Modeling of African population history using *f*-statistics can be highly biased and is not addressed by previously suggested SNP ascertainment schemes

Pavel Flegontov [1,2,3,@*], Ulaş Işıldak [2,*,†], Robert Maier [1,*], Eren Yüncü [2,†], Piya Changmai [2], David Reich [1,4,5,6,@]

[1] *Department of Human Evolutionary Biology, Harvard University, Cambridge, MA, USA*
[2] *Department of Biology and Ecology, Faculty of Science, University of Ostrava, Ostrava, Czechia*
[3] *Kalmyk Research Center of the Russian Academy of Sciences, Elista, Russia*
[4] *Department of Genetics, Harvard Medical School, Boston, MA 02115, USA*
[5] *Howard Hughes Medical Institute, Harvard Medical School, Boston, MA, USA*
[6] *Broad Institute of Harvard and MIT, Cambridge, MA, USA*

† present addresses: E.Y., Department of Biological Sciences, Middle East Technical U., Ankara, Turkey; U.I., Leibniz Institute on Aging - Fritz Lipmann Institute (FLI), Jena, Germany.
** authors contributed equally*
@ corresponding authors: Pavel Flegontov (pavel.flegontov@hms.harvard.edu, pavel.flegontov@osu.cz) and David Reich (reich@genetics.med.harvard.edu)

Abstract

f-statistics have emerged as a first line of analysis for making inferences about demographic history from genome-wide data. These statistics can provide strong evidence for either admixture or cladality, which can be robust to substantial rates of errors or missing data. *f*-statistics are guaranteed to be unbiased under "SNP ascertainment" (analyzing non-randomly chosen subsets of single nucleotide polymorphisms) only if it relies on a population that is an outgroup for all groups analyzed. However, ascertainment on a true outgroup that is not co-analyzed with other populations is often impractical and uncommon in the literature. In this study focused on practical rather than theoretical aspects of SNP ascertainment, we show that many non-outgroup ascertainment schemes lead to false rejection of true demographic histories, as well as to failure to reject incorrect models. But the bias introduced by common ascertainments such as the 1240K panel is mostly limited to situations when more than one sub-Saharan African and/or archaic human groups

(Neanderthals and Denisovans) or non-human outgroups are co-modelled, for example, f_4-statistics involving one non-African group, two African groups, and one archaic group. Analyzing panels of SNPs polymorphic in archaic humans, which has been suggested as a solution for the ascertainment problem, cannot fix all these problems since for some classes of *f*-statistics it is not a clean outgroup ascertainment, and in other cases it demonstrates relatively low power to reject incorrect demographic models since it provides a relatively small number of variants common in anatomically modern humans. And due to the paucity of high-coverage archaic genomes, archaic individuals used for ascertainment often act as sole representatives of the respective groups in an analysis, and we show that this approach is highly problematic. By carrying out large numbers of simulations of diverse demographic histories, we find that bias in inferences based on *f*-statistics introduced by non-outgroup ascertainment can be minimized if the derived allele frequency spectrum in the population used for ascertainment approaches the spectrum that existed at the root of all groups being co-analyzed. Ascertaining on sites with variants common in a diverse group of African individuals provides a good approximation to such a set of SNPs, addressing the great majority of biases and also retaining high statistical power for studying population history. Such a "pan-African" ascertainment, although not completely problem-free, allows unbiased exploration of demographic models for the widest set of archaic and modern human populations, as compared to the other ascertainment schemes we explored.

Introduction

Archaeogenetics has achieved remarkable progress in the last decade (Skoglund and Mathieson 2018, Stoneking et al. 2023), with genome-wide data for thousands of ancient humans now being published each year. No region of the world is now inaccessible to archaeogenetic research, although isolation of enough authentic DNA from skeletons excavated in tropical and sub-tropical areas (Lipson et al. 2018) or from Pleistocene individuals (Hajdinjak et al. 2021) remains a challenge. For generating usable archaeogenetic data from Africa, targeted enrichment of human DNA on dedicated single nucleotide polymorphism (SNP) capture panels is almost always necessary. A majority of ancient DNA studies on African populations (Skoglund et al. 2017, van de Loosdrecht et al. 2018, Prendergast et al. 2019, Lipson et al. 2020, Wang et al. 2020, Sirak et al. 2021, Lipson et al.

