



**HAL**  
open science

## Questioning Neanderthal admixture: on models, robustness and consensus in human evolution

Rémi Tournebize, Lounès Chikhi

► **To cite this version:**

Rémi Tournebize, Lounès Chikhi. Questioning Neanderthal admixture: on models, robustness and consensus in human evolution. 2024. hal-04767245

**HAL Id: hal-04767245**

**<https://hal.science/hal-04767245v1>**

Preprint submitted on 5 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Questioning Neanderthal admixture: on models, robustness  
2 and consensus in human evolution

3 Rémi Tournebize<sup>1,\*</sup> Lounès Chikhi<sup>2,1</sup>

4 April 5, 2023

5 <sup>1</sup>Instituto Gulbenkian de Ciência, Oeiras, Portugal

6  
7 <sup>2</sup>Laboratoire Évolution & Diversité Biologique (EDB UMR 5174), Université de Toulouse Midi-  
8 Pyrénées, CNRS, IRD, UPS, Toulouse Cedex 9, France

9  
10 \*Current address: Laboratoire Évolution & Diversité Biologique (EDB UMR 5174), Université de  
11 Toulouse Midi-Pyrénées, CNRS, IRD, UPS, Toulouse Cedex 9, France

12  
13 **Corresponding authors** remi.tournebize@univ-tlse3.fr, lounes.chikhi@univ-tlse3.fr

14 **Classification** Biological Sciences: anthropology

15 **Keywords** population structure, inference, robustness, admixture, human evolutionary history,  
16 Neanderthal, population genetics

## 17 Abstract

18 Genomic and ancient DNA data have revolutionized palaeoanthropology and our vision of human  
19 evolution, with indisputable landmarks like the sequencing of Neanderthal and Denisovan genomes.  
20 Yet, using genetic data to identify, date and quantify evolutionary events—like ancient bottlenecks or  
21 admixture—is not straightforward, as inferences may depend on model assumptions. In the last two  
22 decades, the idea that Neanderthals and members of the *Homo sapiens* lineage interbred has gained  
23 momentum. From the status of unlikely theory, it has reached consensus among human evolutionary  
24 biologists. This theory is mainly supported by statistical approaches that depend on demographic  
25 models minimizing or ignoring population structure, despite its widespread occurrence and the fact  
26 that when ignored, population structure can lead to infer spurious demographic events. We simulated  
27 genomic data under a structured and admixture-free model of human evolution, and found that all  
28 the tested admixture approaches identified long Neanderthal fragments in our simulated genomes  
29 and an admixture event that never took place. We also observed that several published admixture  
30 models failed to predict important empirical diversity or admixture statistics, and that our model was  
31 best at predicting these statistics jointly. Our results suggest that models accounting for population  
32 structure are fundamental to improve our understanding of human evolution, and that admixture  
33 between Neanderthals and *Homo sapiens* needs to be re-evaluated in the light of structured models.  
34 Beyond the Neanderthal case, we argue that ancient hybridization events, which are increasingly  
35 documented in many species, including with other hominins, may also benefit from such reevaluation.

## 36 Significance statement

37 The idea that Neanderthals and some ancestral *Homo sapiens* populations interbred has gained  
38 momentum in the last two decades. Yet, this theory is mainly supported by statistical approaches  
39 that assume highly simplified models of hominin evolution. A major issue is that these methods  
40 have been poorly tested in the context of population structure, despite its widespread occurrence  
41 in many vertebrate species. We simulated data under a structured model and found that all tested  
42 methods identified spurious admixture events, suggesting a lack of robustness to population struc-  
43 ture. Besides, our structured model was better at predicting several key genomic statistics than  
44 the tested admixture models. This suggests that admixture should be re-evaluated in the light of  
45 population structure, in hominins and beyond.

## 4.6 Introduction

4.7 Population structure is increasingly recognized by palaeo-anthropologists, geneticists and palaeo-  
4.8 climatologists as a fundamental feature of *Homo sapiens* (*Hs*) evolution—not only in the recent, but  
4.9 also in the ancient past (Beaumont, 2004; Harding and McVean, 2004; Ramachandran et al., 2005;  
5.0 Mazet et al., 2016; Rodríguez et al., 2018; Arredondo et al., 2021; Ragsdale et al., 2022). Periods of  
5.1 geographical contraction or expansion of populations, of increased or decreased connectivity among  
5.2 regions of the African continent have likely been at the heart of this history, possibly in relation  
5.3 with climate and habitat change during the Pleistocene (Scerri et al., 2018). Yet, the integration of  
5.4 modern and ancient genomic data into one general framework to infer *Hs* demographic history is  
5.5 among the greatest challenges facing theoretical population geneticists and palaeo-anthropologists  
5.6 (Goldstein and Chikhi, 2002; Scerri et al., 2018, 2019).

5.7 Given that our closest relatives, the African great apes, are spatially and genetically structured  
5.8 over wide regions of Africa (Prado-Martinez et al., 2013; Lester et al., 2021), it seems reasonable  
5.9 to consider that *Hs* and other hominins have been structured across different but similarly wide  
6.0 regions of Africa during significant periods of their evolution. Consequently, it seems important to  
6.1 integrate intra-continental population structure in the models used for demographic inference if we  
6.2 wish to understand and properly represent the recent evolutionary history of hominins (Scerri et al.,  
6.3 2019). Currently, most inferential approaches applied to human genetic data tend to either ignore  
6.4 population structure altogether (Li and Durbin, 2011) or assume that it is mainly significant between  
6.5 continental populations (Gutenkunst et al., 2009), with a few exceptions (Currat and Excoffier,  
6.6 2004).

6.7 The importance of model-based approaches to address these challenges has been stressed for  
6.8 several decades (Beaumont, 2004; Beaumont et al., 2002; Goldstein and Chikhi, 2002; Gerbault  
6.9 et al., 2014). Demographic models should be as simple as possible to avoid over-parameterization  
7.0 but should also capture important aspects of a necessarily complex history (Goldstein and Chikhi,  
7.1 2002; Scerri et al., 2018, 2019; Chikhi, 2023). In addition, some features of the actual demographic  
7.2 dynamics can be confounding for inferential methods, whose robustness is another often overlooked  
7.3 issue. For instance, when ignoring population structure, researchers can identify and quantify with  
7.4 great precision population size changes even when the inferred changes never occurred (Wakeley,  
7.5 1999; Beaumont, 2004; Chikhi et al., 2010; Mazet et al., 2016; Battey et al., 2020). Besides, inferential  
7.6 methods used to detect important events such as admixture or bottlenecks are more likely to be  
7.7 convincing if they can also predict realistic values for other well-established measures of human  
7.8 genetic diversity.

7.9 In the last two decades, ancient DNA (aDNA) studies have revolutionized palaeo-anthropology  
8.0 with the publication of thousands of ancient genomes, and notably, with the sequencing of ancient  
8.1 hominins like Neanderthals (*Hn*). Remarkably, researchers identified that a small but significant  
8.2 amount of DNA stretches sampled in some modern *Hs* populations were genetically closer to *Hn*  
8.3 alleles than to alleles found in other contemporaneous *Hs* populations (Green et al., 2010). This  
8.4 pattern has been interpreted as a result of admixture or interbreeding between *Hn* and some *Hs*  
8.5 populations around 50–60 thousand years ago (kya) (Sankararaman et al., 2012). While there is  
8.6 some debate regarding the tempo and mode of this archaic introgression, the current consensus is  
8.7 that some modern populations harbor Neanderthal DNA as a result of a recent admixture event. A  
8.8 few discordant voices have suggested that part of the admixture signal could be explained by spatial  
8.9 population structure leading to isolation by distance and incomplete lineage sorting (ILS) in some  
9.0 *Hs* populations (Eriksson and Manica, 2012, 2014). Others added that current admixture models  
9.1 failed to explain significant patterns of genomic diversity which can be explained without admixture  
9.2 when population structure is accounted for (Rodríguez et al., 2018). Under such admixture-free  
9.3 but structured models, the DNA tracts identified as introgressed should not be attributed to an  
9.4 actual introgression event: they are the result of shared ancestry and are inherited from a structured  
9.5 common ancestor to both *Hn* and *Hs*.

9.6 It is therefore important to ask whether admixture estimates are as robust to model departures

97 (viz. intra-continental population structure) as it has been generally argued (Peter, 2016; Racimo  
98 et al., 2015). A related and open problem is whether current models of  $Hn$ - $Hs$  admixture are com-  
99 patible among themselves and whether they can reproduce well-established genome-wide patterns  
100 of human genetic diversity and differentiation.

101 Several statistics have been developed to detect and estimate genetic admixture or introgression  
102 (Sankararaman, 2020). We can, as a first approximation, classify these statistics as frequency- or LD-  
103 based. Frequency-based statistics compare allelic composition in  $Hn$  and  $Hs$  populations to identify  
104 patterns that are unlikely deriving from ILS but are more likely resulting from a recent admixture  
105 event. LD-based statistics aim at identifying stretches of DNA shared between  $Hn$  and specific  
106  $Hs$  populations, that are too long to have been inherited independently from a common ancestor.  
107 Yet, for both frequency- or LD-based statistics, the likelihood that these shared SNPs or long DNA  
108 fragments derive from an admixture event rather than from ILS depends on the specificities of  
109 the assumed demographic, mutation (Amos, 2020) and recombination models. For instance, the  $D$   
110 statistic, basis of the ABBA/BABA test, is expected to be zero if the contribution of Neanderthals  
111 to European and African *sapiens* is the same (Green et al., 2010). This expectation is deduced from  
112 a model where Europeans and Africans are modeled as continent-wide panmictic populations since  
113 their split, tens of kya, deriving from a panmictic (non-structured) ancestral population. An implicit  
114 assumption is that departures from this simple model should not influence the value and statistical  
115 significance of  $D$  in any notable way.

116 Eriksson and Manica (2012) demonstrated that this assumption was incorrect and that the  $D$   
117 statistic was in fact not robust to spatial population structure. By simulating a one-dimensional  
118 (1D) stepping-stone model with African, Eurasian  $Hs$  and  $Hn$  demes, they estimated  $D$  values  
119 similar to those observed in real data even though they did not model admixture between  $Hn$  and  
120  $Hs$  populations. In 2014, the same authors showed that the  $DCFS$  (doubly-conditioned frequency  
121 spectrum, Yang et al. (2012)) was not robust to population structure either. Thus, two statistics  
122 designed to detect admixture could be reproduced in the absence of admixture when population  
123 structure was modeled. Since then, new admixture statistics and models have been developed,  
124 that increasingly integrate patterns of LD and haplotype information, as these were thought to  
125 provide clearer support for introgression (Racimo et al., 2015; Sankararaman, 2020). In addition, the  
126 discovery and analysis of ancient  $Hs$  specimens, like Oase1, led to the identification of high  $D$  values  
127 and long Neanderthal-like DNA stretches that were suggested as further evidence of admixture (Fu  
128 et al., 2015). Nevertheless, little effort has been made to test the robustness to population structure  
129 of more recent admixture statistics, neither in present-day nor in ancient  $Hs$  samples.

130 The work presented here aims at clarifying the role of both admixture and population structure  
131 in human evolution and is inspired by two lines of research. The first is exemplified by the seminal  
132 theoretical work developed around the structured coalescent by Herbots (1994); Wilkinson-Herbots  
133 (1998) and others (Notohara, 1990; Takahata, 1991; Hudson, 1990; Tajima, 1990), who showed that  
134 gene trees sampled in structured populations have properties that no stationary panmictic model  
135 can replicate (Wakeley, 1999; Beaumont, 2004; Chikhi et al., 2010; Mazet et al., 2016; Rodríguez  
136 et al., 2018). The second line of research follows the critical work of Eriksson and Manica (2012,  
137 2014). We extend their original model and approach to investigate the behaviour of additional  
138 frequency-based statistics, but also of LD-based admixture statistics. In particular, we show that,  
139 contrary to common claims, long fragments interpreted as introgressed from Neanderthal can be  
140 detected without any actual  $Hn$  admixture, once spatial structure is accounted for. We compare  
141 our results with those of eleven demographic models published in the literature and find that our  
142 model generally best predicts the statistics jointly. We acknowledge that admixture models could  
143 provide interesting avenues for research regarding past human evolution but predict that admixture  
144 estimates will significantly decrease as models increasingly integrate intra-continental population  
145 structure. We thus argue that admixture between  $Hs$  and  $Hn$  needs to be re-evaluated in the light  
146 of structured models.

