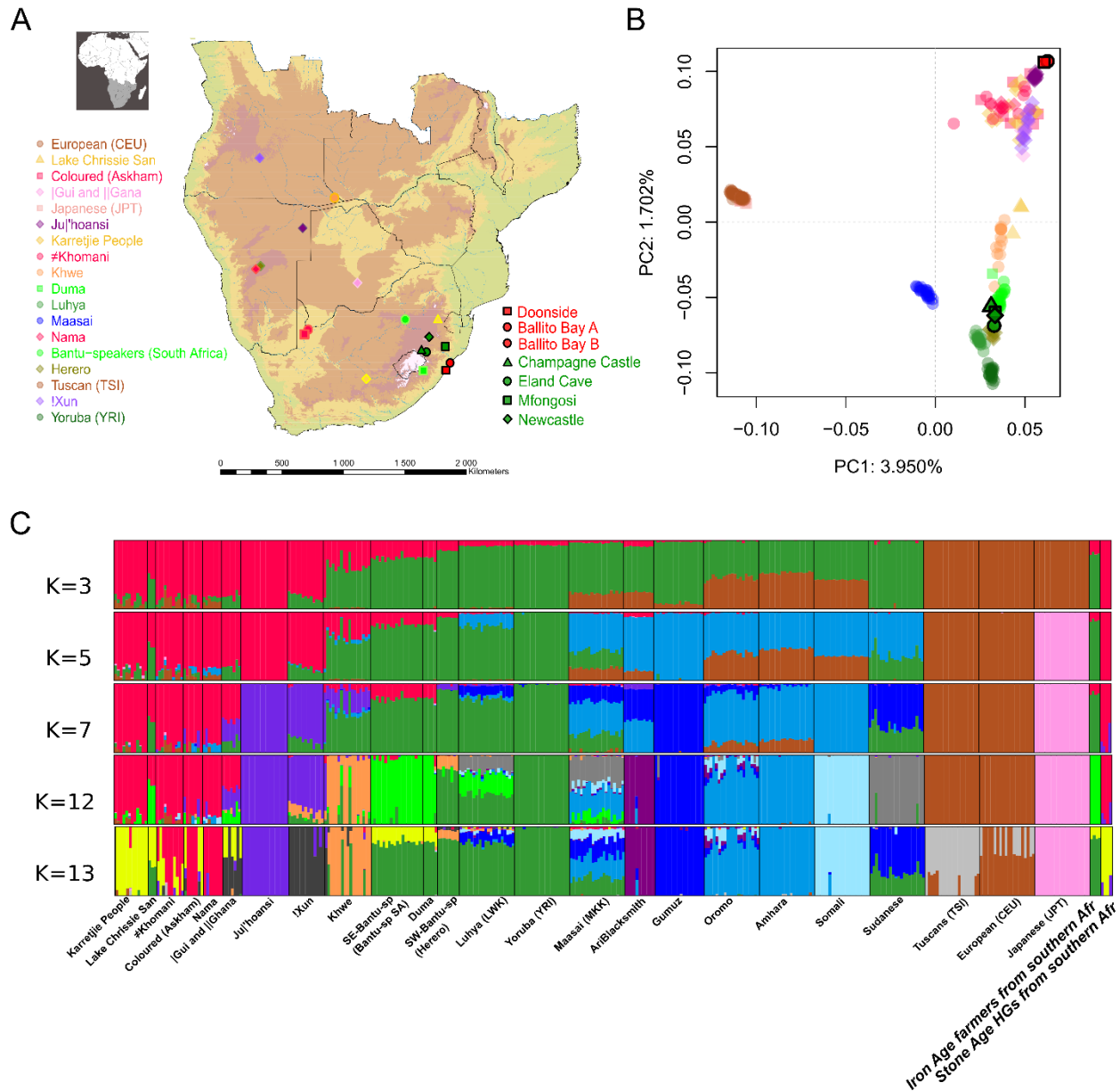
254

255 **Table 1.** Sample information. Summary table of calibrated dates, genomic and mitochondrial DNA  
256 coverage, sex estimations and mitochondrial and Y-chromosomal haplogroup assignments (see SI  
257 section 1-5 for further information). All new radiocarbon dates are reported in SI section 1.

<b>Sample</b>	<b>Cal.BP (2<math>\sigma</math>)</b>	<b>Gen. cov.</b>	<b>mt cov.</b>	<b>Biol. sex</b>	<b>mt hg</b>	<b>Y hg</b>	<b>Morph. sex</b>
Ballito Bay A	1986-1831*	12.94	1035	XY	L0d2c1	A1b1b2	juvenile
Ballito Bay B	2149-1932	1.25	84	XY	L0d2a1	A1b1b2	male
Doonside	2296-1910*	0.01	2.6	-	L0d2	-	-
Champagne Castle	448-282	0.36	186	XX	L0d2a1a	-	female
Eland Cave	533-453	13.23	7597	XX	L3e3b1	-	female
Mfongosi	448-308	6.94	562	XX	L3e1b2	-	female
Newcastle	508-327	10.65	616	XX	L3e2b1a2	-	female

258 \* Ribot et al. (2010)<sup>24</sup>





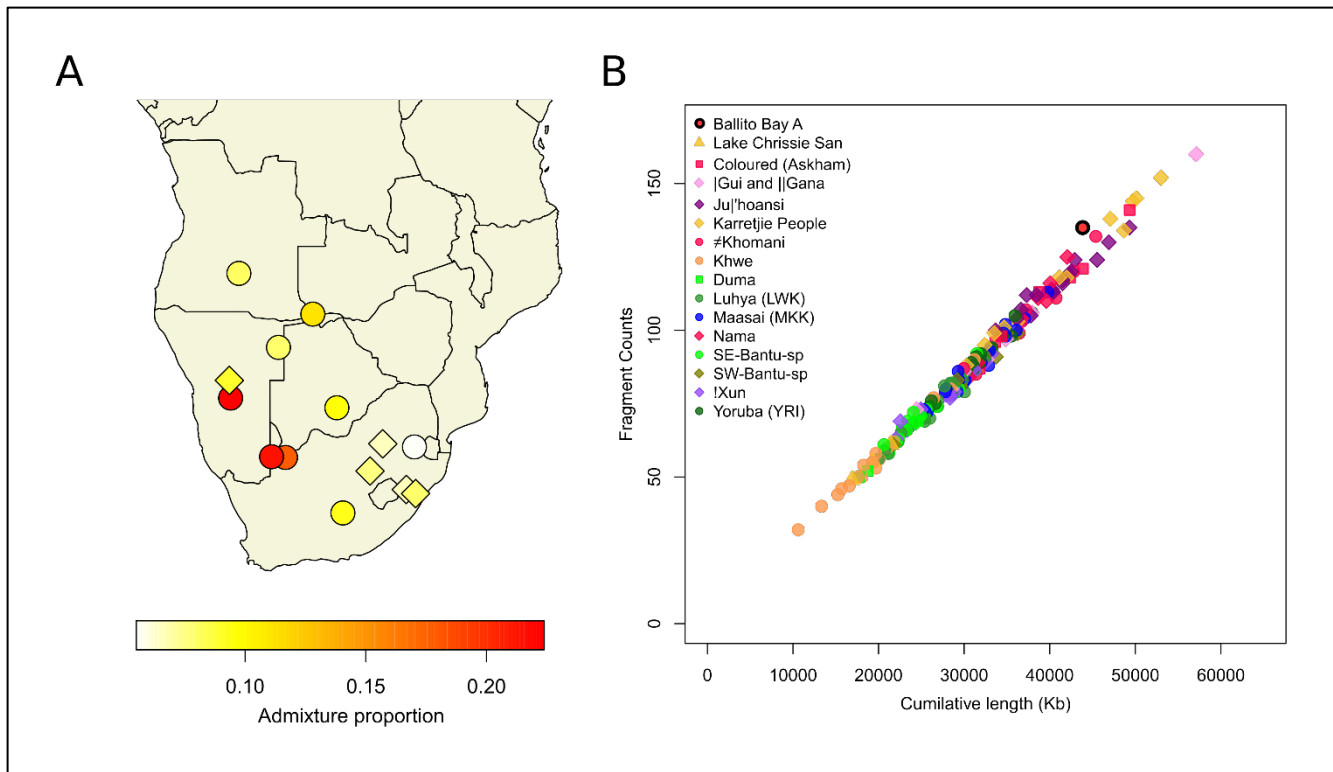
259

260 **Figure 1.** (A) Locations of the archaeological sites and the geographic centers of comparative