2022) relied on a SNP capture panel usually termed "1240K" (Fu et al. 2015, Mathieson et al. 2015), and some studies on Upper Paleolithic humans relied on a supplementary panel ("1000K", comprising transversion polymorphisms found in two Yoruba individuals and transversion polymorphisms in the Altai Neanderthal genome) or on its union with 1240K (Fu et al. 2015, Hajdinjak et al. 2021). The 1240K panel was constructed of the following elements: all SNPs on the Human Origins array (itself composed of 13 sub-panels, each ascertained as heterozygous in a single high-coverage human genome, Patterson et al. 2012), all SNPs on the Illumina 650Y array, all SNPs on the Affymetrix 50k XBA array, and smaller numbers of SNPs chosen for other purposes (Fu et al. 2015). The 1240K capture panel is now used routinely for analyzing thousands of ancient humans across the world (Skoglund and Mathieson 2018, Olalde and Posth 2020), and successor panels including the full set of 1240K sites are now available (Rohland et al. 2022).

Bergström *et al*. (2020), relying on high-quality genomic data for present-day humans, showed that f_4-statistics including three sub-Saharan African groups and one non-African group, or four sub-Saharan African (hereafter “African”) groups can be biased when computed on common SNP panels such as Illumina MEGA, the panel used by Li *et al*. (2008), and the Affymetrix Human Origins array (Patterson et al. 2012). An influence of ascertainment on common population genetic analyses (*ADMIXTURE*, F_{ST}) was also demonstrated. However, the bias in f_4-statistics including archaic humans and apes was not explored.

Bergström *et al*. (2020) found that selecting approximately 1.3M SNPs polymorphic in the group composed of high-coverage archaic human genomes (the Altai and Vindija Neanderthals, the “Denisova 3” Denisovan) effectively eliminated the biases affecting f_4-statistics calculated on anatomically modern humans (AMH) and including 3 or 4 sub-Saharan African groups. A similar approach (selecting ca. 814K transversion sites variable between the Altai Neanderthal and Denisovan) was proposed by Skoglund *et al*. (2017). A SNP capture reagent relying on this principle, the *myBaits Expert Human Affinities Kit* “Ancestral 850K” module, became available in 2021 from Daicel Arbor Biosciences (https://arborbiosci.com/genomics/targeted-sequencing/mybaits/mybaits-expert/mybaits-expert-human-affinities/). This module targets approximately 850K biallelic transversion SNPs (autosomal and X-chromosomal) ascertained as polymorphic in the group composed of high-coverage archaic human genomes: the Altai (Prüfer et al. 2014), Vindija (Prüfer et al.

2017), and Chagyrskaya Neanderthals (Mafessoni et al. 2020), as well as the "Denisova 3" Denisovan genome (Meyer et al. 2012). This set of variable sites was shown to yield nearly unbiased F_{ST} values for pairs composed of an African and a non-African group within the Simons Genome Diversity Panel (SGDP) dataset (Mallick et al. 2016), in contrast to the 1240K panel (see a technical note on manufacturer's website: https://arborbiosci.com/wp-content/uploads/2021/03/Skoglund_Ancestral_850K_Panel_Design.pdf).

These recommendations are motivated by a theoretical property of *f*-statistics: if a SNP is the result of a single historical mutation and there has not been natural selection, the statistics are expected to be unbiased if SNPs are either unascertained or ascertained as polymorphic in a population that is an outgroup for all populations being analyzed (Patterson et al. 2012, Wang and Nielsen 2012), and the results in Bergström *et al*. (2020) and in the technical note published on the Daicel Arbor Biosciences product page are consistent with this theoretical property of outgroup ascertainment. The problematic case is non-outgroup ascertainment, that is ascertainment on a population that is co-analyzed with others. A series of papers explored non-outgroup ascertainment affecting measures of population divergence on simulated data and real data for humans and domestic animals (Nielsen and Signorovitch 2003, Nielsen 2004, Nielsen et al. 2004, Clark et al. 2005, Guillot and Foll 2009, Albrechtsen et al. 2010, Wang and Nielsen 2012, Lachance and Tishkoff 2013, McTavish and Hillis 2015, Malomane et al. 2018, Geibel et al. 2021). However, *D*- and *f*-statistics which have more robustness than other allele frequency-based statistics in many cases (Patterson et al. 2012), were not considered in those studies. Limited exploration of non-outgroup ascertainment schemes was performed on simulated data in publications introducing the *D*- and *f*-statistics, with the conclusion that biases are not noticeable in practice (Durand et al. 2011, Patterson et al. 2012).