## 147 Results

### 148 A model with intra-continental structure and without admixture: speci- 149 ficities and results

150 The demographic scenarios that we explored belong to a family of 1D stepping-stone models (Figure  
151 1 for a simplified version, Figure S1 for the full model), inspired by the original model of Eriksson and  
152 Manica (2012, 2014) and by archaeological and palaeo-anthropological literature (Scerri et al., 2018,  
153 2019). Still, our model differs from Eriksson and Manica’s (2012, 2014) in several ways (see Notes  
154 S1.1 for details). For instance, their model assumed a single African stepping-stone, whereas we  
155 found that an ancient bipartite metapopulation structure across the African continent significantly  
156 improved the prediction of several statistics (see Ragsdale et al. (2022) for a similar and independent  
157 suggestion of an ancient split in the hominin lineage). Specifically, our model assumed two sets of  
158 interconnected demes, where one ancient set of 10 demes (metapopulation  $M_A$ ) generated a second  
159 set of 10 demes (metapopulation  $M_B$ ) through a serial colonization process in a period that was  
160 allowed to vary from 500 kya to 9 Mya. Our simulations suggested that this event occurred prior to  
161 2 Mya, and that this bipartite structure may have been maintained during most of the Pleistocene,  
162 with  $F_{ST}$  values varying between 0.03 and 0.89 (Notes S6).

163 Following the formation of this bipartite African structure among the species ancestral to  $Hs$  and  
164  $Hn$ , we modeled the ancestors of  $Hn$  as a 1D stepping-stone metapopulation of 10 demes, created  
165 when individuals from the African  $M_B$  demes serially colonized Eurasia. Despite a wide prior for  
166 the split between the ancestors of  $Hs$  and  $Hn$  (from 400 kya to 1 Mya), the best scenarios tended to  
167 favor a split around 650 kya (see Notes S5). After this split, our model assumed that  $M_A$  and  $M_B$   
168 would give rise to  $Hs$ , whereas the Eurasian metapopulation,  $M_N$ , would become  $Hn$ . A later serial  
169 colonization of Eurasia by  $Hs$  was also simulated from the last  $M_B$  deme, with gene flow connecting  
170 the new  $Hs$  Eurasian metapopulation  $M_C$  with the African  $Hs$  metapopulation  $M_B$  (itself connected  
171 to the  $Hs$  metapopulation  $M_A$ ). Note however that none of the  $Hs$  metapopulations ( $M_A$ ,  $M_B$  and  
172  $M_C$ ) ever exchanged gene flow with the  $Hn$  metapopulation ( $M_N$ ). In brief, our model does not  
173 allow admixture between  $Hs$  and  $Hn$ . The period providing the best support for  $Hs$  expansion in  
174 Eurasia was estimated around 40–65 kya, in agreement with current models and archaeological data  
175 (cf. published models in Notes S8).

176 We must stress here that our objective was not to find the best model explaining the data  
177 but rather, to identify a reasonable set of demographic parameters that could reproduce important  
178 summary statistics within a wide parameter space. To do so, we used a set of ten summary statistics  
179 ( $AFS_{CEU}$ ,  $AFS_{YRI}$ ,  $F_{ST}$ ,  $D$ ,  $DCFS$ ,  $L_{S'}$ ,  $\pi_{CEU}$ ,  $\pi_{YRI}$ ,  $T_{LD}$ ,  $\sigma_{T,LD}$ , see Notes S1.4 for details)  
180 (Gutenkunst et al., 2009; Bhatia et al., 2013; Green et al., 2010; Yang et al., 2012; Browning et al.,  
181 2018; Sankararaman et al., 2012; Keinan et al., 2009) as a guide for the exploration of the parameter  
182 space (Table S2) and for the identification of the twenty most promising scenarios out of a million  
183 simulations. Some of these statistics ( $D$ ,  $DCFS$ ,  $L_{S'}$ ,  $T_{LD}$ ) are generally used to detect or quantify  
184 admixture. We additionally computed three statistics (the  $S'$ -related match rate  $m_{S'}$  (Browning  
185 et al., 2018) and two  $PSMC$  curves (Li and Durbin, 2011) for genomes sampled at present in  $M_A$   
186 and  $M_C$  demes) to see if our model could predict observed values or curves from real present-day  
187 samples.

### 188 Explaining admixture statistics without admixture

189 Despite important differences between our model and Eriksson and Manica (2012, 2014), our sim-  
190 ulations confirmed their original findings (Figure 2F), viz. that large and realistic  $D$  values can be  
191 reproduced after accounting for population structure in the absence of  $Hn$  admixture. Our model  
192 generated  $D$  values that varied from -1.2% to 10.8% (IQR95), with an average at 4.7% similar to  
193 the observed value of 4.6% (Figure 2F). Using the mean relative absolute error (MRAE) as a scaled  
194 measure of discrepancy between model predictions and observed values (Notes S1.7), we estimated

195 MRAE=52.4%±47.2% (mean ± SD). Similarly, we obtained non linear *DCFS* curves, which are  
196 typically suggested to be indicative of admixture (Yang et al., 2012) (Figure 2G).

197 Our simulations further extend the studies of Eriksson and Manica (2012, 2014) by showing that  
198 empirical LD-based statistics can also be reproduced. For instance, using the *S'* approach (Brown-  
199 ing et al., 2018), we identified large introgressed Neanderthal fragments in our simulated samples  
200 (Figure 2H). With an average length of 391 kb (IQR95=175—978 kb; MRAE=36.2%±55.3%) for  
201 the simulated genomes, which included the value computed from the observed data (354 kb), our  
202 model was thus able to explain empirical results and could generate large fragments (>500 kb).  
203 In addition, our model correctly predicted the mean empirical introgression rate (1.17%) estimated  
204 with a conditional random field (*CRF*), which fell within the IQR95 of our simulations (0.1%—1.7%;  
205 MRAE=51.9%±30.7%).

206 Another important admixture statistic,  $m_{S'}$ , estimates the match rate between the *S'*-inferred  
207 ancestral alleles and the actual Neanderthal alleles within the *S'*-identified introgressed segments.  
208 This statistic is supposed to be specific to introgressed Neanderthal fragments since the tracts are  
209 first identified statistically, and only then compared to the Neanderthal sequences. In real data,  
210 the match rate was estimated as 72.6%, within the values estimated on our simulated data (average  
211 66.4%; IQR95=44.3—83.7%; MRAE=13.3%±11.1%) (Figure 2H). We also cautiously applied the  
212 *CRF* approach to estimate the length of the Neanderthal fragments and estimated an average of  
213 34 kb in the simulated data (IQR95=28—41 kb; MRAE=70.3%±3.5%), which was lower than the  
214 value obtained on real data (114 kb, Sankararaman et al. (2014)) (Figure S3). Using the ancestry-*LD*  
215 approach, we estimated that the decay constant was 109 kya on average (assuming a generation time  
216 of 25 years; IQR95=38—189 kya, MRAE=129.8%±88.2%) (Figure 2I), which included the value of  
217 48 kya inferred from observed data (Sankararaman et al., 2012). We note however that for the  
218 simulations under our structured model,  $T_{LD}$  should reflect the simulated split time between *Hs*  
219 and *Hn* (> 415 kya for all scenarios, and ~650 kya on average for the retained models) since this  
220 is the time of the last genetic exchange between *Hs* and *Hn*. The estimated average of 109 kya is  
221 much more recent than that, questioning the classical interpretation of this statistic in the presence  
222 of population structure. In other words, these two statistics identified in our simulated *Hs* genomes:  
223 (i) a ~100 kya old admixture event that never took place, and (ii) Neanderthal fragments of ~34 kb  
224 (*CRF*), that cannot be the result of an introgression. Although our estimates based on simulated  
225 data do not perfectly match the values for the observed data, we note that when we allowed for  
226 a variable recombination rates along the genomes, the *CRF*-inferred Neanderthal fragments in the  
227 simulated data increased in length (with an average of 75 kb; IQR95: 55—88 kb) and the ancestry-  
228 *LD* decay constant decreased to 88 kya (IQR95=23—190 kya; MRAE=97.0%±94.9%), values closer  
229 to the one inferred on real data.

230 Interestingly, moving away from admixture statistics, we computed the nucleotide diversity ( $\pi$ )  
231 of our simulated Eurasian and African individuals and found that our model was able to accu-  
232 rately predict the empirical autosomal  $\pi$  previously estimated by Keinan et al. (2009) in CEU  
233 ( $\pi_{CEU,obs}=0.827 \text{ kb}^{-1}$ ,  $\pi_{CEU,sim}=0.947 \text{ kb}^{-1}$ , IQR95=0.77—1.16  $\text{kb}^{-1}$ , MRAE=16.5%±13.7%) and  
234 YRI ( $\pi_{YRI,obs}=1.081 \text{ kb}^{-1}$ ,  $\pi_{YRI,sim}=1.092 \text{ kb}^{-1}$ , IQR95=0.90—1.40  $\text{kb}^{-1}$ , MRAE=11.2%±8.8%).

235 Lastly, *PSMC* curves inferred on empirical genomes from *Hs* individuals typically exhibit two  
236 humps (Li and Durbin, 2011; Mazet et al., 2016; Arredondo et al., 2021). Even though the *PSMC*  
237 was not used as a statistics for the exploration of the parameter space and for the selection of the  
238 twenty scenarios, our sub-models were able to reproduce the humps correctly for both simulated  
239 European and Yoruba individuals (Figure 2D,E). We note that the *PSMC* curves exhibited larger  
240 fluctuations for some of these scenarios and were not always perfectly timed with the observed  
241 ones. This mismatch however represents a secondary issue given the poorer fit of the *PSMC* curves  
242 computed under the other published models (Figure 3A,B).

243 We performed additional robustness tests to account for the fact that modern datasets can  
244 typically use genomic data with different sampling strategies, coverage properties and sequencing  
245 quality. We filtered the simulated genetic data for the *S'* and *CRF* analyses so as to mimic the  
246 original analyses performed on the 1000 Genomes Project dataset with low coverage (Durbin et al.,

247 2010) (cf. Notes S1.3.2). We found that our results were consistent irrespective of the shifts in  
248 allele frequency spectra induced by low coverage (Notes S10.2), to ancestral allele specification  
249 errors (Notes S10.3), to the sampling age of the simulated Neanderthal genomes (Notes S10.5) or to  
250 pseudodiploid genotypes (Notes S10.6).

## 251 Comparing models with and without admixture

252 Using twelve statistics, we compared the performance of our structured model to that of eleven  
253 published models assuming Neanderthal admixture (Figure 3A). To this end, we simulated individual  
254 genomes of 0.6 Gbp (twenty 30-Mbp chromosomes) assuming a uniform recombination rate. Our  
255 model ranked first or second for nine statistics out of twelve, including the three that were predicted  
256 and not used for model selection (Table S4). Its worst ranking (seventh) was observed for the  $T_{LD}$   
257 statistics, which we saw above appears unreliable to date gene exchange under population structure.  
258 Several models performed poorly on statistics other than the ones originally used for model fitting  
259 (Notes S8). A total of six out of the eleven models failed to predict observed levels of genetic diversity  
260 in CEU or YRI (mean relative percentage error, MRPE, greater than 30% on either  $\pi_{CEU}$  or  $\pi_{YRI}$ ),  
261 although we allowed mutation rate to vary—from  $5 \times 10^{-9}$  to  $5 \times 10^{-8} \text{ g}^{-1} \text{ bp}^{-1}$ —to improve their ability  
262 to predict observed diversity levels. In addition to over and under-estimated nucleotide diversities,  
263 several models predicted biased  $F_{ST}$  estimates (MRPE to true value from -91% for Yang et al. (2012)  
264 model up to 87% for Gower et al. (2021)). Also, ten models predicted biased  $D$  estimates (eight  
265 overestimations, two underestimations with  $|\text{MRPE}| > 30\%$ ) with three models having  $|\text{MRPE}| > 100\%$ .  
266 For most models, the  $T_{LD}$  statistic identified a recent admixture event ( $\sim 35$ – $53$  kya), in agreement  
267 with the simulated admixture ages, suggesting that this statistic is accurate for dating admixture  
268 in models with no or very limited population structure. In one model, it estimated an interbreeding  
269 event happening less than 20 kya, a period when Neanderthals are currently thought to have been  
270 extinct for several thousands of years (Ragsdale and Gravel, 2019).

271 A set of seven models predicted match rates ( $m_{S'}$ ) lower than 60%, i.e. significantly less than the  
272 observed value of 72.6% or the mean value predicted by our model (66.4%). Five models predicted  
273  $CRF$  introgression rates at odds with the observed value of 1.17% ( $|\text{MRPE}| > 50\%$ ). Finally, all  
274 the admixture models produce  $PSMC$  curves at odds with the empirical curves for CEU and YRI  
275 individuals (Li and Durbin, 2011), even when ignoring the most ancient past prior to a million years  
276 ago (Figure 3, Notes S8).