261 populations from Schlebusch et al<sup>8</sup>. (B) PCA with southern African, African and global comparative

262 data. (C) Admixture analysis, displayed for selected number of clusters (K), see figure S6.5 for all K

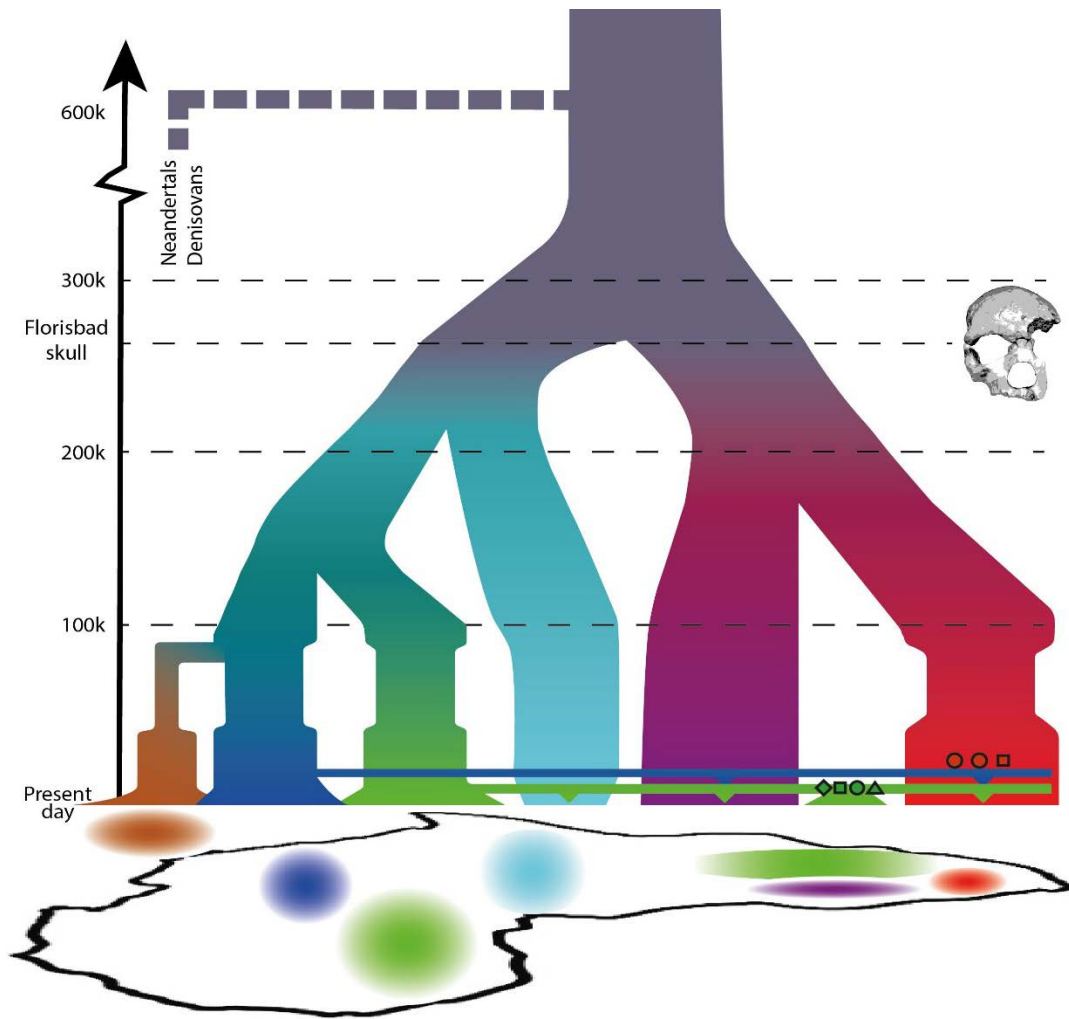
263 values.



264  
265  
266  
267  
268  
269

**Figure 2.** (A) East African/Eurasian admixture proportions (inferred with  $f_4$  ratio test using Amhara) in a southern African comparative dataset<sup>8</sup>. Circles are San and Khoekhoe populations and diamonds are Bantu-speakers. (B) Runs of Homozygosity of the 200-500 Kb bin in Africans (see fig S7.2 for inclusion of non-Africans).

270 A



271

272 B

Split method	Human-Neandertal (Nean-BBayA)	Human-Neandertal (Nean-San)	Human-Neandertal (Nean-Dinka)	Deep Human (Dinka-BBayA)	Deep Human (Dinka-San)	Deep Human (Mandenka-BBayA)	Deep Human (Mandenka-San)	NKSP-SKSP (San-BBayA)	Out of AFR (Dinka-Sardinian)
G-PhoCS	545 ± 9	534 ± 8	535 ± 9	336 ± 7	282 ± 7	356 ± 7	298 ± 7	185 ± 6	115 ± 6
TT-method	660 ± 33	639 ± 26	632 ± 28	265 ± 5	255 ± 5	256 ± 6	261 ± 5	156 ± 5	76 ± 6

273

274 **Figure 3. Demographic model of African history and estimated divergences.** (A) The estimated age of the Florisbad skull from southern Africa is indicated by a dashed line. The ancient southern African individuals in this study are denoted using the labels in Fig 1A, where the Stone Age hunter-gatherers are shown by red symbols, and Iron Age farmers of West African origin as green symbols. The depiction of population split-times, hierarchy and population sizes (width along a horizontal axis for populations) are a summary of the results presented in Extended Data Figures 1-2, 4-6. (B) Estimated population split times using the Gronau et al. approach (G-PhoCS)<sup>15</sup> and the novel Two-by-Two (TT) method described in SI 9. The deepest divergence among humans is estimated to 260-360 kya. Divergence time estimates for human vs Neandertal, and non-Africans vs Africans (indicating the out of Africa event) are given for reference, and these estimates are consistent with recent estimates in the literature<sup>4</sup>. NKSP-SKSP shows the estimated split time between Ballito Bay A and Ju|'hoansi.