The existing recommendations for a bias-free SNP enrichment panel also rely on the assumption that archaic humans are nearly perfect outgroups with respect to all AMH, and the low-level archaic admixture in non-Africans (Green et al. 2010, Reich et al. 2010) does not contribute major bias. However, evidence is accumulating that supports archaic admixture in Africans (Chen et al. 2020, Hubisz et al. 2020), and, according to some models, "super-archaic" ancestry (i.e., symmetrically related to Neanderthals, Denisovans, and AMH) may reach 19% in the common ancestor of AMH (Durvasula and Sankararaman 2020). Moreover, for outgroup ascertainment to be unbiased from the theoretical perspective, the

outgroup (or a closely related population) should not be then co-analyzed with other populations (Patterson et al. 2012, Wang and Nielsen 2012), and the individuals used for ascertainment should not be used as sole representatives of the respective groups. However, given the paucity of high-coverage archaic genomes (Meyer et al. 2012, Prüfer et al. 2014, 2017, Mafessoni et al. 2020) and the usefulness of archaic or African outgroups for calculation of f_4-and *D*-statistics and for testing of admixture models (Maier et al. 2022 preprint), these recommendations are often ignored in published *f*-statistic, *qpAdm*, *qpGraph*, and *TreeMix* analyses (e.g., Skolgund et al. 2017, Lipson et al. 2020, 2022, Hajdinjak et al. 2021, Kılınç et al. 2021, Yaka et al. 2021). For instance, archaic individuals are co-analyzed with anatomically modern humans on archaic-ascertained SNPs (Skoglund et al. 2017, Hajdinjak et al. 2021) or a Yoruba group is co-analyzed with non-Africans on Yoruba-ascertained SNPs (Kılınç et al. 2021, Yaka et al. 2021).

Since outgroup ascertainment that is "clean" from the theoretical point of view is rarely used in practice, and since the statistical power of outgroup ascertainment to reject incorrect models of population history was not investigated, it is reasonable to examine the performance of archaic ascertainment and common SNP panels such as 1240K in situations that are often encountered in practice. A technical development important for the work reported here is the *ADMIXTOOLS 2* package (Maier et al. 2022 preprint), which extends the functionality of the original *ADMIXTOOLS* package (Patterson et al. 2012), enabling bootstrap resampling for most tools and a rapid algorithm for finding optima in complex admixture graph topology spaces. The *ADMIXTOOLS 2* package also makes calculating millions of f_4-statistics and fitting tens of thousands of admixture graphs to data a routine task. These developments, taken together, allow us to explore biased *f*-statistics more systematically and provide more informed guidelines for future studies.

Results

1. Empirical analyses: exploration of the effect of ascertainment bias on real data

We assembled a set of diploid autosomal genotype calls for 352 individuals (Suppl. Table 1) sequenced at high coverage (Mallick et al. 2016; Fan et al. 2019), including mostly present-day individuals from the Simons Genome Diversity Project (SGDP), several high-coverage ancient genomes with diploid genotype calls (Lazaridis et al. 2014, Fu et al. 2014), and three