277 We further compared all models using the mean scaled Euclidean distance (MSED) across the  
278 twelve comparative statistics, as a scaled measure of discrepancy between simulations and observed  
279 data. For each model, we retained the minimum distance for each statistic across twenty simulations  
280 accepted using the same standard procedure (see Materials and Methods). For the published models,  
281 the only parameter which was let to vary across the 50 simulations was the mutation rate, to improve  
282 the model ability to predict diversity statistics. Then, we averaged the distances over the twelve  
283 statistics to obtain a single value per model (Notes S1.6). Our structured model had a MSED  
284 of 0.44 and ranked first (Table S9). The second-best model was Schaefer et al. (2021) with a  
285 distance of 0.68, followed by Moorjani et al. (2016) (0.97). We also compared the fit of our model  
286 against the admixture models assuming now a heterogeneous landscape of recombination (Hellenthal  
287 and Stephens, 2007). Here, our model and Schaefer's et al. exhibited near-identical standardized  
288 distances (MSED=0.531 and 0.525 respectively). The third model was Moorjani et al. (2016),  
289 with a value of 0.853. When considering the average *rank* (instead of the continuous value) of the  
290 mean standardized Euclidean distances, our model always ranked first for both the uniform and the  
291 heterogeneous recombination rate simulations.

## 292 Ancient DNA admixture estimates

293 The surprisingly high levels of empirical Neanderthal ancestry estimated in some ancient Eurasian  
294 *Hs* specimens—like Oase1 (Fu et al., 2015) or the recently sequenced Bacho Kiro samples (Hajdinjak

295 et al., 2021)—, were suggested as strong evidence of *Hn* admixture into *Hs*. We wondered if our  
296 structured model could generate such patterns, too. To this end, we sampled single individuals  
297 (to mimic the rarity of aDNA specimens) in the *Hs* Eurasian metapopulation  $M_C$  (from 0 to 45  
298 kya) for the twenty accepted runs of our structured model, assuming uniform recombination. To  
299 mimic various types of aDNA genetic data, we considered three SNP sampling strategies: (i) “*All*”:  
300 using all the SNPs simulated for 0.6 Gbp genomes per sample (this would mimic a high-coverage  
301 sequencing); (ii) “*1M*”: using a random subset of one million SNPs from the *All* data (to mimic  
302 the 1240K SNP array); (iii) “*Archaic*”: a subset of ascertained SNPs from *All* mimicking the Oase1  
303 genotyping (Fu et al., 2015) (cf. Materials and Methods). When measuring the  $D$  statistic using  
304 diploid genotypes and the densest SNP sampling (*All*), we observed that spurious archaic admixture  
305 could be identified again ( $D=4.8\%$  on average and  $Z=5.3$ , IQR95:  $-0.2—10.7$ ), despite the fact  
306 that *Hs* and *Hn* never exchanged gene flow in the simulations. The  $D$  statistic was stable across  
307 the sampling ages (P-value=0.67), in agreement with the best admixture model we co-analyzed (P-  
308 value=0.50, Schaefer et al. (2021)). When we accounted for ascertainment biases (see Notes S1.8.4),  
309 the simulated  $D$  values included the empirical  $D$  estimated on several ancient specimens (Figure  
310 4). Importantly, this included Oase1 when accounting for the *Archaic* ascertainment (Figure 4), a  
311 sample often presented as unambiguously admixed. The  $D$  values for ancient simulated samples and  
312 their stability across sampling ages were robust to pseudodiploidization typically used on real data  
313 (Skoglund et al., 2012) and SNP downsampling (ANOVA: P-value=0.99), but were on average four  
314 times larger using the *Archaic* pseudodiploid data compared to the *1M* diploid data. Our analyses  
315 thus suggest that the archaic admixture interpretation for statistics computed on aDNA samples  
316 might need to be re-assessed.

317 Further, we calculated the decay constant of the single-sample ancestry- $LD$  decay curve,  $T_{LD,1}$   
318 (Moorjani et al., 2016). We observed that the curves decayed more slowly for some ancient samples  
319 compared to present-day simulated samples, in line with empirical observations. For instance, the  
320  $\sim 45$  kya old Ust’-Ishim fossil exhibits  $\sim 1.8—4.2\times$  longer segments (hence slower decay rate) than  
321 present-day genomes (Fu et al., 2014). When re-scaling the decay constant ( $T_{LD,1}$ ) in kya after  
322 adding the age of the simulated samples, all values agreed across the time series (P-value=0.17 of a  
323 linear model) and identified an admixture event that was in most cases  $<100-200$  kya, much more  
324 recent than the simulated *Hn-Hs* split times (Notes S7). Our results thus suggest that high levels  
325 of *Hn* admixture estimates and longer *Hn* segments can be inferred in aDNA samples under simple  
326 models of population structure without admixture.

## 327 Discussion

328 The sequencing of ancient genomes like those of Neanderthal and Denisova has revolutionarized  
329 palaeanthropology. While these technological feats are remarkable, several lines of evidence suggest  
330 that caution should still be made regarding the interpretation of genomic and aDNA-based studies.  
331 Twenty years ago, when the general consensus was that Neanderthals could not have interbred  
332 with *Homo sapiens*, it was noted that the then available genetic data could not be used to rule out  
333 interbreeding between *Hn* and *Hs* (Nordborg, 1998; Goldstein and Chikhi, 2002). This statement was  
334 based on the work of Nordborg (1998) who had shown by simulations that Neanderthal mtDNA data  
335 only allowed to reject panmixia between *Hs* and *Hn*. But because Nordborg’s conclusions were based  
336 on a model assuming a single panmictic population only, the need for considering structured models  
337 to study human evolution was already stressed (Goldstein and Chikhi, 2002). The simulation results  
338 presented here suggest that similar caution is still recommended today and that studies identifying  
339 admixture while ignoring intra-continental population structure should be taken with a grain of salt.

340 In the last two decades, population geneticists have learned much about (and from) structured  
341 models (Herbots, 1994; Hudson, 1990; Notohara, 1990; Wakeley, 1999; Beaumont, 2004; Edmonds  
342 et al., 2004; Currat and Excoffier, 2005, 2004; Chikhi et al., 2010, 2018; Mazet et al., 2016; Arredondo  
343 et al., 2021; Battey et al., 2020). In particular, it has been shown that when populations are  
344 structured, the interpretation of genetic data can be counter-intuitive and can lead to the spurious

345 identification of selection (Currat et al., 2006; Battey et al., 2020) or of specific demographic events  
346 like population size changes (Wakeley, 1999; Chikhi et al., 2010; Mazet et al., 2016) and admixture  
347 (Eriksson and Manica, 2012, 2014). Worse, spurious signals are more likely to be detected and  
348 estimated precisely as genetic information increases (Chikhi et al., 2010). It was thus suggested that  
349 as genomic data would become more widely available, researchers might feel more convinced of the  
350 existence of events that perhaps never happened, as exemplified with the inference of population  
351 size changes in *PSMC* curves based on whole-genome sequences (Chikhi et al., 2018). Recently,  
352 several authors have proposed alternatives to panmictic models to interpret the *PSMC* of *Hs* and  
353 *Hn* (Mazet et al., 2016; Rodríguez et al., 2018; Arredondo et al., 2021), showing that structured  
354 models—with changes in connectivity that include *Hs*, *Hn* and their ancestor—could explain major  
355 patterns of the observed human genomic diversity, including patterns unexplained by other models  
356 (see also Scerri et al. (2018, 2019)). In a similar line of research, Eriksson and Manica (2012, 2014)  
357 showed that spurious admixture signals—as measured by two statistics (*D* and *DCFS*)—could be  
358 generated when intra-continental population structure was ignored. Thus, it should not come as  
359 a surprise if admixture parameters provide powerful adjustment variables when geneticists ignore  
360 population structure. This may be particularly problematic since admixture event(s) can compensate  
361 for the limitations of models that ignore or minimize other forms of connectivity at the intra- and  
362 inter-continental level.

363 In the present study, we confirmed that the *D* statistic and the *DCFS* computed from present-  
364 day *Hs* genomes can be reproduced by a demographic model with no admixture between *Hn* and *Hs*.  
365 We extended the work of Eriksson and Manica (2012, 2014) by showing that a structured model  
366 could also explain observed LD-based statistics. We showed that long DNA fragments that are  
367 identified as introgressed by LD-based methods like *S'* or *CRF* are expected in our admixture-free  
368 model. Thus, one important result and strength of our work is to show that structured models can  
369 reproduce a large set of genetic statistics *jointly*, despite the fact that complex relationships may  
370 exist between all these statistics (Notes S3). In addition, we showed that our structured model  
371 could also predict several statistics measured in aDNA samples. Our results are thus at odds with  
372 several reviews stating that admixture statistics are robust to population structure or that *Hn-Hs*  
373 admixture is now well established (Peter, 2016; Ahlquist et al., 2021; Theunert and Slatkin, 2017;  
374 Vernot and Akey, 2014).

375 The long Neanderthal tracts in non-African modern humans that are reported in the literature  
376 can thus also be generated from shared ancestry in the context of spatial structure. Consequently,  
377 these long tracts should not be seen as more “Neanderthal” than “*sapiens*”, since they can be inherited  
378 from an ancestral structured population and maintained in both the *Hs* and *Hn* populations amidst  
379 a complex history of local changes in population sizes and connectivity. Thus, a parsimonious  
380 explanation could be that *Hn* fixed some ancestral alleles due to their historically smaller population  
381 sizes, whereas *Hs* maintained more standing ancestral variation, i.e. a larger proportion of the  
382 diversity present in the common ancestor of *Hs* and *Hn*.

383 We also evaluated how several published models of human evolution—which consider *Hn* admix-  
384 ture into *Hs*—could explain the aforementioned set of statistics. We found that the different models  
385 were not necessarily cross-compatible (e.g. admixture dating) and that several of them were not able  
386 to predict realistic values of several important measures of human genetic diversity or differentiation.  
387 When we compared the discrepancy between model predictions and observed data across twelve  
388 statistics, our model provided the highest fit (for nine out of twelve), including when we accounted  
389 for variations in recombination rates along the genome.

390 We note that our objective was not to infer the best scenario using a fully-fledged inferential  
391 approach (e.g. Beaumont et al. (2002)). Rather, we aimed at illustrating the principle that struc-  
392 tured models represent a promising avenue of research to improve our understanding of human and  
393 hominin evolution. More specific structured models might be required in the future to assess the  
394 sensitivity and specificity of yet to come admixture statistics. In particular, we have introduced here  
395 a simple 1D stepping stone model in which we found that a bipartite ancient structure was required  
396 to account for the genomic diversity currently estimated in *Hs*. We interpret this surprising result as  
397 a hint that ancient population structure should be explicitly integrated in future models of hominin

398 evolution. The parsimonious scenario mentioned above, which we find is currently the best in our  
399 comparative model set, does not necessarily mean that no admixture ever took place between *Hn*  
400 and *Hs*. What our results suggest is that, if admixture ever occurred, it is currently hard to identify  
401 using existing methods, due to their limited robustness to population structure.

402 Future work should explore these questions within more complex spatial models, or metapop-  
403 ulations integrating population extinctions (Scerri et al., 2018, 2019; Chikhi, 2023). As noted by  
404 Scerri et al. (2019), choosing meta-population over tree models requires to shift from the estimation  
405 of split times, for instance, towards the quantification of processes such as changes in connectivity,  
406 extinction or recolonization rates. Focusing on split times tends to stress differences between human  
407 groups, whereas metapopulation models might allow us to focus on the dynamics of ancient pop-  
408 ulations and may thus provide a less problematic approach to evolutionary anthropology (Chikhi,  
409 2023).

410 The issues identified here should thus not be construed to the identification of a simple technical  
411 issue, since they question how narratives are constructed in our field and how consensus is built.  
412 While admixture between *Hn* and *Hs* was considered impossible by most researchers twenty years  
413 ago, it is now difficult to even question that it took place. The story that defines humans has changed,  
414 to the point that Neanderthals are now regularly and increasingly identified as being responsible for  
415 many adaptations and diseases observed in some present-day *Hs* populations.

416 We conclude by stressing that this work does not oppose the idea of ancient interbreeding between  
417 *Hs* and *Hn*. After all, they both are hominin species with long generation times who probably  
418 diverged relatively recently and likely had overlapping distributions for thousands of years (Talamo  
419 et al., 2023). This likely provided opportunity for them to interact and possibly interbreed. Yet,  
420 our work does suggest that current statistical approaches to identify these putative events remain  
421 subject to confounding factors, and notably, to population structure. Our results thus question the  
422 power of genetic data and current admixture statistics to provide clear and unambiguous evidence in  
423 favor of ancient interbreeding. Although admixture models provide interesting avenues for research,  
424 we wonder whether Neanderthal admixture estimates might not decrease (and admixture possibly  
425 vanish) when these models better integrated intra-continent population structure. This is however  
426 an open question given the infinitely large parameter space of structured models. We thus argue that  
427 admixture between *Hs* and *Hn*, and possibly other hominins, needs to be re-evaluated in the light  
428 of population structure. Ultimately, the questions asked and the modeling framework developed in  
429 this study are important beyond humans and hominins. Population structure may have affected the  
430 inference of hybridization in many living taxa, suggesting that some hybridization or introgression  
431 results might also benefit from a statistical re-evaluation under structured models.