284

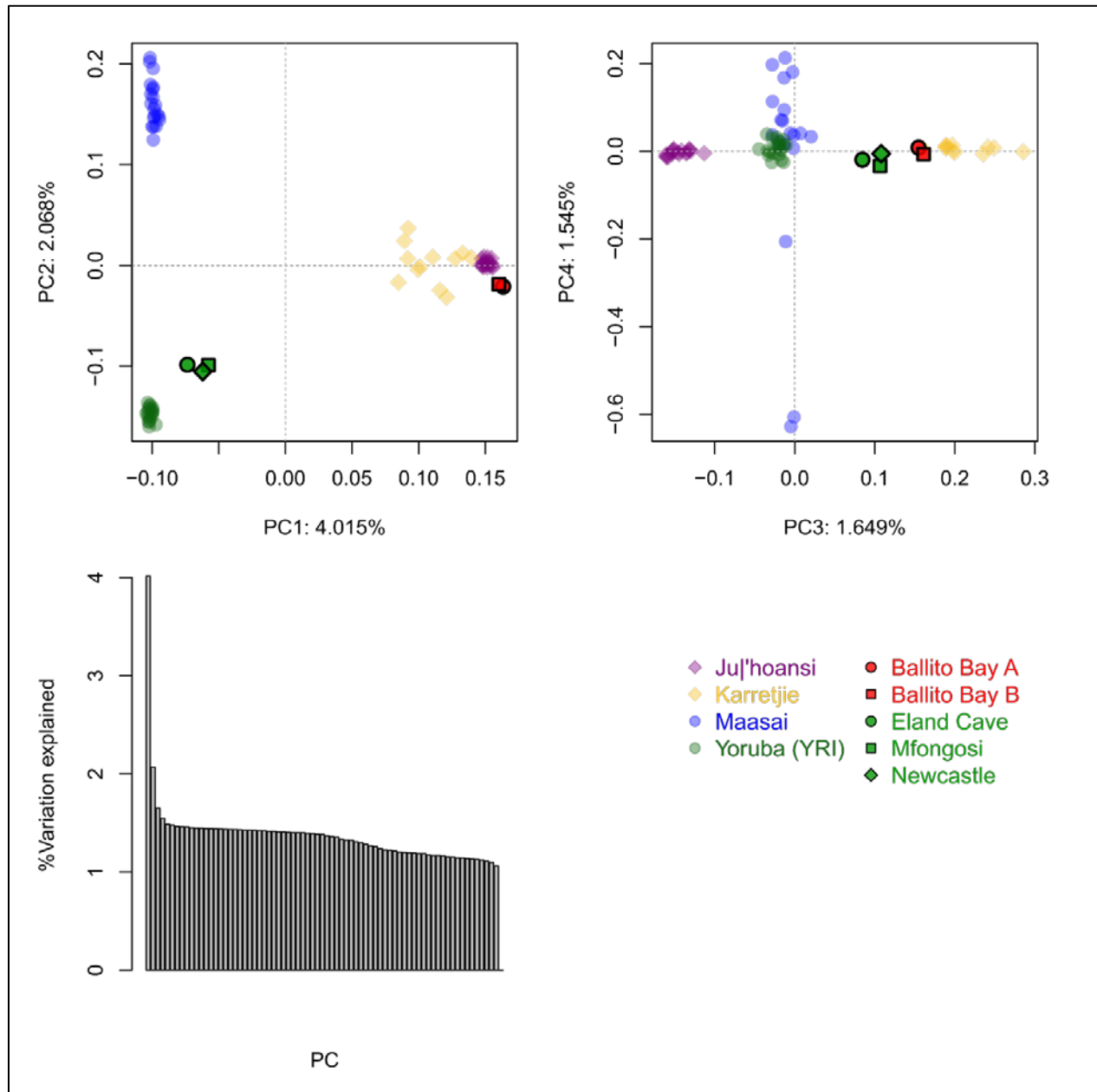
285 **References**

286

- 287 1. Cann, R. L., Stoneking, M. & Wilson, A. C. Mitochondrial DNA and human evolution. *Nature*  
288 325, 31-6 (1987).
- 289 2. White, T. D. et al. Pleistocene *Homo sapiens* from Middle Awash, Ethiopia. *Nature* 423, 742-7  
290 (2003).
- 291 3. McDougall, I., Brown, F. H. & Fleagle, J. G. Stratigraphic placement and age of modern humans  
292 from Kibish, Ethiopia. *Nature* 433, 733-6 (2005).
- 293 4. Nielsen, R. et al. Tracing the peopling of the world through genomics. *Nature* 541, 302-310  
294 (2017).
- 295 5. Jakobsson, M. et al. Genotype, haplotype and copy-number variation in worldwide human  
296 populations. *Nature* 451, 998-1003 (2008).
- 297 6. Tishkoff, S. A. et al. The genetic structure and history of Africans and African Americans.  
298 *Science* 324, 1035-44 (2009).
- 299 7. Lombard, M., Schlebusch, C. M. & Soodyall, H. Bridging disciplines to better elucidate the  
300 evolution of early *Homo sapiens* in southern Africa. *South African Journal of Science* 109, 1-8  
301 (2013).
- 302 8. Schlebusch, C. M. et al. Genomic Variation in Seven Khoe-San Groups Reveals Adaptation and  
303 Complex African History. *Science* 338, 374-379 (2012).
- 304 9. Pickrell, J. K. et al. The genetic prehistory of southern Africa. *Nat Commun* 3, 1143 (2012).
- 305 10. Henn, B. M. et al. Hunter-gatherer genomic diversity suggests a southern African origin for  
306 modern humans. *Proc Natl Acad Sci U S A* 108, 5154-62 (2011).
- 307 11. Dusseldorp, G., Lombard, M. & Wurz, S. Pleistocene *Homo* and the updated Stone Age sequence  
308 of South Africa. *South African Journal of Science* 109, 46-52 (2013).
- 309 12. Lombard, M. Thinking through the Middle Stone Age of sub-Saharan Africa. *Quaternary*  
310 *International* 270, 140-155 (2012).
- 311 13. Wadley, L. Those marvellous millennia: the Middle Stone Age of Southern Africa. *Azania:*  
312 *Archaeological Research in Africa* 50, 155-226 (2015).
- 313 14. Schlebusch, C. M., Lombard, M. & Soodyall, H. MtDNA control region variation affirms  
314 diversity and deep sub-structure in populations from Southern Africa. *BMC Evol Biol* 13, 56  
315 (2013).
- 316 15. Gronau, I., Hubisz, M. J., Gulko, B., Danko, C. G. & Siepel, A. Bayesian inference of ancient  
317 human demography from individual genome sequences. *Nat Genet* 43, 1031-4 (2011).
- 318 16. Veeramah, K. R. et al. An early divergence of KhoeSan ancestors from those of other modern  
319 humans is supported by an ABC-based analysis of autosomal resequencing data. *Mol Biol Evol*  
320 29, 617-30 (2011).
- 321 17. Barbieri, C. et al. Refining the Y chromosome phylogeny with southern African sequences. *Hum*  
322 *Genet* 135, 541-53 (2016).
- 323 18. Stringer, C. Modern human origins: progress and prospects. *Philos Trans R Soc Lond B Biol Sci*  
324 357, 563-79 (2002).
- 325 19. Högberg, A. & Lombard, M. Still Bay Point-Production Strategies at Hollow Rock Shelter and  
326 Umhlatuzana Rock Shelter and Knowledge-Transfer Systems in Southern Africa at about 80-70  
327 Thousand Years Ago. *PLoS One* 11, e0168012 (2016).
- 328 20. d'Errico, F. et al. Early evidence of San material culture represented by organic artifacts from  
329 Border Cave, South Africa. *Proc Natl Acad Sci U S A* 109, 13214-9 (2012).
- 330 21. Sawyer, S., Krause, J., Guschanski, K., Savolainen, V. & Paabo, S. Temporal patterns of  
331 nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* 7, e34131  
332 (2012).