159 archaic human genomes: the “Denisova 3” Denisovan (Meyer et al. 2012), Vindija (Prüfer et
160 al. 2017) and Altai Neanderthals (Prüfer et al. 2014). Relying on this "SGDP+archaic" dataset,
161 we explored a wide array of ascertainment schemes: 1) A/T and G/C SNPs (henceforth
162 “AT/GC”) that are, unlike the other mutation classes, unaffected by biased gene conversion
163 (Pouyet et al. 2018), and are also unaffected by deamination ancient DNA damage; 2)
164 random thinning of the unascertained or "AT/GC" sets down to the size approximately equal
165 to that of the 1240K SNP panel if missing data are not allowed on a given population set; 3)
166 the 1240K panel (Fu et al. 2015); 4) the 1000K panel composed of 997,780 SNPs comprising
167 all transversion polymorphisms found in two African (Yoruba) individuals sequenced to high
168 coverage and transversion polymorphisms found in the Altai Neanderthal genome (Fu et al.
169 2015); 5) the union of the 1000K and 1240K panels termed 2200K (Hajdinjak et al. 2021); 6)
170 various components of the 1240K panel (the sites included in the Illumina 650Y and/or
171 Human Origins SNP arrays, sites included exclusively in one of them, and the remaining
172 sites); 7) the largest Human Origins sub-panels – panel 4 ascertained as sites heterozygous
173 in a single San individual, panel 5 ascertained as sites heterozygous in a single Yoruba
174 individual, their union (panels 4+5), and panel 13 including sites where a randomly chosen
175 San allele is derived relative to the Denisovan (Patterson et al. 2012); 8) all sites
176 polymorphic in a group uniting three high-coverage archaic genomes: the ”Denisova 3”
177 Denisovan, the Altai and Vindija Neanderthals (this ascertainment scheme is similar to those
178 proposed by Bergström *et al*. 2020 and in the technical note published on the Daicel Arbor
179 Biosciences product page); 9) transversion sites variable in the group comprising these three
180 high-coverage archaic genomes; 10) restricting to SNPs that have high minor allele
181 frequency (MAF > 5%) in the whole "SGDP+archaic" dataset, i.e. high global MAF; 11)
182 restricting to SNPs having high global MAF combined with taking A/T and G/C SNPs only; 12)
183 restricting to SNPs that have > 5% MAF in a selected African or non-African continental
184 meta-population, irrespective of their frequency in the other meta-populations (there are
185 nine such meta-populations in our dataset, and thus nine different ascertainments, see
186 Suppl. Table 1); 13) restricting to SNPs that have > 5% MAF in a selected continental meta-
187 population, A/T and G/C SNPs only. For a list of SGDP-derived SNP sets explored in this study
188 and their sizes in terms of groups, individuals, and SNPs see Suppl. Table 2.
189 To investigate the influence of ascertainment on the ranking of admixture graph models
190 according to their fits to data, we analyzed real data, considering sets of five populations

and generated all possible admixture graph topologies with two admixture events (32,745 distinct topologies with no fixed outgroup; we considered graphs of this complexity as it was unfeasible to work with exhaustive collections of more complex graphs). First, we tested three combinations of groups (Suppl. Fig. 1). Admixture graph residuals on all sites, on a random subset of them approximately equal to the size of the 1240K set, and on AT/GC sites, are tightly correlated (*R* for a linear model ~1). Residuals of graph models including non-Africans only are also highly correlated on all sites and 1240K sites (*R* = 0.95-0.99, Suppl. Fig. 1). In contrast, the worst f_4-statistic residuals (WR) for graphs including one archaic human, three African groups, and one African group with ca. 60% of non-African ancestry (Fan et al. 2019) are poorly correlated on all sites and 1240K sites (*R* = 0.31-0.35). Thus, admixture graph fit rankings are severely affected by the 1240K ascertainment if certain population combinations are involved. We considered the possibility that this case of poor correlation was driven by admixture graph topologies that were obviously inconsistent with the data – that is, topologies could be shown to be inconsistent with the data based on gold standard SNP sets without ascertainment bias. However, the lack of strong correlation for some combinations of populations is not just driven by graphs with poor fits to the data. For example, WRs of admixture graphs that are well-fitting the data (WR <2.5 SE) on a random subsample of 840,000 sites have worst residuals ranging from nearly 0 SE to about 10 SE (Suppl. Fig. 1) on ca. 845,000 sites included in the 1240K panel. Rejecting a model that fits the data on unascertained data runs the risk of rejecting the true model, as we show on simulated data in the next section. The converse problem also applies: some admixture graphs are well-fitting (WR <2.5 SE) on the 1240K sites but fit a random sample of sites poorly (WR >5 SE, Suppl. Fig. 1).