## 432 Materials and Methods

433 We provide here a brief outline of our approach but a detailed description can be found in the  
434 Supplementary Notes S1.

### 435 The structured model

436 We assume a set of linear stepping stones—represented in Figure 1 (simplified version) and Figure  
437 S1 (detailed)—to which we will refer as “metapopulations” throughout the manuscript, because our  
438 model differs from classical stationary stepping stones which do not account for change in the number  
439 of demes. We consider that new demes are created as new regions are colonized, starting with ten  
440 demes in the most distant past (metapopulation  $M_A$ ) towards a full extent of 30 *Hs* demes (10 in  
441 metapopulation  $M_A$ , 10 in  $M_B$ , 10 in  $M_C$ ) and ten *Hn* demes in metapopulation  $M_N$ . Details of the  
442 timings and demographic processes are provided in Notes S1.1. In short, the first metapopulation,  
443  $M_A$ , is comprised of ten demes where each deme is connected to its neighboring demes or deme.  
444 A second metapopulation,  $M_B$ , is founded sequentially, one generation at a time, from the tenth

445  $M_A$  deme. Gene flow occurs between the tenth  $M_A$  and the first  $M_B$  deme, as well as between  
446 neighbouring demes within the  $M_B$  metapopulation. A new linear metapopulation,  $M_N$ , is later  
447 founded from the tenth  $M_B$  deme, following the same process of stepping stone creation but with  
448 different deme sizes.  $M_N$  demes exchange genes among neighbours but do not exchange genes with  
449  $M_A$  or  $M_B$  demes after the foundation of the first  $M_N$  deme.  $M_N$  will evolve into  $Hn$ , whereas  
450  $M_A$  and  $M_B$  will evolve into  $Hs$ . As  $M_N$  is created, the deme sizes in  $M_A$  and  $M_B$  are allowed to  
451 change to a new value. Also, after the  $M_N$  foundation, gene flow between  $M_A$  and  $M_B$  is allowed  
452 to increase. At another later time period, a new metapopulation,  $M_C$ , is founded from the tenth  
453  $M_B$  deme to represent the colonization of Eurasia by  $Hs$ . Each creation of  $M_C$  demes starts with a  
454 multigenerational bottleneck. Gene flow occurs between the tenth  $M_B$  and the first  $M_C$  deme and  
455 between  $M_C$  neighboring demes. No gene flow is ever allowed between  $M_N$  and  $M_C$  (or  $M_A$  and  $M_B$ )  
456 demes. When  $M_C$  is founded,  $M_A$  is allowed to have different deme sizes from  $M_B$  and gene flow  
457 is allowed to change between  $M_B$  demes. Altogether, this model accounts for three foundations of  
458 metapopulations, and for changes in connectivity and deme sizes. However, it does not allow gene  
459 flow between  $M_N$  with any of the other metapopulations, i.e. no gene flow is ever allowed between  
460  $Hn$  and  $Hs$ .

461 The timing of these events, the deme sizes and migration rates between demes or metapopulations  
462 were allowed to vary within wide pre-defined ranges, consistent with the literature (Notes S1.1,  
463 Table S2). For instance, the split between  $M_N$  and  $M_B$  was allowed to occur between 400 kya and  
464 1 Mya, to account for the dating uncertainty of the divergence between  $Hs$  and  $Hn$ . Details on the  
465 other parameters, including the 25 free parameters, can be found in Notes S1.1. We note that we  
466 included three parameters which enabled the sampling of individuals in random demes within their  
467 respective metapopulation. This allowed introducing variability across runs, to avoid arbitrarily  
468 sampling individuals within specific demes of our structured model. We thus sampled 50 individuals  
469 at present in  $M_A$  (corresponding to present-day YRI) and 50 in  $M_C$  (present-day CEU). We also  
470 sampled one Neanderthal at 50 kya in  $M_N$  (Vindija33.19) (Prüfer et al., 2017).

471 A large number of simulations was produced using the coalescent simulator `msprime` 1.1.1  
472 (Hudson’s algorithm with recombination) (Kelleher et al., 2016; Baumdicker et al., 2022) and a two-  
473 step filtering strategy to minimize computation burden. For the initial exploration of the parameter  
474 space, each individual genome consisted of seven chromosomes of length 10 Mbp each. We screened  
475 the simulated data after simulating the first chromosome and excluded the runs where some statistics  
476 were clearly out of range compared to real data (see Notes S1 for details). We repeated this screening  
477 process after simulating the second chromosome. We completed the full ( $10 \times 7$  Mbp genetic data)  
478 simulations for the runs which successfully passed the two-step filtering process. Thus, from testing  
479 one million parameter combinations, we generated 10,361 genetic data sets from which the full set  
480 of statistics was computed (see below).

481 For these simulations, we considered a uniform recombination landscape with a rate of  $10^{-8}$  events  
482 per base pair per lineage per generation (Halldorsson et al., 2019) and a binary mutation model with  
483 a rate equal to  $1.2 \times 10^{-8}$  mutations per base pair per lineage per generation (Jónsson et al., 2017).  
484 We did not account for gene conversion. All analyzes assumed a generation time of 25 years (Li and  
485 Durbin, 2011).

## 486 Summary statistics

487 The simulated genetic data were summarized using a collection of statistics detailed in Notes S1.3,  
488 computed using original softwares, custom scripts and Python libraries (`scikit-allel` 1.3.2).  
489 Briefly, these include (i) within-population diversity statistics, in CEU and YRI populations: nu-  
490 cleotide diversity  $\pi$ , normalized derived allele frequency spectrum (*AFS*), *PSMC* (Li and Durbin,  
491 2011); (ii) between-population differentiation statistics, between CEU and YRI: Hudson’s  $F_{ST}$ ; (iii)  
492 “admixture” statistics for Neanderthal signals into CEU compared to YRI:  $D$ -statistics (Green et al.,  
493 2010), doubly conditioned frequency spectrum (*DCFS*) (Yang et al., 2012), mean length ( $L_{S'}$ ) and  
494 mean match rate ( $m_{S'}$ ) of the putative Neanderthal-introgressed segments inferred by  $S'$  (Brown-

495 [ing et al., 2018](#)), introgression rate ( $\alpha_{CRF}$ ) and mean length ( $L_{CRF}$ ) of the putative Neanderthal-  
496 introgressed segments inferred by a conditional random field (*CRF*) ([Sankararaman et al., 2014](#)),  
497 decay constant of the multisample ancestry-*LD* ( $T_{LD}$ ) ([Sankararaman et al., 2012](#)).

498 The corresponding empirical estimates for these summary statistics were extracted from pub-  
499 lished literature for sampling designs that matched our simulation framework. See Notes S2 for a  
500 detailed description of how these empirical estimates were obtained. Note that for some summary  
501 statistics (related to  $S'$ , *CRF* and ancestry-*LD*), we applied a filter to our simulated genetic data  
502 (minor allele count filtering), in order to mimic the type of empirical data analyzed in the original  
503 studies (see Notes S1.3.2).

## 504 Run selection

505 Among the 10,361 filtered runs, we retained a set of twenty simulations which were inferred to be  
506 the closest to the empirical data according to ten summary statistics (Notes S1.4). We estimated  
507 the global distance of each simulation to the observed data by calculating  $D^*$  defined as:

$$D^* = \sum_{a=1}^{N_s} \frac{w_a \cdot d_a}{MAD(d_a)} \quad (1)$$

508 where  $N_s$  is the total number of statistics considered ( $N_s = 10$ ),  $d_a$  is the Euclidean distance  
509 between the observed and simulated values of the (vectorial or scalar) statistics  $S_a$ ,  $w_a$  is the weight  
510 of the statistics  $S_a$  and  $MAD(\cdot)$  is the mean absolute deviation. Note that for *PSMC*,  $d_a$  is defined  
511 as the cross-correlation distance (Notes S1.4.1).

512 Based on the twenty retained runs, we then simulated genetic data with larger genome sizes (600  
513 Mbp) composed of twenty chromosomes of length 30 Mbp each. All subsequent analyses reported  
514 in this study are based on these larger datasets.

## 515 Robustness analysis

516 We assessed the robustness of our analyses to variable sources of genomic uncertainty (low coverage  
517 data leading to SNP sampling biases, ancestral allele misspecification, pseudodiploid data), to the  
518 sampling ages of Neanderthal and to the nature of the recombination maps (uniform, heterogeneous  
519 vs. empirical). Details are provided in Notes S10.

## 520 Comparison with other published models

521 We assessed the performance of our admixture-free structured model in comparison with a collection  
522 of eleven published models of human evolution which assume (i) admixture from at least one archaic  
523 species (always Neanderthals) and (ii) no or limited intra-continental population structure ([Yang  
524 et al., 2012](#); [Moorjani et al., 2016](#); [Skov et al., 2020](#); [Ragsdale and Gravel, 2019](#); [Durvasula and  
525 Sankararaman, 2020](#); [Gower et al., 2021](#); [Iasi et al., 2021](#); [Schaefer et al., 2021](#)). A few models  
526 ([Jacobs et al., 2019](#); [Kamm et al., 2020](#)) allowed for several *Hs* populations within a continent, but  
527 these were typically defined by a unique “ghost” population or by aDNA specimens, and corresponded  
528 to branches in an evolution tree where gene flow was often absent, or only represented by admixture  
529 events. Yet, given the great diversity of these published demographic models, a detailed description of  
530 each model (with visual representation) is provided in Notes S8 along with the simulation protocols  
531 and commands. For each published model (characterized by its parametric point estimates), we  
532 simulated 50 datasets by varying the mutation rates uniformly from  $5 \times 10^{-9}$  to  $5 \times 10^{-8}$   $\text{bp}^{-1}\text{g}^{-1}$  and  
533 selected the twenty runs closest to the observed data, following the same procedure as above. We

## REFERENCES

## REFERENCES

534 then compared the performance of the twelve models by calculating, for each model  $M_i$ , the minimum  
535 scaled Euclidean distance (MSED)  $D_{M_i}^*$ :

$$D_{M_i}^* = \frac{1}{N_s} \sum_{a=1}^{N_s} \frac{\min(\{d_a\}_{20})}{MAD(\{d_a\}_\Omega)} \quad (2)$$

536 where  $\min(\{d_a\}_{20})$  is the minimum of the Euclidean distances calculated for the (scalar or  
537 vectorial) statistics  $S_a$  across the twenty accepted runs of the model  $M_i$  and  $MAD(\{d_a\}_\Omega)$  is the  
538 mean absolute deviation of the distances  $d_a$  across the twenty accepted runs of *all* compared models.

## 539 Ancient DNA study

540 We investigated the values of two summary statistics ( $D$ -statistics (Green et al., 2010) and the  
541 single-sample ancestry- $LD$  decay constant  $T_{LD,1}$  (Moorjani et al., 2016)) in ancient samples that  
542 we simulated (i) under the twenty accepted runs of our structured model and (ii) Schaefer et al.'s  
543 (2021) model. Details on the simulations and analyses are provided in Notes S1.8.

## 544 Code availability

545 All the scripts used in this study (including but not limited to: model simulation and plotting,  
546 statistics calculation, run selection, model comparison, figure plotting) as well as the demes YAML-  
547 formatted demographic histories of the twenty selected runs will be publicly available in a GitHub  
548 repository (freely available once the manuscript is accepted, can also be provided to the editor and  
549 referees if requested or required).

## 550 Acknowledgements

551 We thank Olivier Mazet, Simon Boitard, Barbara Parreira, Armando Arredondo and members of  
552 the Population and Conservation Genetic group for their support and for useful discussions on  
553 this topic. We would like also to acknowledge the Bioinformatic Unit and the Informatics Team  
554 of the IGC, as well as CALMIP (project P23002) for their help and support with computational  
555 resources. We thank Olivier Mazet and Simon Boitard for their useful comments on the first version  
556 of this manuscript. LC and RT were funded by Fundação para a Ciência e Tecnologia (ref. PTDC-  
557 BIA-EVL/30815/2017). This work was also supported by the LABEX entitled TULIP (ANR-10-  
558 LABX-41 and ANR-11-IDEX-0002-02) as well as the IRP BEEG-B (International Research Project  
559 - Bioinformatics, Ecology, Evolution, Genomics and Behaviour). We acknowledge an Investissement  
560 d'Avenir grant of the Agence Nationale de la Recherche (CEBA: ANR-10-LABX-25-01).