- 333 22. Meyer, M. et al. A high-coverage genome sequence from an archaic Denisovan individual.  
334 Science 338, 222-6 (2012).
- 335 23. Schlebusch, C. M., Prins, F., Lombard, M., Jakobsson, M. & Soodyall, H. The disappearing San  
336 of southeastern Africa and their genetic affinities. Hum Genet 135, 1365-1373 (2016).
- 337 24. Ribot, I., Morris, A. G., Sealy, J. & Maggs, T. Population history and economic change in the  
338 last 2000 years in KwaZulu-Natal, RSA. Southern African Humanities 22, 89-112 (2010).
- 339 25. Patterson, N. et al. Ancient admixture in human history. Genetics 192, 1065-93 (2012).
- 340 26. Breton, G. et al. Lactase persistence alleles reveal partial East african ancestry of southern african  
341 Khoe pastoralists. Curr Biol 24, 852-8 (2014).
- 342 27. Russell, T. & Lander, F. 'What is consumed is wasted': from foraging to herding in the southern  
343 African Later Stone Age. Azania: Archaeological Research in Africa 50, 267-317 (2015).
- 344 28. Schiffels, S. & Durbin, R. Inferring human population size and separation history from multiple  
345 genome sequences. Nat Genet 46, 919-25 (2014).
- 346 29. Mazet, O., Rodriguez, W., Grusea, S., Boitard, S. & Chikhi, L. On the importance of being  
347 structured: instantaneous coalescence rates and human evolution--lessons for ancestral  
348 population size inference? Heredity (Edinb) 116, 362-71 (2016).
- 349 30. Prüfer, K. et al. The complete genome sequence of a Neanderthal from the Altai Mountains.  
350 Nature 505, 43-9 (2014).
- 351 31. Moorjani, P., Gao, Z. & Przeworski, M. Human Germline Mutation and the Erratic Evolutionary  
352 Clock. PLoS Biol 14, e2000744 (2016).
- 353 32. Wakeley, J. Coalescent Theory (Roberts & Company, Greenwood Village, CO, 2008).
- 354 33. Lachance, J. et al. Evolutionary history and adaptation from high-coverage whole-genome  
355 sequences of diverse african hunter-gatherers. Cell 150, 457-69 (2012).
- 356 34. Hammer, M. F., Woerner, A. E., Mendez, F. L., Watkins, J. C. & Wall, J. D. Genetic evidence  
357 for archaic admixture in Africa. Proc Natl Acad Sci U S A 108, 15123-8 (2011).
- 358 35. Hsieh, P. et al. Model-based analyses of whole-genome data reveal a complex evolutionary  
359 history involving archaic introgression in Central African Pygmies. Genome Res 26, 291-300  
360 (2016).
- 361 36. Grün, R. et al. Direct dating of Florisbad hominid. Nature 382, 500-1 (1996).
- 362 37. Porat, N. et al. New radiometric ages for the Fauresmith industry from Kathu Pan, southern  
363 Africa: implications for the Earlier to Middle Stone Age transition. Journal of Archaeological  
364 Science 37, 269-283 (2010).
- 365 38. Dreyer, T. F. A human skull from Florisbad, Orange Free State, with a note on the endocranial  
366 cast, by CU Ariens Kappers. Proc Koninklijke Akad Wetenschappen te Amsterdam 38, 3-12  
367 (1935).
- 368 39. Berger, L. R. & Parkington, J. E. Brief communication: a new Pleistocene hominid-bearing  
369 locality at Hoedjiespunt, South Africa. Am J Phys Anthropol 98, 601-9 (1995).
- 370 40. Berger, L. R., Hawks, J., Dirks, P. H., Elliott, M. & Roberts, E. M. Homo naledi and Pleistocene  
371 hominin evolution in subequatorial Africa. Elife 6 (2017).
- 372  
373

**Extended data:**

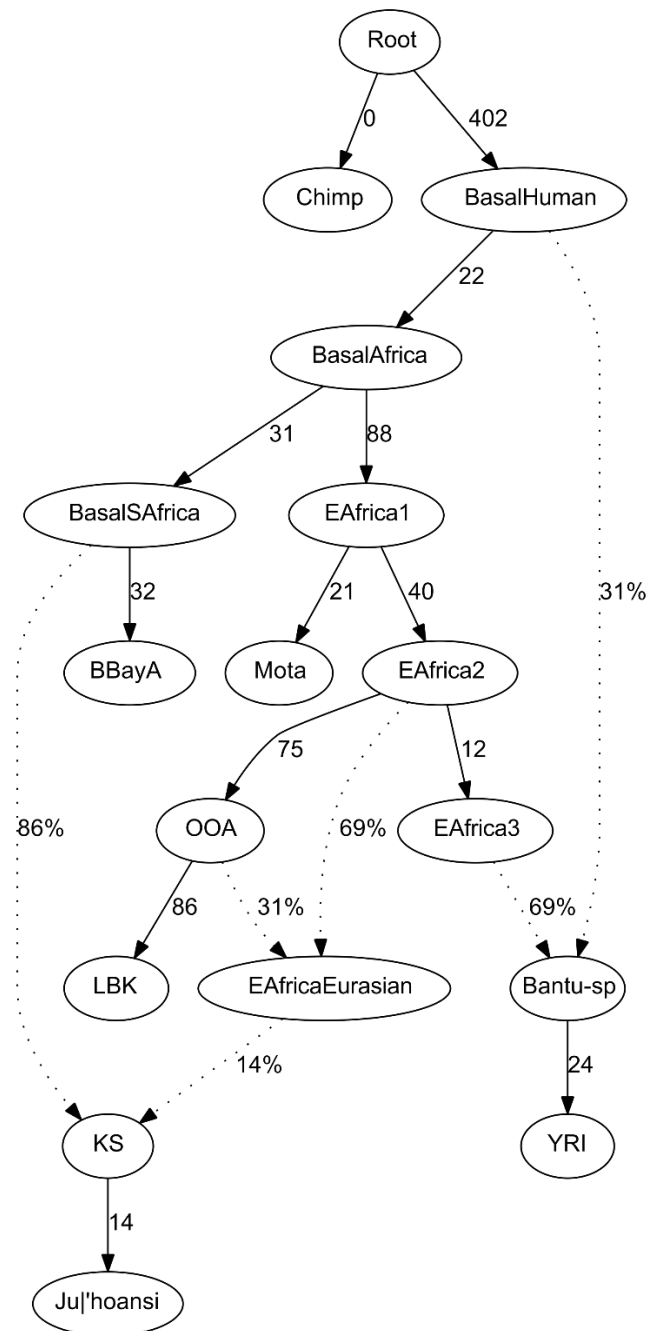


**Extended Data Fig. 1:** Principal Component analysis with comparative East and West Africans (Maasai and Yoruba) and southern and northern Khoe-San (Karretjie People and Ju|'hoansi), excluding Champagne Castle and Doonside to maximize the number of retained SNPs. (A) Ju|'hoansi is shifted towards Maasai (east Africans) compared to Ballito Bay A for PC 1 and 2, while some of the Karretjie individuals are shifted towards Maasai and some towards Yoruba (PC1 and 2). (B) Ballito Bay A and Ballito Bay B cluster with Karretjie People (southern San) and not with Ju|'hoansi (northern San) for greater PCs. The Khoe-San admixture in the Iron Age farmers also cluster with the Karretjie People (C) Bar plot displaying the amount of genetic variation explain by each PC.