## 561 References

- 562 Ahlquist, K. D., Bañuelos, M. M., Funk, A., Lai, J., Rong, S., Villanea, F. A., and Witt, K. E.  
563 (2021). Our Tangled Family Tree: New Genomic Methods Offer Insight into the Legacy of Archaic  
564 Admixture. *Genome Biology and Evolution*, 13(7):evab115.
- 565 Amos, W. (2020). Signals interpreted as archaic introgression appear to be driven primarily by faster  
566 evolution in Africa. *Royal Society Open Science*, 7(7):191900.

REFERENCES

REFERENCES

- 567 Arredondo, A., Mourato, B., Nguyen, K., Boitard, S., Rodríguez, W., Noûs, C., Mazet, O., and  
568 Chikhi, L. (2021). Inferring number of populations and changes in connectivity under the n-island  
569 model. *Heredity*, 126(6):896–912.
- 570 Battey, C. J., Ralph, P. L., and Kern, A. D. (2020). Space is the Place: Effects of Continuous Spatial  
571 Structure on Analysis of Population Genetic Data. *Genetics*, 215(1):193–214.
- 572 Baumdicker, F., Bisschop, G., Goldstein, D., Gower, G., Ragsdale, A. P., Tsambos, G., Zhu, S.,  
573 Eldon, B., Ellerman, E. C., Galloway, J. G., Gladstein, A. L., Gorjanc, G., Guo, B., Jeffery, B.,  
574 Kretzschmar, W. W., Lohse, K., Matschiner, M., Nelson, D., Pope, N. S., Quinto-Cortés, C. D.,  
575 Rodrigues, M. F., Saunack, K., Sellinger, T., Thornton, K., van Kemenade, H., Wohms, A. W.,  
576 Wong, Y., Gravel, S., Kern, A. D., Koskela, J., Ralph, P. L., and Kelleher, J. (2022). Efficient  
577 ancestry and mutation simulation with msprime 1.0. *Genetics*, 220(3):iyab229.
- 578 Beaumont, M. A. (2004). Recent developments in genetic data analysis: What can they tell us about  
579 human demographic history? *Heredity*, 92(5):365–379.
- 580 Beaumont, M. A., Zhang, W., and Balding, D. J. (2002). Approximate Bayesian Computation in  
581 Population Genetics. *Genetics*, 162(4):2025–2035.
- 582 Bhatia, G., Patterson, N., Sankararaman, S., and Price, A. L. (2013). Estimating and interpreting  
583 FST: The impact of rare variants. *Genome Research*, 23(9):1514–1521.
- 584 Browning, S. R., Browning, B. L., Zhou, Y., Tucci, S., and Akey, J. M. (2018). Analysis of Human  
585 Sequence Data Reveals Two Pulses of Archaic Denisovan Admixture. *Cell*, 173(1):53–61.e9.
- 586 Chikhi, L. (2023). Basato su una storia vera: come fanno i genetisti delle popolazioni umane  
587 a conoscere le storie che raccontano? In *Evoluzione umana e origini di Homo sapiens*. Accad.  
588 Lincei. Coll., Roma, Italy, g. manzi edition.
- 589 Chikhi, L., Rodríguez, W., Grusea, S., Santos, P., Boitard, S., and Mazet, O. (2018). The IICR  
590 (inverse instantaneous coalescence rate) as a summary of genomic diversity: Insights into demo-  
591 graphic inference and model choice. *Heredity*, 120(1):13–24.
- 592 Chikhi, L., Sousa, V. C., Luisi, P., Goossens, B., and Beaumont, M. A. (2010). The Confounding  
593 Effects of Population Structure, Genetic Diversity and the Sampling Scheme on the Detection and  
594 Quantification of Population Size Changes. *Genetics*, 186(3):983–995.
- 595 Currat, M. and Excoffier, L. (2004). Modern Humans Did Not Admix with Neanderthals during  
596 Their Range Expansion into Europe. *PLOS Biology*, 2(12):e421.
- 597 Currat, M. and Excoffier, L. (2005). The effect of the Neolithic expansion on European molecular  
598 diversity. *Proceedings of the Royal Society B: Biological Sciences*, 272(1564):679–688.
- 599 Currat, M., Excoffier, L., Maddison, W., Otto, S. P., Ray, N., Whitlock, M. C., and Yeaman,  
600 S. (2006). Comment on "Ongoing adaptive evolution of ASPM, a brain size determinant in  
601 Homo sapiens" and "Microcephalin, a gene regulating brain size, continues to evolve adaptively  
602 in humans". *Science (New York, N.Y.)*, 313(5784):172; author reply 172.
- 603 Durbin, R. M., Altshuler, D., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A.,  
604 Clark, A. G., Collins, F. S., De La Vega, F. M., Donnelly, P., Egholm, M., Flicek, P., Gabriel,  
605 S. B., Gibbs, R. A., Knoppers, B. M., Lander, E. S., Lehrach, H., Mardis, E. R., McVean, G. A.,  
606 Nickerson, D. A., Peltonen, L., Schafer, A. J., Sherry, S. T., Wang, J., Wilson, R. K., Gibbs,  
607 R. A., Deiros, D., Metzker, M., Muzny, D., Reid, J., Wheeler, D., Wang, J., Li, J., Jian, M., Li,  
608 G., Li, R., Liang, H., Tian, G., Wang, B., Wang, J., Wang, W., Yang, H., Zhang, X., Zheng, H.,  
609 Lander, E. S., Altshuler, D., Ambrogio, L., Bloom, T., Cibulskis, K., Fennell, T. J., Gabriel, S. B.,  
610 Jaffe, D. B., Shefler, E., Sougnez, C. L., Bentley, D. R., Gormley, N., Humphray, S., Kingsbury, Z.,  
611 Kokko-Gonzales, P., Stone, J., McKernan, K. J., Costa, G. L., Ichikawa, J. K., Lee, C. C., Sudbrak,  
612 R., Lehrach, H., Borodina, T. A., Dahl, A., Davydov, A. N., Marquardt, P., Mertes, F., Nietfeld,

REFERENCES

REFERENCES

- 613 W., Rosenstiel, P., Schreiber, S., Soldatov, A. V., Timmermann, B., Tolzmann, M., Egholm, M.,  
614 Affourtit, J., Ashworth, D., Attiya, S., Bachorski, M., Buglione, E., Burke, A., Caprio, A., Celone,  
615 C., Clark, S., Connors, D., Desany, B., Gu, L., Guccione, L., Kao, K., Keibel, A., Knowlton, J.,  
616 Labrecque, M., McDade, L., Mealmaker, C., Minderman, M., Nawrocki, A., Niazi, F., Pareja,  
617 K., Ramenani, R., Riches, D., Song, W., Turcotte, C., Wang, S., Mardis, E. R., Wilson, R. K.,  
618 Dooling, D., Fulton, L., Fulton, R., Weinstock, G., Durbin, R. M., Burton, J., Carter, D. M.,  
619 Churcher, C., Coffey, A., Cox, A., Palotie, A., Quail, M., Skelly, T., Stalker, J., Swerdlow, H. P.,  
620 Turner, D., De Witte, A., Giles, S., Gibbs, R. A., Wheeler, D., Bainbridge, M., Challis, D., Sabo,  
621 A., Yu, F., Yu, J., Wang, J., Fang, X., Guo, X., Li, R., Li, Y., Luo, R., Tai, S., Wu, H., Zheng,  
622 H., Zheng, X., Zhou, Y., Li, G., Wang, J., Yang, H., Marth, G. T., Garrison, E. P., Huang, W.,  
623 Indap, A., Kural, D., Lee, W.-P., Fung Leong, W., Quinlan, A. R., Stewart, C., Stromberg, M. P.,  
624 Ward, A. N., Wu, J., Lee, C., Mills, R. E., Shi, X., Daly, M. J., DePristo, M. A., Altshuler, D.,  
625 Ball, A. D., Banks, E., Bloom, T., Browning, B. L., Cibulskis, K., Fennell, T. J., Garimella, K. V.,  
626 Grossman, S. R., Handsaker, R. E., Hanna, M., Hartl, C., Jaffe, D. B., Kernytzky, A. M., Korn,  
627 J. M., Li, H., Maguire, J. R., McCarroll, S. A., McKenna, A., Nemes, J. C., Philippakis, A. A.,  
628 Poplin, R. E., Price, A., Rivas, M. A., Sabeti, P. C., Schaffner, S. F., Shefler, E., Shlyakhter, I. A.,  
629 Cooper, D. N., Ball, E. V., Mort, M., Phillips, A. D., Stenson, P. D., Sebat, J., Makarov, V., Ye,  
630 K., Yoon, S. C., Bustamante, C. D., Clarke, L., Flicek, P., Cunningham, F., Herrero, J., Keenen,  
631 S., Kulesha, E., Leinonen, R., McLaren, W. M., Radhakrishnan, R., Smith, R. E., Zalunin, V.,  
632 Zheng-Bradley, X., Korb, J. O., Stütz, A. M., Humphray, S., Bauer, M., Keira Cheetham, R.,  
633 Cox, T., Eberle, M., James, T., Kahn, S., Murray, L., Chakravarti, A., Ye, K., De La Vega, F. M.,  
634 Fu, Y., Hyland, F. C. L., Manning, J. M., McLaughlin, S. F., Peckham, H. E., Sakarya, O., Sun,  
635 Y. A., Tsung, E. F., Batzer, M. A., Konkel, M. K., Walker, J. A., Sudbrak, R., Albrecht, M. W.,  
636 Amstislavskiy, V. S., Herwig, R., Parkhomchuk, D. V., Sherry, S. T., Agarwala, R., Khouri, H. M.,  
637 Morgulis, A. O., Paschall, J. E., Phan, L. D., Rotmistrovsky, K. E., Sanders, R. D., Shumway,  
638 M. F., Xiao, C., McVean, G. A., Auton, A., Iqbal, Z., Lunter, G., Marchini, J. L., Moutsianas,  
639 L., Myers, S., Tumian, A., Desany, B., Knight, J., Winer, R., Craig, D. W., Beckstrom-Sternberg,  
640 S. M., Christoforides, A., The 1000 Genomes Project Consortium, Corresponding author, Steering  
641 committee, Production group: Baylor College of Medicine, BGI-Shenzhen, Broad Institute  
642 of MIT and Harvard, Illumina, Life Technologies, Max Planck Institute for Molecular Genetics,  
643 Roche Applied Science, Washington University in St Louis, Wellcome Trust Sanger Institute,  
644 Analysis group: Agilent Technologies, Baylor College of Medicine, Boston College, Brigham and  
645 Women's Hospital, Cardiff University, T. H. G. M. D., Cold Spring Harbor Laboratory, Cornell  
646 and Stanford Universities, European Bioinformatics Institute, European Molecular Biology Laboratory,  
647 Johns Hopkins University, Leiden University Medical Center, Louisiana State University,  
648 US National Institutes of Health, Oxford University, and The Translational Genomics Research  
649 Institute (2010). A map of human genome variation from population-scale sequencing. *Nature*,  
650 467(7319):1061–1073.
- 651 Durvasula, A. and Sankararaman, S. (2020). Recovering signals of ghost archaic introgression in  
652 African populations. *Science Advances*, 6(7):eaax5097.
- 653 Edmonds, C. A., Lillie, A. S., and Cavalli-Sforza, L. L. (2004). Mutations arising in the wave front  
654 of an expanding population. *Proceedings of the National Academy of Sciences*, 101(4):975–979.
- 655 Eriksson, A. and Manica, A. (2012). Effect of ancient population structure on the degree of poly-  
656 morphism shared between modern human populations and ancient hominins. *Proceedings of the*  
657 *National Academy of Sciences*, 109(35):13956–13960.
- 658 Eriksson, A. and Manica, A. (2014). The doubly conditioned frequency spectrum does not distin-  
659 guish between ancient population structure and hybridization. *Molecular Biology and Evolution*,  
660 31(6):1618–1621.
- 661 Fu, Q., Hajdinjak, M., Moldovan, O. T., Constantin, S., Mallick, S., Skoglund, P., Patterson, N.,  
662 Rohland, N., Lazaridis, I., Nickel, B., Viola, B., Prüfer, K., Meyer, M., Kelso, J., Reich, D., and