**Extended Data Table 1:** Admixture into Ju|'hoansi inferred using  $f_3$  statistics.

Source	Source	Recipient	$f_3$	Std. error	Z-score
<i>BBayA</i>	Northwestern Europeans	Ju 'hoansi	-0.007914	0.000931	-8.502
<i>BBayA</i>	Tuscans (TSI)	Ju 'hoansi	-0.007569	0.000922	-8.213
<i>BBayA</i>	<i>Neolithic European</i>	Ju 'hoansi	-0.00723	0.001201	-6.02
<i>BBayA</i>	Japanese (JPT)	Ju 'hoansi	-0.005671	0.000968	-5.858
<i>BBayA</i>	Amhara	Ju 'hoansi	-0.005213	0.000782	-6.67
<i>BBayA</i>	Somali	Ju 'hoansi	-0.004683	0.000792	-5.91
<i>BBayA</i>	Oromo	Ju 'hoansi	-0.004112	0.000765	-5.372
<i>BBayA</i>	Maasai (MKK)	Ju 'hoansi	-0.002504	0.000714	-3.507
<i>BBayA</i>	Gumuz	Ju 'hoansi	-0.001467	0.000751	-1.952
<i>BBayA</i>	Sudanese	Ju 'hoansi	-0.001212	0.000726	-1.67
<i>BBayA</i>	Ari Blacksmith	Ju 'hoansi	-0.000863	0.000747	-1.154
<i>BBayA</i>	Luhya (LWK)	Ju 'hoansi	0.000431	0.000689	0.625
<i>BBayA</i>	<i>Champagne Castle</i>	Ju 'hoansi	0.000641	0.001702	0.376
<i>BBayA</i>	Yoruba (YRI)	Ju 'hoansi	0.000711	0.000703	1.011
<i>BBayA</i>	<i>Mota</i>	Ju 'hoansi	0.001396	0.001606	0.869
<i>BBayA</i>	Southwest-Bantu-sp.	Ju 'hoansi	0.001827	0.00071	2.571
<i>BBayA</i>	<i>Eland Cave</i>	Ju 'hoansi	0.00358	0.001136	3.153
<i>BBayA</i>	Khwe	Ju 'hoansi	0.003927	0.000548	7.161
<i>BBayA</i>	<i>Newcastle</i>	Ju 'hoansi	0.004109	0.00112	3.668
<i>BBayA</i>	!Xun	Ju 'hoansi	0.004707	0.0004	11.755
<i>BBayA</i>	Duma	Ju 'hoansi	0.005365	0.000724	7.405
<i>BBayA</i>	Southeast-Bantu-sp.	Ju 'hoansi	0.005947	0.000632	9.417
<i>BBayA</i>	<i>Mfongosi</i>	Ju 'hoansi	0.005979	0.001099	5.442
<i>BBayA</i>	Gui &   Gana	Ju 'hoansi	0.013675	0.000519	26.356
<i>BBayA</i>	Lake Chrissie San	Ju 'hoansi	0.014024	0.000756	18.548
<i>BBayA</i>	Nama	Ju 'hoansi	0.016387	0.000598	27.415
<i>BBayA</i>	≠Khomani	Ju 'hoansi	0.017791	0.000511	34.812
<i>BBayA</i>	Coloured (Askham)	Ju 'hoansi	0.018558	0.000559	33.183
<i>BBayA</i>	Karretjie People	Ju 'hoansi	0.021907	0.000556	39.424
<i>BBayA</i>	<i>Doonside</i>	Ju 'hoansi	0.051681	0.013006	3.974
<i>BBayA</i>	<i>Ballito Bay B</i>	Ju 'hoansi	0.054866	0.00186	29.498

To estimate whether the Ju|'hoansi received admixture from another population, we computed the  $f_3$  statistic<sup>25</sup> with Ju|'hoansi as the recipient population, the diploid Ballito Bay A as one source and other populations of the East African extended dataset as the other source population. Negative Z scores was observed for all non-Africans and East Africans (except the prehistoric Mota individual), as a source population in addition to Ballito Bay A. This shows admixture from either East Africans, Eurasians, or a mixture of the two into the Ju|'hoansi. Ancient individuals are marked in italic.



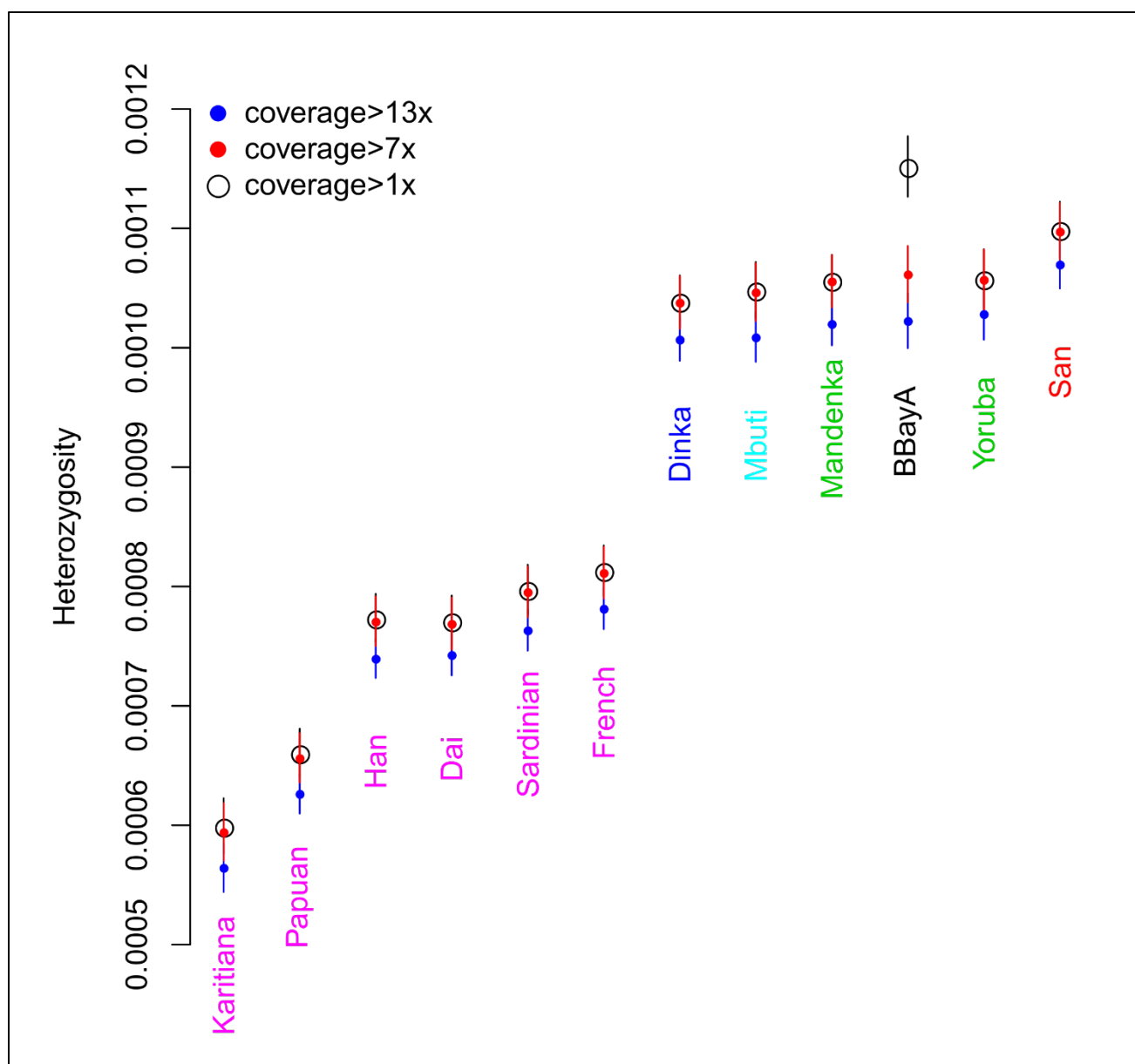
**Extended Data Fig. 2:** A population and admixture graph model of Ju|'hoansi as an admixed population between southern Africans and an admixed (Eurasian/East African) population is consistent with the data. The numbers next to edges represent the amount of drift between the nodes (multiplied by 1000). The model is including Yoruba as a potential source of Bantu-speaking ancestry. Ju|'hoansi as a Khoe-San population without admixture from Bantu-speakers is also consistent with the data. Alternative tested models constructed in a hierarchical way are discussed in SI section 6.5. The most likely model for the Ju|'hoansi is shown here.



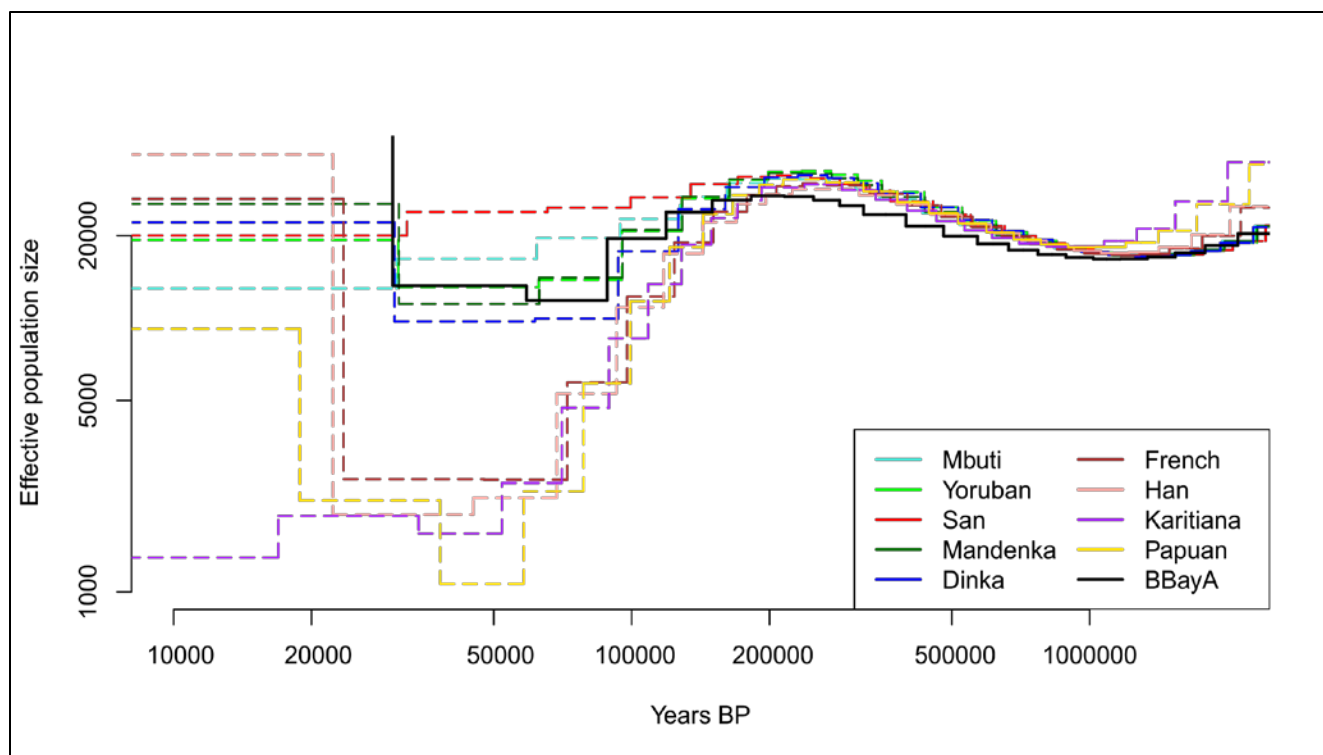
**Extended Data Table 2:**  $f_4$  ratios estimating East African/Eurasian ancestry proportions in Africans from the East African extended dataset.

<b>Target Population</b>	<b>Alpha (Amhara)</b>	<b>Std. Err. (Amahra)</b>
Somali	0.794275	0.014208
Maasai (MKK)	0.43969	0.019122
Ari Blacksmith	0.349587	0.023819
Nama	0.228031	0.019015
≠Khomani	0.216926	0.018334
Coloured (Askham)	0.179966	0.019053
Luhya (LWK)	0.126774	0.025231
Khwe	0.120816	0.021767
Gumuz	0.118447	0.027112
Southwest Bantu-sp.	0.110572	0.026251
Karretjie People	0.105278	0.019415
Gui &   Gana	0.103623	0.019771
Ju'hoansi	0.091258	0.019246
!Xun	0.089015	0.0199
Yoruba (YRI)	0.088348	0.025809
Duma	0.082081	0.025796
Southeast Bantu-sp.	0.080496	0.024114
Sudanese	0.079786	0.027334
Lake Chrissie San	0.06033	0.025744
<i>Mota</i>	0.010431	0.059915

As southern African populations contain admixture from a source of mixed East African and Eurasian ancestry, we used  $f_4$  ratios to estimate this ancestry using a modern-day east African group (Amhara) that display Eurasian admixture as source. We use the Amhara as a proxy for admixing source. Other sources from the same and other datasets are shown in SI section 6.4. The ancient individual is marked in italic.

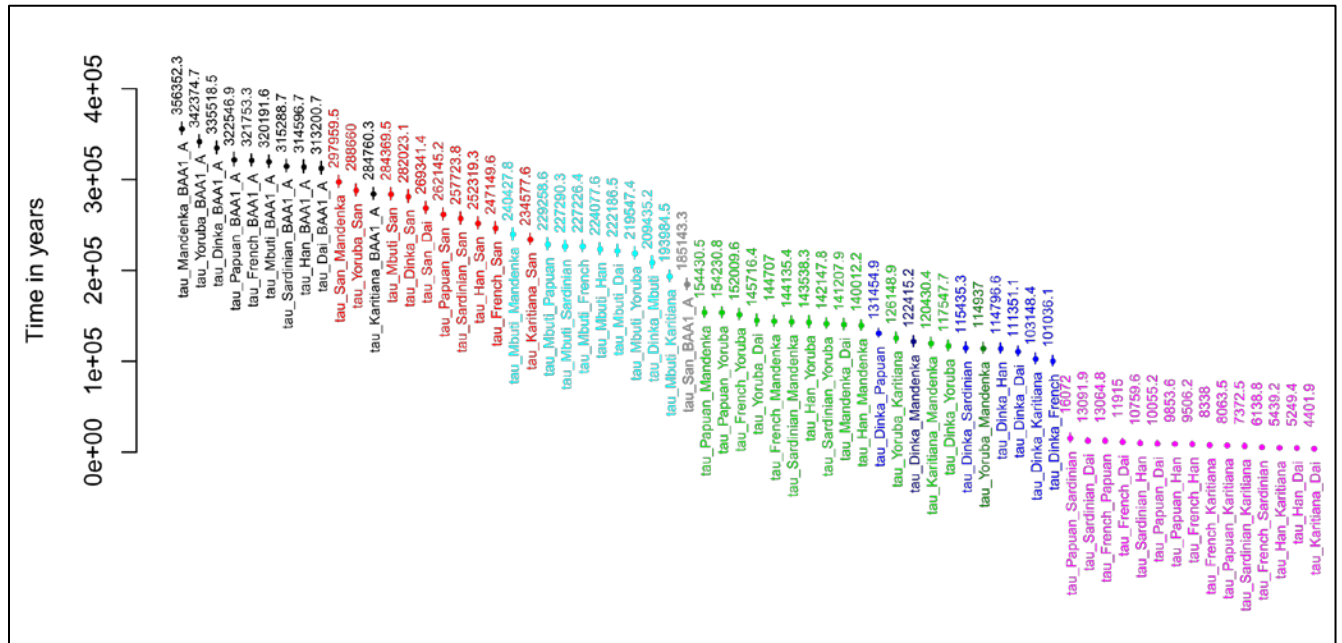


**Extended Data Fig. 3:** Heterozygosity estimates based on stringently filtered sites. Red points indicates heterozygosity based on all sites with coverage >7x, black circles on sites with coverage >1x and blue points have additional applied quality filters, including >13x (see SI section 7.1). Points are sorted in ascending order based on the more stringently filtered sites (blue points).