REFERENCES

REFERENCES

- 663 Pääbo, S. (2015). An early modern human from Romania with a recent Neanderthal ancestor.  
664 Nature, 524(7564):216–219.
- 665 Fu, Q., Li, H., Moorjani, P., Jay, F., Slepchenko, S. M., Bondarev, A. A., Johnson, P. L. F., Aximu-  
666 Petri, A., Prüfer, K., de Filippo, C., Meyer, M., Zwyns, N., Salazar-García, D. C., Kuzmin, Y. V.,  
667 Keates, S. G., Kosintsev, P. A., Razhev, D. I., Richards, M. P., Peristov, N. V., Lachmann, M.,  
668 Douka, K., Higham, T. F. G., Slatkin, M., Hublin, J.-J., Reich, D., Kelso, J., Viola, T. B., and  
669 Pääbo, S. (2014). Genome sequence of a 45,000-year-old modern human from western Siberia.  
670 Nature, 514(7523):445–449.
- 671 Gerbault, P., Allaby, R. G., Boivin, N., Rudzinski, A., Grimaldi, I. M., Pires, J. C., Climer Vigueira,  
672 C., Dobney, K., Gremillion, K. J., Barton, L., Arroyo-Kalin, M., Purugganan, M. D., Rubio de  
673 Casas, R., Bollongino, R., Burger, J., Fuller, D. Q., Bradley, D. G., Balding, D. J., Richerson,  
674 P. J., Gilbert, M. T. P., Larson, G., and Thomas, M. G. (2014). Storytelling and story testing in  
675 domestication. Proceedings of the National Academy of Sciences, 111(17):6159–6164.
- 676 Goldstein, D. B. and Chikhi, L. (2002). HUMAN MIGRATIONS AND POPULATION STRUC-  
677 TURE: What We Know and Why it Matters. Annual Review of Genomics and Human Genetics,  
678 3(1):129–152.
- 679 Gower, G., Picazo, P. I., Fumagalli, M., and Racimo, F. (2021). Detecting adaptive introgression in  
680 human evolution using convolutional neural networks. eLife, 10:e64669.
- 681 Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li,  
682 H., Zhai, W., Fritz, M. H.-Y., Hansen, N. F., Durand, E. Y., Malaspinas, A.-S., Jensen, J. D.,  
683 Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H. A., Good, J. M., Schultz, R.,  
684 Aximu-Petri, A., Butthof, A., Höber, B., Höffner, B., Siegemund, M., Weihmann, A., Nusbaum,  
685 C., Lander, E. S., Russ, C., Novod, N., Affourtit, J., Egholm, M., Verna, C., Rudan, P., Brajkovic,  
686 D., Kucan, Ž., Gušić, I., Doronichev, V. B., Golovanova, L. V., Lalueza-Fox, C., de la Rasilla,  
687 M., Fortea, J., Rosas, A., Schmitz, R. W., Johnson, P. L. F., Eichler, E. E., Falush, D., Birney,  
688 E., Mullikin, J. C., Slatkin, M., Nielsen, R., Kelso, J., Lachmann, M., Reich, D., and Pääbo, S.  
689 (2010). A Draft Sequence of the Neandertal Genome. Science, 328(5979):710–722.
- 690 Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., and Bustamante, C. D. (2009). Inferring  
691 the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency  
692 Data. PLOS Genetics, 5(10):e1000695.
- 693 Hajdinjak, M., Mafessoni, F., Skov, L., Vernot, B., Hübner, A., Fu, Q., Essel, E., Nagel, S., Nickel,  
694 B., Richter, J., Moldovan, O. T., Constantin, S., Endarova, E., Zahariev, N., Spasov, R., Welker,  
695 F., Smith, G. M., Sinet-Mathiot, V., Paskulin, L., Fewlass, H., Talamo, S., Rezek, Z., Sirakova,  
696 S., Sirakov, N., McPherron, S. P., Tsanova, T., Hublin, J.-J., Peter, B. M., Meyer, M., Skoglund,  
697 P., Kelso, J., and Pääbo, S. (2021). Initial Upper Palaeolithic humans in Europe had recent  
698 Neanderthal ancestry. Nature, 592(7853):253–257.
- 699 Halldorsson, B. V., Palsson, G., Stefansson, O. A., Jonsson, H., Hardarson, M. T., Eggertsson,  
700 H. P., Gunnarsson, B., Oddsson, A., Halldorsson, G. H., Zink, F., Gudjonsson, S. A., Frigge,  
701 M. L., Thorleifsson, G., Sigurdsson, A., Stacey, S. N., Sulem, P., Masson, G., Helgason, A.,  
702 Gudbjartsson, D. F., Thorsteinsdottir, U., and Stefansson, K. (2019). Characterizing mutagenic  
703 effects of recombination through a sequence-level genetic map. Science, 363(6425):eaau1043.
- 704 Harding, R. M. and McVean, G. (2004). A structured ancestral population for the evolution of  
705 modern humans. Current Opinion in Genetics & Development, 14(6):667–674.
- 706 Hellenthal, G. and Stephens, M. (2007). msHOT: Modifying Hudson’s ms simulator to incorporate  
707 crossover and gene conversion hotspots. Bioinformatics, 23(4):520–521.
- 708 Herbots, HMJD. (1994). Stochastic models in population genetics : Genealogy and genetic differ-  
709 entiation in structured populations.

REFERENCES

REFERENCES

- 710 Hudson, R. R. (1990). Gene genealogies and the coalescent process. Oxford surveys in evolutionary  
711 biology, 7(1):44.
- 712 Iasi, L. N. M., Ringbauer, H., and Peter, B. M. (2021). An Extended Admixture Pulse Model Reveals  
713 the Limitations to Human–Neandertal Introgression Dating. Molecular Biology and Evolution,  
714 38(11):5156–5174.
- 715 Jacobs, G. S., Hudjashov, G., Saag, L., Kusuma, P., Darusallam, C. C., Lawson, D. J., Mondal, M.,  
716 Pagani, L., Ricaut, F.-X., Stoneking, M., Metspalu, M., Sudoyo, H., Lansing, J. S., and Cox, M. P.  
717 (2019). Multiple Deeply Divergent Denisovan Ancestries in Papuans. Cell, 177(4):1010–1021.e32.
- 718 Jónsson, H., Sulem, P., Kehr, B., Kristmundsdottir, S., Zink, F., Hjartarson, E., Hardarson, M. T.,  
719 Hjorleifsson, K. E., Eggertsson, H. P., Gudjonsson, S. A., Ward, L. D., Arnadottir, G. A., Hel-  
720 gason, E. A., Helgason, H., Gylfason, A., Jonasdottir, A., Jonasdottir, A., Rafnar, T., Frigge,  
721 M., Stacey, S. N., Th. Magnusson, O., Thorsteinsdottir, U., Masson, G., Kong, A., Halldorsson,  
722 B. V., Helgason, A., Gudbjartsson, D. F., and Stefansson, K. (2017). Parental influence on human  
723 germline de novo mutations in 1,548 trios from Iceland. Nature, 549(7673):519–522.
- 724 Kamm, J., Terhorst, J., Durbin, R., and Song, Y. S. (2020). Efficiently Inferring the Demo-  
725 graphic History of Many Populations With Allele Count Data. Journal of the American Statistical  
726 Association, 115(531):1472–1487.
- 727 Keinan, A., Mullikin, J. C., Patterson, N., and Reich, D. (2009). Accelerated genetic drift on  
728 chromosome X during the human dispersal out of Africa. Nature Genetics, 41(1):66–70.
- 729 Kelleher, J., Etheridge, A. M., and McVean, G. (2016). Efficient Coalescent Simulation and Ge-  
730 nealogical Analysis for Large Sample Sizes. PLOS Computational Biology, 12(5):e1004842.
- 731 Lester, J. D., Vigilant, L., Gratton, P., McCarthy, M. S., Barratt, C. D., Dieguez, P., Agbor,  
732 A., Álvarez-Varona, P., Angedakin, S., Ayimisin, E. A., Bailey, E., Bessone, M., Brazzola, G.,  
733 Chancellor, R., Cohen, H., Danquah, E., Deschner, T., Egbe, V. E., Eno-Nku, M., Goedmakers, A.,  
734 Granjon, A.-C., Head, J., Hedwig, D., Hernandez-Aguilar, R. A., Jeffery, K. J., Jones, S., Junker,  
735 J., Kadam, P., Kaiser, M., Kalan, A. K., Kehoe, L., Kienast, I., Langergraber, K. E., Lapuente, J.,  
736 Laudisoit, A., Lee, K., Marrocoli, S., Mihindou, V., Morgan, D., Muhanguzi, G., Neil, E., Nicholl,  
737 S., Orbell, C., Ormsby, L. J., Pacheco, L., Piel, A., Robbins, M. M., Rundus, A., Sanz, C., Sciaky,  
738 L., Siaka, A. M., Städele, V., Stewart, F., Tagg, N., Ton, E., van Schijndel, J., Vyalengerera,  
739 M. K., Wessling, E. G., Willie, J., Wittig, R. M., Yuh, Y. G., Yurkiw, K., Zuberbuehler, K.,  
740 Boesch, C., Köhl, H. S., and Arandjelovic, M. (2021). Recent genetic connectivity and clinal  
741 variation in chimpanzees. Communications Biology, 4(1):1–11.
- 742 Li, H. and Durbin, R. (2011). Inference of human population history from individual whole-genome  
743 sequences. Nature, 475(7357):493–496.
- 744 Mazet, O., Rodríguez, W., Grusea, S., Boitard, S., and Chikhi, L. (2016). On the importance  
745 of being structured: Instantaneous coalescence rates and human evolution—lessons for ancestral  
746 population size inference? Heredity, 116(4):362–371.
- 747 Moorjani, P., Sankararaman, S., Fu, Q., Przeworski, M., Patterson, N., and Reich, D. (2016). A  
748 genetic method for dating ancient genomes provides a direct estimate of human generation interval  
749 in the last 45,000 years. Proceedings of the National Academy of Sciences, 113(20):5652–5657.
- 750 Nordborg, M. (1998). On the probability of Neanderthal ancestry. American Journal of Human  
751 Genetics, 63(4):1237–1240.
- 752 Notohara, M. (1990). The coalescent and the genealogical process in geographically structured  
753 population. Journal of Mathematical Biology, 29(1):59–75.
- 754 Peter, B. M. (2016). Admixture, Population Structure, and F-Statistics. Genetics, 202(4):1485–1501.

REFERENCES

REFERENCES

- 755 Prado-Martinez, J., Sudmant, P. H., Kidd, J. M., Li, H., Kelley, J. L., Lorente-Galdos, B., Veeramah,  
756 K. R., Woerner, A. E., O'Connor, T. D., Santpere, G., Cagan, A., Theunert, C., Casals, F.,  
757 Laayouni, H., Munch, K., Hobolth, A., Halager, A. E., Malig, M., Hernandez-Rodriguez, J.,  
758 Hernando-Herraez, I., Prüfer, K., Pybus, M., Johnstone, L., Lachmann, M., Alkan, C., Twigg, D.,  
759 Petit, N., Baker, C., Hormozdiari, F., Fernandez-Callejo, M., Dabad, M., Wilson, M. L., Stevison,  
760 L., Camprubí, C., Carvalho, T., Ruiz-Herrera, A., Vives, L., Mele, M., Abello, T., Kondova, I.,  
761 Bontrop, R. E., Pusey, A., Lankester, F., Kiyang, J. A., Bergl, R. A., Lonsdorf, E., Myers, S.,  
762 Ventura, M., Gagneux, P., Comas, D., Siegemund, H., Blanc, J., Agueda-Calpena, L., Gut, M.,  
763 Fulton, L., Tishkoff, S. A., Mullikin, J. C., Wilson, R. K., Gut, I. G., Gonder, M. K., Ryder,  
764 O. A., Hahn, B. H., Navarro, A., Akey, J. M., Bertranpetit, J., Reich, D., Mailund, T., Schierup,  
765 M. H., Hvilsom, C., Andrés, A. M., Wall, J. D., Bustamante, C. D., Hammer, M. F., Eichler,  
766 E. E., and Marques-Bonet, T. (2013). Great ape genetic diversity and population history. *Nature*,  
767 499(7459):471–475.
- 768 Prüfer, K., de Filippo, C., Grote, S., Mafessoni, F., Korlević, P., Hajdinjak, M., Vernot, B., Skov,  
769 L., Hsieh, P., Peyrégne, S., Reher, D., Hopfe, C., Nagel, S., Maricic, T., Fu, Q., Theunert, C.,  
770 Rogers, R., Skoglund, P., Chintalapati, M., Dannemann, M., Nelson, B. J., Key, F. M., Rudan,  
771 P., Kučan, Ž., Gušić, I., Golovanova, L. V., Doronichev, V. B., Patterson, N., Reich, D., Eichler,  
772 E. E., Slatkin, M., Schierup, M. H., Andrés, A. M., Kelso, J., Meyer, M., and Pääbo, S. (2017).  
773 A high-coverage Neandertal genome from Vindija Cave in Croatia. *Science*, 358(6363):655–658.
- 774 Racimo, F., Sankararaman, S., Nielsen, R., and Huerta-Sánchez, E. (2015). Evidence for archaic  
775 adaptive introgression in humans. *Nature Reviews Genetics*, 16(6):359–371.
- 776 Ragsdale, A. P. and Gravel, S. (2019). Models of archaic admixture and recent history from two-locus  
777 statistics. *PLOS Genetics*, 15(6):e1008204.
- 778 Ragsdale, A. P., Weaver, T. D., Atkinson, E. G., Hoal, E., Möller, M., Henn, B. M., and Gravel, S.  
779 (2022). A weakly structured stem for human origins in Africa. *bioRxiv*, pages 2022–03.
- 780 Ramachandran, S., Deshpande, O., Roseman, C. C., Rosenberg, N. A., Feldman, M. W., and Cavalli-  
781 Sforza, L. L. (2005). Support from the relationship of genetic and geographic distance in human  
782 populations for a serial founder effect originating in Africa. *Proceedings of the National Academy*  
783 *of Sciences*, 102(44):15942–15947.
- 784 Rodríguez, W., Mazet, O., Grusea, S., Arredondo, A., Corujo, J. M., Boitard, S., and Chikhi, L.  
785 (2018). The IICR and the non-stationary structured coalescent: Towards demographic inference  
786 with arbitrary changes in population structure. *Heredity*, 121(6):663–678.
- 787 Sankararaman, S. (2020). Methods for detecting introgressed archaic sequences. *Current Opinion*  
788 *in Genetics & Development*, 62:85–90.
- 789 Sankararaman, S., Mallick, S., Dannemann, M., Prüfer, K., Kelso, J., Pääbo, S., Patterson, N., and  
790 Reich, D. (2014). The genomic landscape of Neanderthal ancestry in present-day humans. *Nature*,  
791 507(7492):354–357.
- 792 Sankararaman, S., Patterson, N., Li, H., Pääbo, S., and Reich, D. (2012). The Date of Interbreeding  
793 between Neandertals and Modern Humans. *PLOS Genetics*, 8(10):e1002947.
- 794 Scerri, E. M. L., Chikhi, L., and Thomas, M. G. (2019). Beyond multiregional and simple out-of-  
795 Africa models of human evolution. *Nature Ecology & Evolution*, 3(10):1370–1372.
- 796 Scerri, E. M. L., Thomas, M. G., Manica, A., Gunz, P., Stock, J. T., Stringer, C., Grove, M.,  
797 Groucutt, H. S., Timmermann, A., Rightmire, G. P., d'Errico, F., Tryon, C. A., Drake, N. A.,  
798 Brooks, A. S., Dennell, R. W., Durbin, R., Henn, B. M., Lee-Thorp, J., deMenocal, P., Petraglia,  
799 M. D., Thompson, J. C., Scally, A., and Chikhi, L. (2018). Did Our Species Evolve in Subdivided  
800 Populations across Africa, and Why Does It Matter? *Trends in Ecology & Evolution*, 33(8):582–  
801 594.

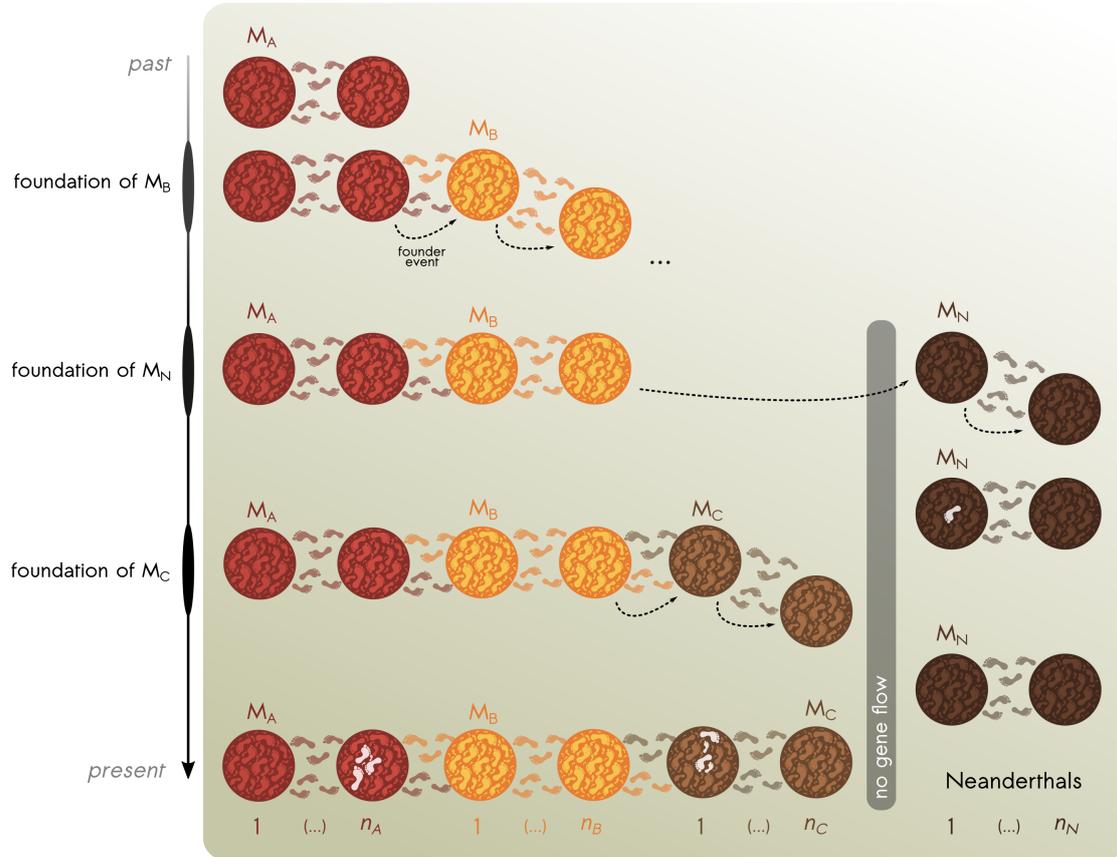
REFERENCES

REFERENCES

- 802 Schaefer, N. K., Shapiro, B., and Green, R. E. (2021). An ancestral recombination graph of human,  
803 Neanderthal, and Denisovan genomes. Science Advances, 7(29):eabc0776.
- 804 Skoglund, P., Malmström, H., Raghavan, M., Storå, J., Hall, P., Willerslev, E., Gilbert, M. T. P.,  
805 Götherström, A., and Jakobsson, M. (2012). Origins and Genetic Legacy of Neolithic Farmers  
806 and Hunter-Gatherers in Europe. Science, 336(6080):466–469.
- 807 Skov, L., Coll Macià, M., Sveinbjörnsson, G., Mafessoni, F., Lucotte, E. A., Einarisdóttir, M. S.,  
808 Jonsson, H., Halldorsson, B., Gudbjartsson, D. F., Helgason, A., Schierup, M. H., and Stefansson,  
809 K. (2020). The nature of Neanderthal introgression revealed by 27,566 Icelandic genomes. Nature,  
810 582(7810):78–83.
- 811 Tajima, F. (1990). Relationship between migration and DNA polymorphism in a local population.  
812 Genetics, 126(1):231–234.
- 813 Takahata, N. (1991). Genealogy of neutral genes and spreading of selected mutations in a geograph-  
814 ically structured population. Genetics, 129(2):585–595.
- 815 Talamo, S., Kromer, B., Richards, M. P., and Wacker, L. (2023). Back to the future: The advantage  
816 of studying key events in human evolution using a new high resolution radiocarbon method. PLOS  
817 ONE, 18(2):e0280598.
- 818 Theunert, C. and Slatkin, M. (2017). Distinguishing Recent Admixture from Ancestral Population  
819 Structure. Genome Biology and Evolution, 9(3):427–437.
- 820 Vernot, B. and Akey, J. M. (2014). Resurrecting Surviving Neandertal Lineages from Modern Human  
821 Genomes. Science, 343(6174):1017–1021.
- 822 Wakeley, J. (1999). Nonequilibrium Migration in Human History. Genetics, 153(4):1863–1871.
- 823 Wilkinson-Herbots, H. M. (1998). Genealogy and subpopulation differentiation under various models  
824 of population structure. Journal of Mathematical Biology, 37(6):535–585.
- 825 Yang, M. A., Malaspina, A.-S., Durand, E. Y., and Slatkin, M. (2012). Ancient Structure in Africa  
826 Unlikely to Explain Neanderthal and Non-African Genetic Similarity. Molecular Biology and  
827 Evolution, 29(10):2987–2995.

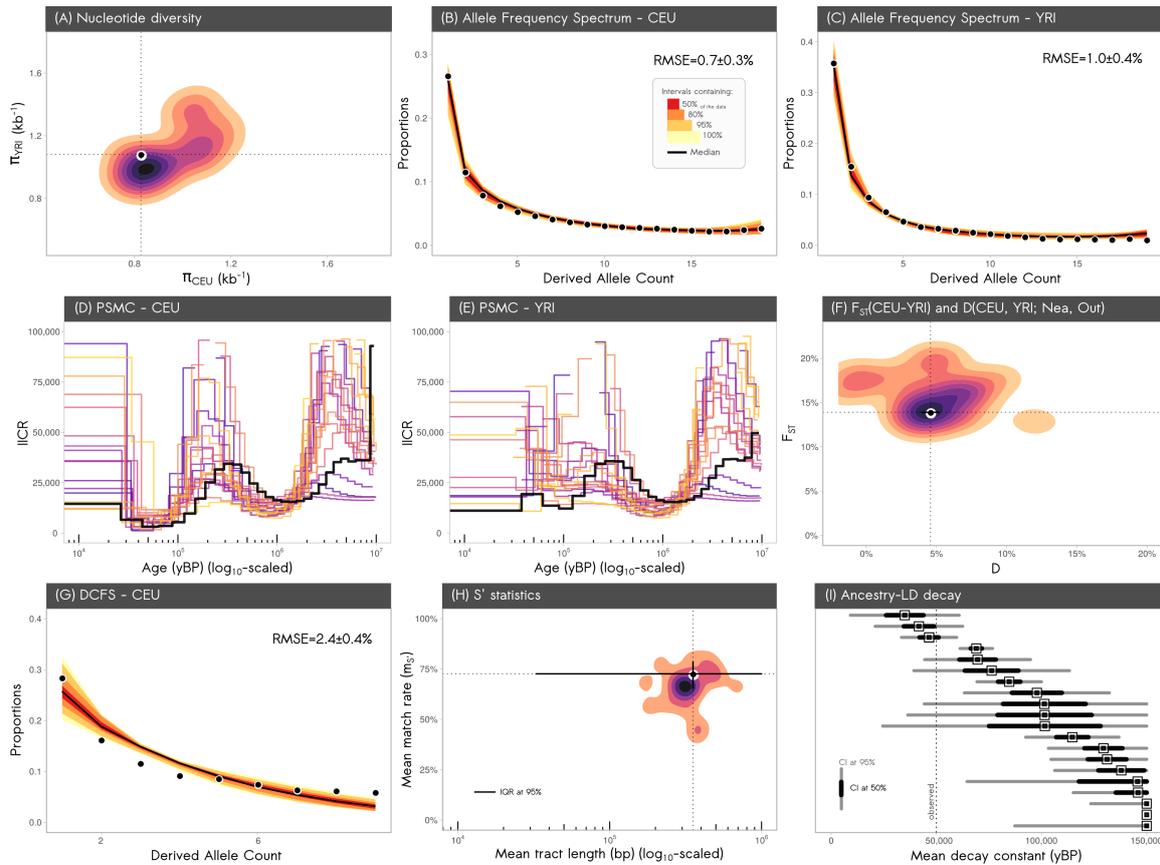
828 **Figures**

829 **Figure 1**



830 **Simplified representation of the 1D structured model considered in this study.** Time  
 831 flows from top (past) to bottom (present), with an initial metapopulation  $M_A$  consisting of  $n_A$  (=10)  
 832 demes exchanging migrants with their neighbours. At some point in the past, the rightmost deme  
 833 of  $M_A$  founds a new metapopulation  $M_B$  of  $n_B$  (=10) demes, with which it will continue exchanging  
 834 migrants till the present. Later, the rightmost deme of  $M_B$  founds the metapopulation  $M_N$  of  $n_N$   
 835 (=10) demes which will become Neanderthals. The  $M_N$  metapopulation will never exchange migrants  
 836 with any other deme from the other metapopulations. Closer towards the present, the rightmost  
 837 deme of  $M_B$  founds  $M_C$  which corresponds to the expansion of *H. sapiens* towards Eurasia. White  
 838 feet represent the sampled metapopulations (not the specific demes) for respective sampling times.  
 839 The location of the sampled demes (within the corresponding metapopulations) is a random variable  
 840 (see Notes S1.1). Fifty individuals are sampled in  $M_A$  and in  $M_C$  to represent modern-day YRI and  
 841 CEU samples respectively. For the Neanderthals ( $M_N$ ), one individual is sampled at 50 kya. More  
 842 details about the models can be found in the Materials and Methods section and in Notes S1.1.