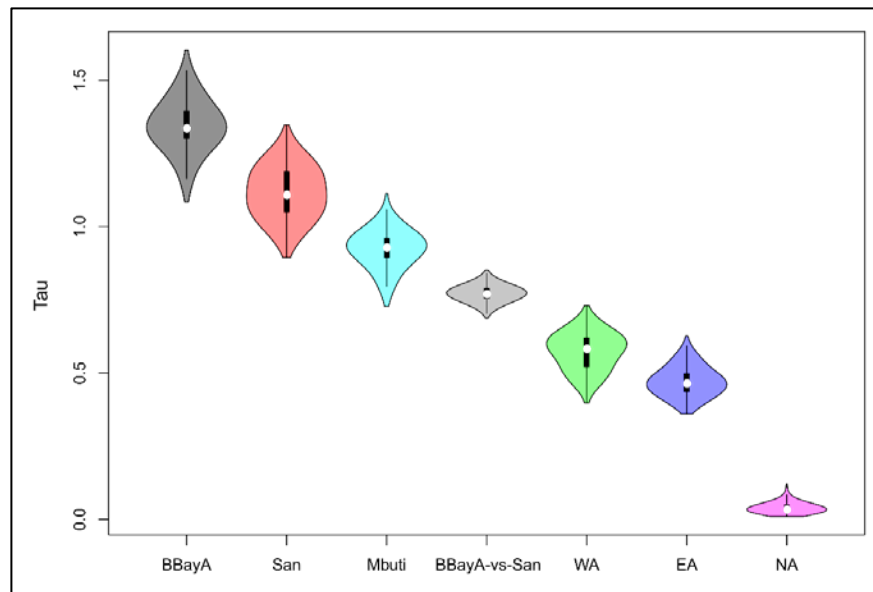


**Extended Data Fig. 4:** MSMC plot of 11 HGDP genomes together with the diploid full genome of Ballito Bay A (BBayA).

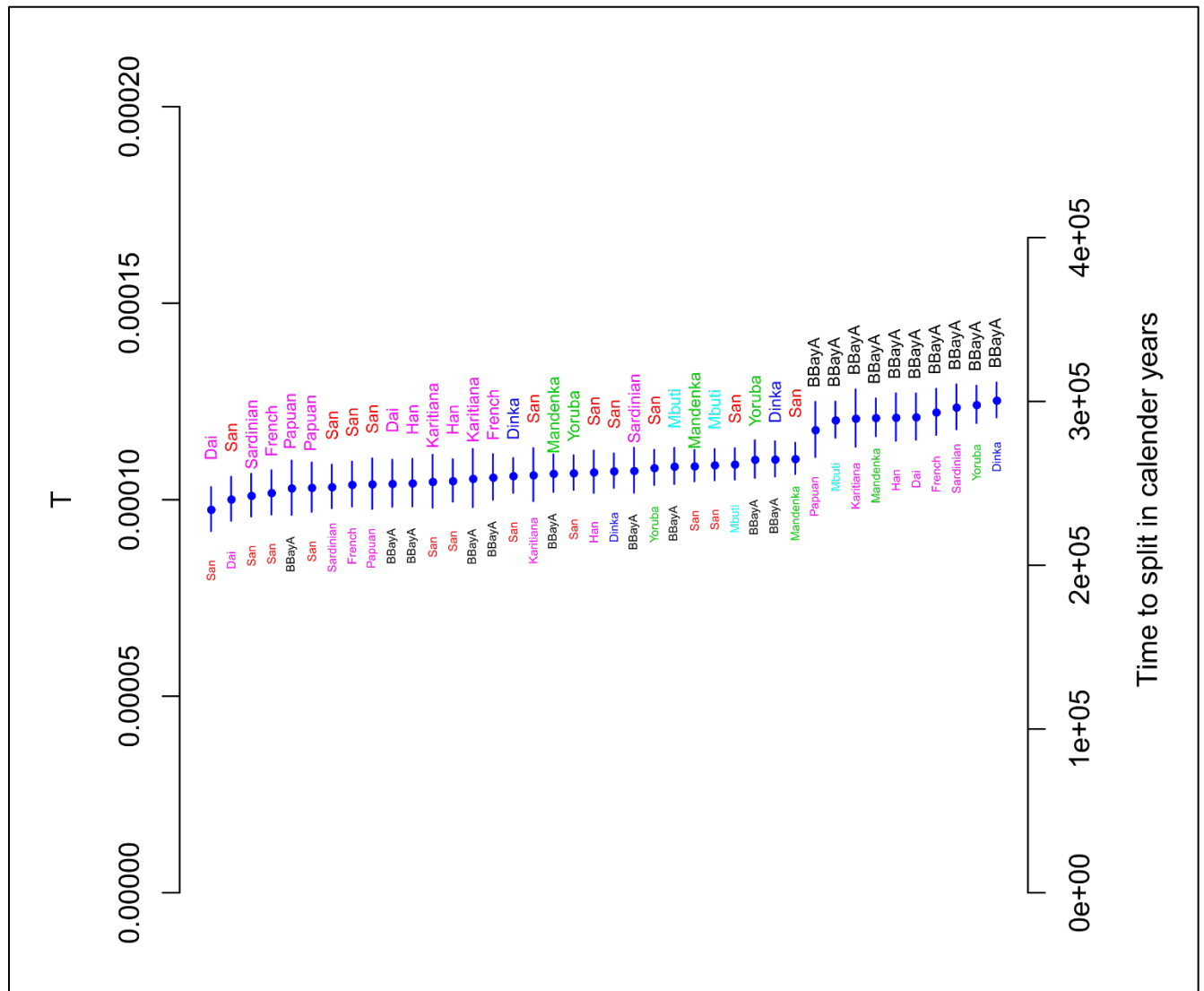
A



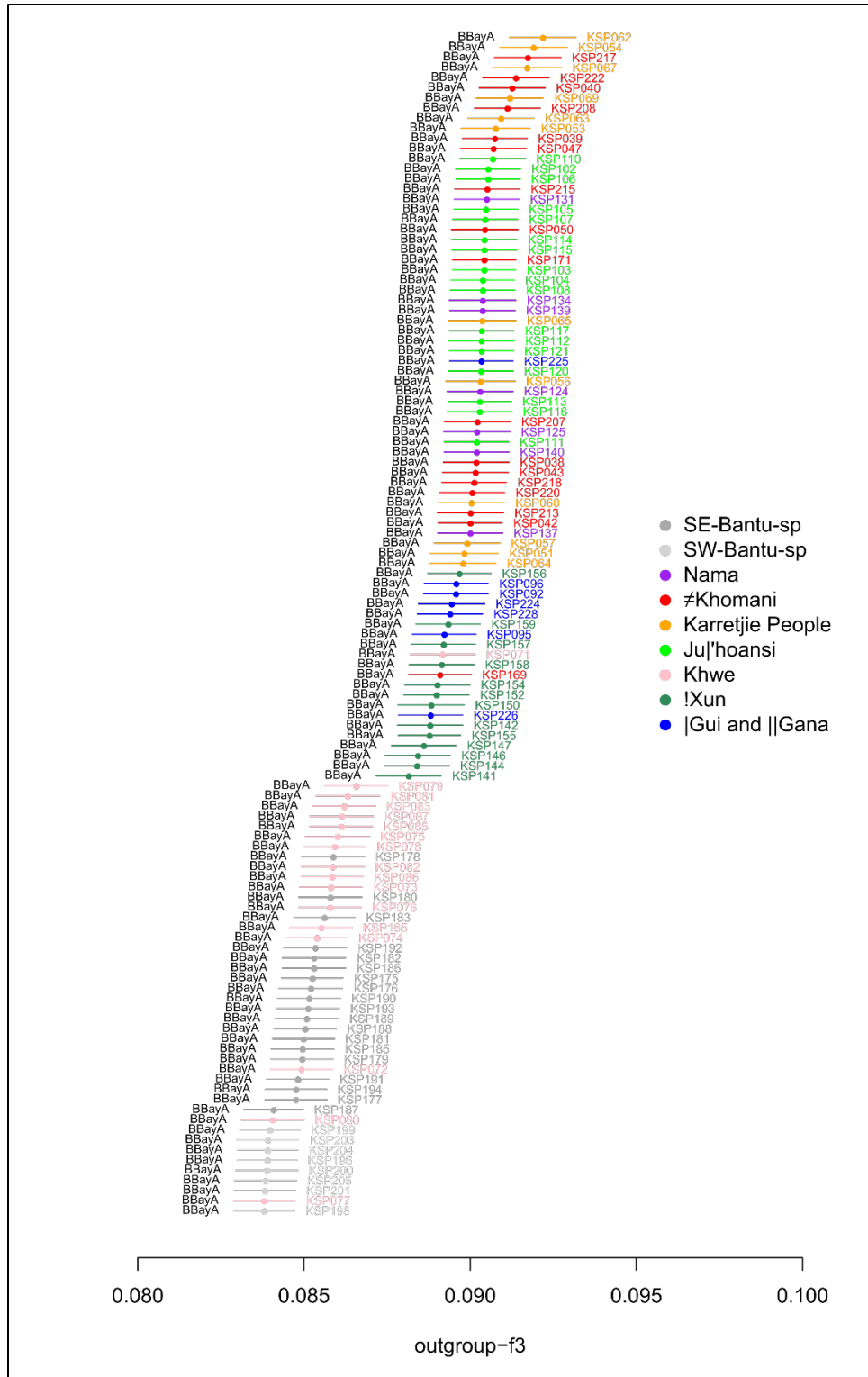
B



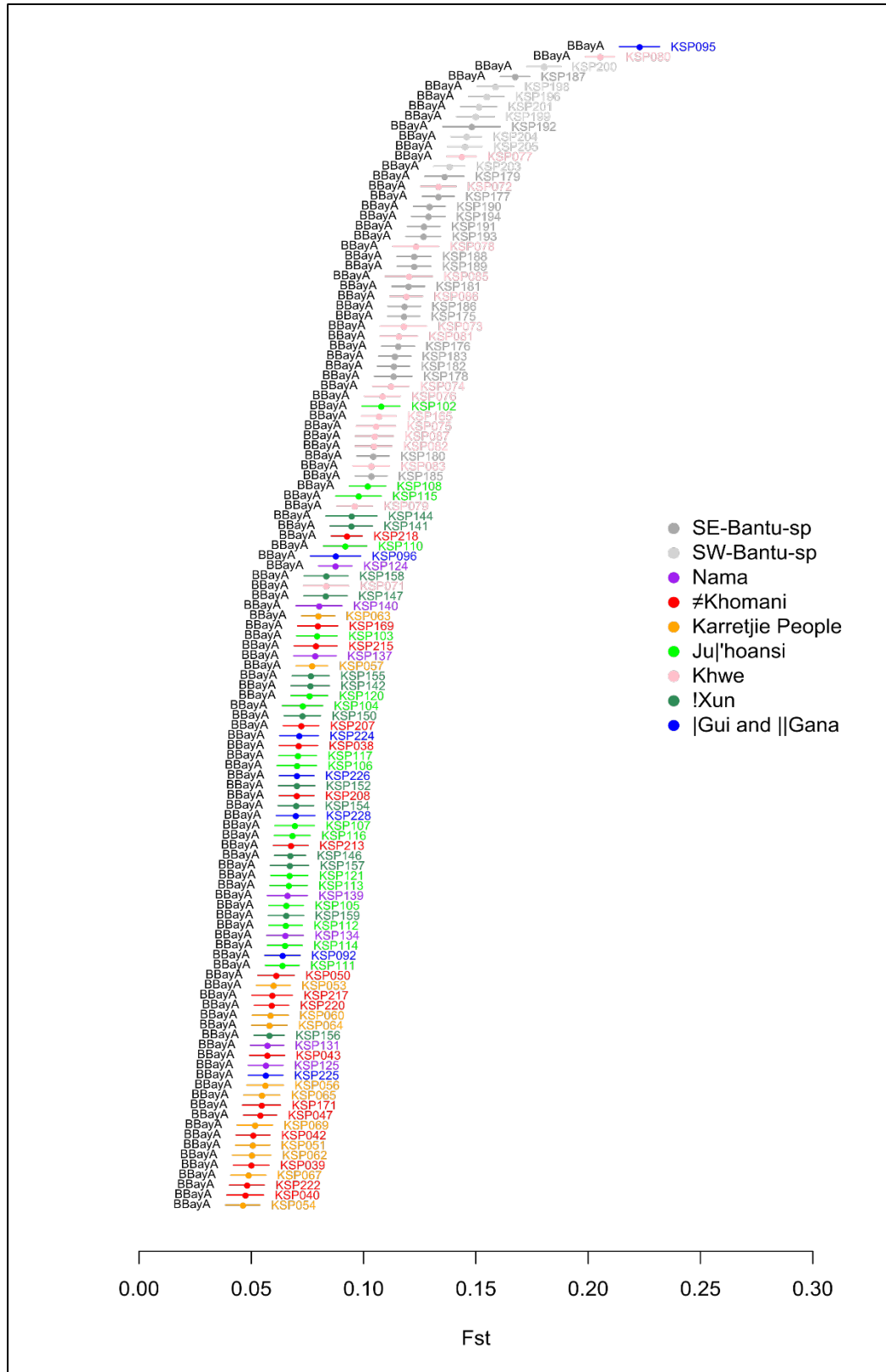
**Extended Data Fig. 5:** **A)** Means (dots) and standard deviations (bars) of G-PhoCS pairwise population split times, sorted in descending order. Colors are according to the hierarchical split times: Ballito Bay A (BAA) vs. all non-San (Black); San vs. all non-San (Red); Mbuti vs. all non-San (Turquoise); Ballito Bay A vs. San (Gray), West Africans (Mandenka and Yoruba) vs. non-Africans and East Africans (Green); East Africans vs. non-Africans (Blue); pairwise non-Africans (Pink). **B)** Same hierarchical split times shown as violin plots. Y-axis: Tau (Time in generations =  $\text{Tau} / (10,000 \times \text{mutation rate})$ ). X-Axis: Populations. West Africans (WA), East Africans (EA), and non-Africans (NA). Means and standard deviations of these grouped split times are summarized in Table S8.1



**Extended Data Fig. 6:** Estimates of split time between pairs of individuals using the TT method. The populations displayed on top and in larger font are focal populations while the populations below in smaller font are the contrasting populations. We assume a mutation rate of  $1.25 \times 10^{-8}$  per site and generation, and a generation time of 30 years to translate the estimated parameter  $T$  to time in calendar years.



**Extended Data Fig. 7:** Outgroup-f<sub>3</sub> between Ballito Bay A and Khoe-San and Bantu-speaker individuals from Schlebusch et al<sup>8</sup>. Ancestral sites are inferred from the genomes of three great apes and are used as outgroup.



**Extended Data Fig. 8:** Pairwise  $F_{ST}$  between Ballito Bay A and the individuals in Schlebusch et al.<sup>8</sup>.