843 **Figure 2**



844 **Performance of our admixture-free model compared to observed data (black) for the**  
 845 **twelve statistics used in this study.** The statistical distributions were obtained from the twenty  
 846 accepted runs of our model, after simulating twenty chromosomes of 30 Mbp (uniform recombina-  
 847 tion). **Abbreviations.** CI: confidence interval; IQR: interquartile range; IICR: inverse instan-  
 848 taneous coalescence rate. **(A)** Nucleotide diversity in CEU and YRI. We represent the 2D density  
 849 (Gaussian kernel) of the joint nucleotide diversities  $\pi$  estimated in simulated CEU and YRI samples.  
 850 The black dot represents the observed values from real data. **(B)** Normalized CEU *AFS*. The black  
 851 dots represent the observed *AFS* using ten real CEU individuals. The colors represent the 50%,  
 852 80%, 95% and 100% interquartile ranges of the normalized derived allele frequency spectrum calcu-  
 853 lated for ten simulated CEU samples across the twenty selected runs. The black line is the median  
 854 across the simulated *AFS*s. RMSE is the root mean squared error and is given in percentage. **(C)**  
 855 Normalized YRI *AFS*. Same as (B), for ten YRI samples. **(D)** CEU *PSMC* curves. The black curve  
 856 is the observed CEU *PSMC* whereas the coloured *PSMC* curves were obtained from the twenty  
 857 selected runs. The  $x$ -axis ( $\log_{10}$ -scaled) is expressed in years before present (with  $\mu=1.2 \times 10^{-8}$  and  
 858  $g=25$  years). **(E)** YRI *PSMC* curves. Same as (D), for a single YRI sample per selected run. **(F)**  
 859 Joint distribution of the  $F_{ST}$  and  $D$ -statistic. The  $D$ -statistic ( $x$ -axis) was calculated as  $D(50$  CEU,  
 860 50 YRI; 1 Neanderthal, Ancestral reference) and the  $F_{ST}$  ( $y$ -axis) was computed between the CEU  
 861 and YRI samples. Both statistics were computed for each of the twenty runs and plotted jointly  
 862 using a 2D density (Gaussian kernel). **(G)** CEU *DCFS*. As in panel (B), an interquartile ribbon plot  
 863 was obtained for the doubly conditioned site frequency spectra across the twenty selected runs. The  
 864 *DCFS* was computed for five CEU individuals. The black line is the median *DCFS* across the twenty  
 865 runs. The black dots represent the observed *DCFS*. **(H)** Joint distribution of mean Neanderthal  
 866 fragments lengths and match rates. A 2D density (Gaussian kernel) of the distribution for the mean

REFERENCES

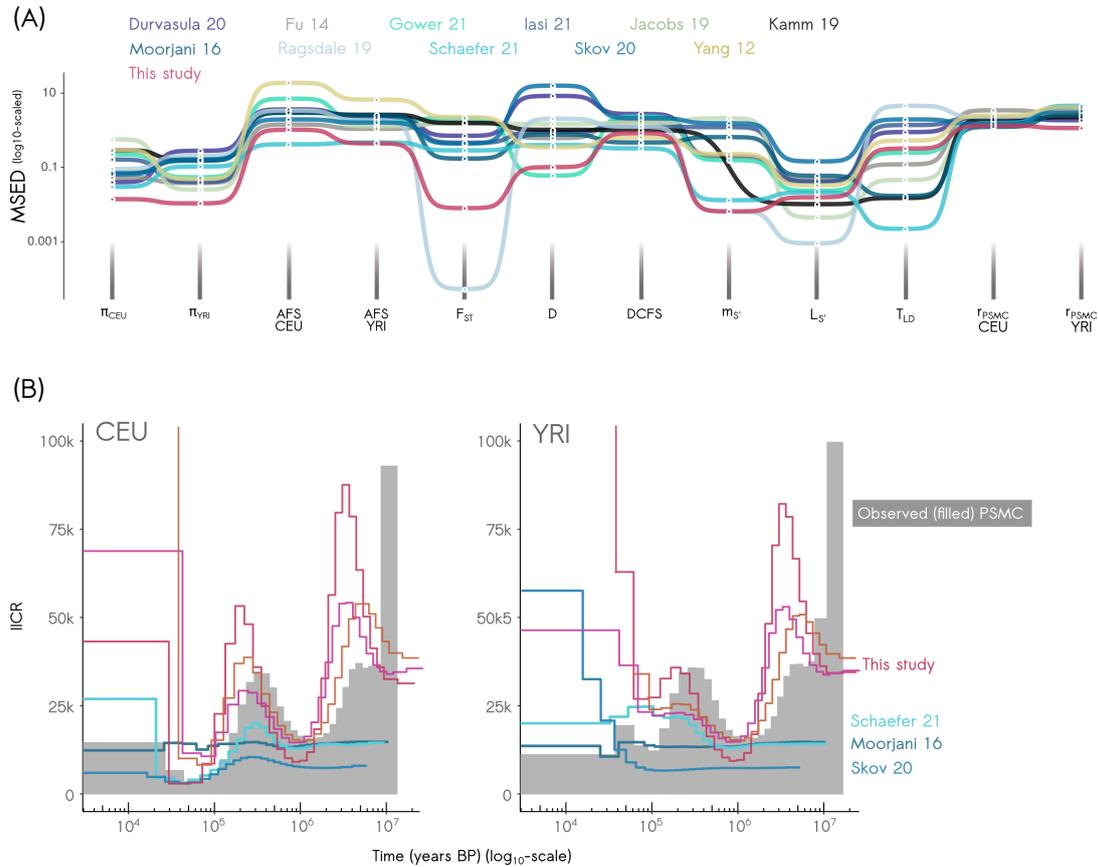
REFERENCES

867 length ( $x$ -axis, in bp,  $\log_{10}$ -scaled) and the mean match rate  $m_S$  ( $y$ -axis) is represented for the  
868 putative Neanderthal segments estimated with  $S'$  in the 50 simulated CEU samples. The black dot  
869 represents the estimates from the observed data (with associated 95% IQR). (I) Ancestry- $LD$  decay  
870 constant. The  $x$ -axis represents the exponential decay constant of the ancestry- $LD$  curve, and the  
871 twenty selected runs are represented on the  $y$ -axis. The point estimates (square dots) are classically  
872 interpreted as the age of the admixture events. The whisks represent the 50% (black thickest seg-  
873 ment) and the 95% confidence intervals (gray segment) around the jackknife mean (calculated using  
874 a weighted jackknife procedure). The point estimate of the decay constant computed from real data  
875 is  $\sim 50$  kya (vertical dashed line).

REFERENCES

REFERENCES

876 **Figure 3**

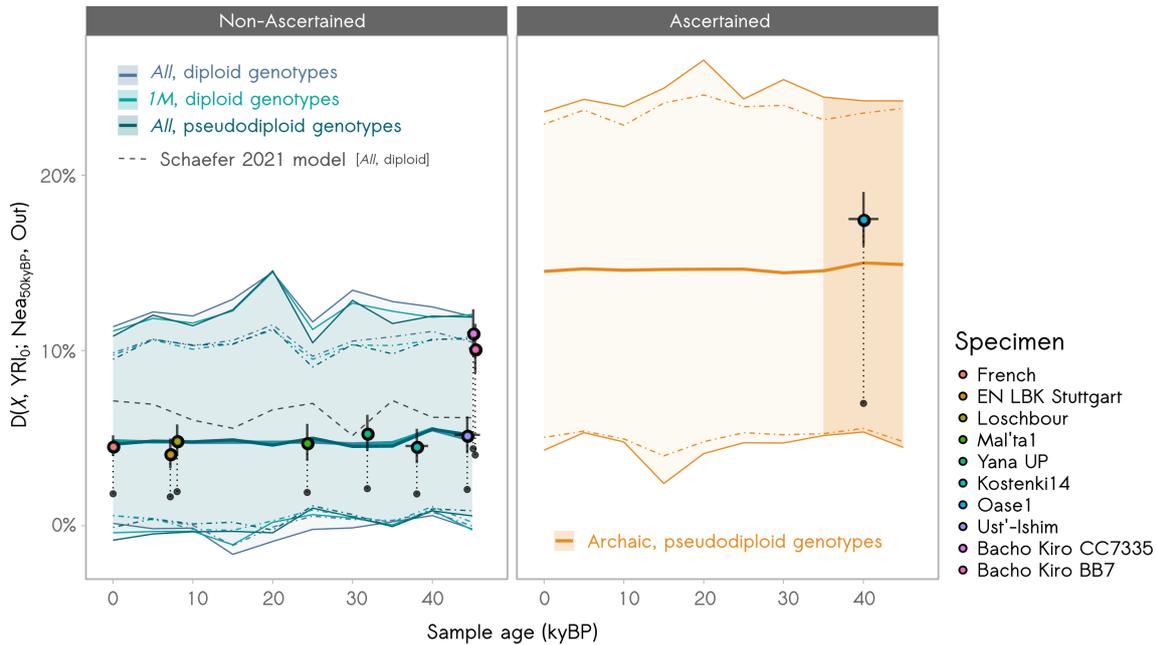


877 **Model comparison.** This figure compares the predictions of the eleven models for the twelve  
 878 statistics (Figure 2) with those of our 1D structured model (red). All models except ours assume  
 879 Neanderthal admixture into *Hs*. (A) Ranking of models for the twelve statistics. The statistics are  
 880 represented in the *x*-axis, and each model is represented by a different color. The *y*-axis represents  
 881 the minimum scaled Euclidean distance (MSED) to the observed statistics for twenty accepted runs  
 882 from each model. The *y*-axis is log<sub>10</sub>-scaled. (B) Representation of the CEU (left) and YRI (right)  
 883 *PSMC* curves from three among the twenty selected runs of our structured model (red shades) and  
 884 from three published models. Two of these published models were chosen because they rank right  
 885 after our model in the global model comparison (Schaefer et al., 2021; Moorjani et al., 2016). The  
 886 *PSMC* curves for all published models are represented in Notes S8. The grey (filled) *PSMC* curves  
 887 are the observed ones. Time (*x*-axis, log<sub>10</sub>-scaled) is given in years BP.

REFERENCES

REFERENCES

888 **Figure 4**



889 **aDNA  $D$ -statistics.** The figure represents the temporal trajectory of the  $D$ -statistic estimated on  
 890 a time series of ancient Eurasian ( $M_C$  metapopulation) samples—simulated under our structured  
 891 model together with empirical estimates (dots). The left panel represents the values computed from  
 892 non-ascertained SNPs. The right panel represents the values computed using a SNP ascertainment  
 893 mimicking the empirical data for Oase1 (Fu et al., 2015). To obtain the different values and curves,  
 894 we simulated single Eurasian samples ( $X$ ) at increasing ages of the past (from present-day up to  
 895 45 kya) under the twenty accepted runs of our structured model. We then estimated  $D$  as  $D(X, 50$   
 896  $YRI; 1$  Neanderthal, Ancestral reference). The shaded areas (contoured with solid lines) correspond  
 897 to the interquartile ranges at 100% of the  $D$  values estimated for various (i) SNP sampling strategies  
 898 (“ $All$ ”: all simulated SNPs; “ $1M$ ”: random subset of one million SNPs; “ $Archaic$ ”: ascertained set of  
 899 SNPs) and (ii) genotype quality (either diploid or pseudodiploid). The thick lines correspond to the  
 900 median values across these ranges. The colored dot-dashed lines represent the interquartile ranges  
 901 at 95%. For comparison, the dotted line (left panel) is the trajectory of the  $D$ -statistics estimated  
 902 under the Schaefer et al. (2021) model. We also represented the (i) *raw* (black points) values of  
 903  $D$  that we re-estimated for ten empirical ancient  $Hs$  samples from the AADR v54.1 dataset—SNP  
 904 array data—, and (ii) their *scaled* (colored points) estimates, obtained by multiplying all the raw  
 905  $D$  values by a constant factor so that the AADR-based raw estimate for the French present-day  
 906 samples fits the published CEU genome-wide estimate (4.57%). The empirical sample ages are the  
 907 calibrated  $^{14}C$  radiocarbon dates (AADR v54 annotation file). The error bars correspond to the  
 908 95% confidence intervals (in the case of the  $D$ -statistics, this is calculated using a weighted-block  
 909 jackknife procedure).