Next, we explored the same exhaustive set of admixture graph topologies including five groups and two admixture events on the wider collection of ascertainments listed above and on a wider collection of populations. Twelve combinations of five groups including up to two archaic humans, up to five African groups, and up to five non-African groups were tested. In Suppl. Fig. 2 we compare various ways of looking on the effects of ascertainment, using a population quintuplet "Denisovan, Khomani San, Mbuti, Dinka, Mursi" as an example. In Table 1 we focus on the fraction of topologies that are rejected under ascertainment (WR > 3 SE) but accepted on all sites (WR < 3 SE) as a metric appropriate for quantifying the most serious effects of ascertainment bias, namely the probability of

rejecting the true model. In the supplementary materials, we also show alternative measures: a metric reflecting the statistical power of ascertainment, namely the fraction of topologies that are accepted under a given ascertainment (WR < 3 SE) but rejected on all sites (WR > 3 SE) (Suppl. Table 3), and R^2 of a linear trend for admixture graph WR (Suppl. Figs. 3 and 4, Suppl. Table 4) or log-likelihood (LL) scores (Suppl. Figs. 3 and 4, Suppl. Table 5).

Although we recognize that there can be no strict rule for classifying ascertainments into biased and unbiased ones since they form a continuum, for high-throughput analysis a classifier is useful. Moreover, fits of admixture graphs vary even in the absence of ascertainment bias, due to random site sampling effects (Fig. 1, Suppl. Fig. 2), as was shown in previous work (Maier et al. 2022 preprint). In this study, we considered a SNP set biased if a metric (such as the fraction of topologies rejected under ascertainment but accepted on all sites) was above (or below, as appropriate) the 2.5th percentile of this metric's distribution across 200 sets of randomly sampled SNPs equal to the size of the 1240K set for a given population combination.

Inspecting the key metric of ascertainment performance (the fraction of topologies that are rejected under ascertainment but accepted on all sites), we found only three site sampling schemes that were classified as unbiased for all the population quintuplets tested: the A/T and G/C mutation classes, Human Origins panel 4, and the union of Human Origins panels 4 and 5 (Table 1). However, due to the low number of sites in the latter two panels, the union of Human Origins panels 4 and 5, and especially panel 4, lack power to reject admixture graph models as compared to the 1240K panel and to the A/T and G/C mutation classes, as we show in Suppl. Table 3. Thus, the only ascertainment scheme that is problem-free according to both metrics is a random one: taking the A/T and G/C mutation classes.

Among the population quintuplets tested, "Altai Neanderthal, Ju|'hoan North, Biaka, Yoruba, Agaw" (Fig. 1, Suppl. Fig. 3a) and "Altai Neanderthal, Ju|'hoan North, Luhya, Palestinian, Spanish" (Suppl. Fig. 3b) are most susceptible to ascertainment bias (Table 1). A very similar quintuplet "Altai Neanderthal, ancient South African hunter-gatherers, Biaka, Yoruba, Agaw" is encountered within more complex admixture graph models that occupy a central place in Lipson *et al*. (2020, 2022) based on 1240K data (see an investigation of bias affecting the admixture graphs from these studies in Suppl. Text 1 and Suppl. Figs. 5 and 6). As explored below on real and simulated data, a class of f_4-statistics that are strongly

affected by non-random ascertainment underlies admixture graphs for both problematic population quintuplets: f_4(African X, archaic; African Y, non-African). On the other hand, population sets including no archaic human were virtually unbiased (Table 1), but some ascertainment schemes showed limited power to reject admixture graph models in these cases (Suppl. Table 3).

Archaic ascertainment has been suggested in the literature (Skoglund et al. 2017, Bergström et al. 2020) as a way to reduce ascertainment bias, however this approach is guaranteed to work only if the outgroup or a related group is not included itself in admixture graphs or *f*-statistics, and if individuals used for ascertainment are not sole representatives of the respective groups in an analysis. Indeed, our practical-oriented analysis showed that archaic ascertainment is biased in the case of the most problematic population quintuplet "Altai Neanderthal, Ju|'hoan North, Biaka, Yoruba, Agaw" (Table 1); in fact, the archaic ascertainment approach is by far the most biased scheme for population sets including both Neanderthal and Denisovan individuals (Table 1, see also results on simulated data below), and in our analysis it also emerged as the scheme with the lowest statistical power to reject admixture graph models (Suppl. Table 3).

If we combine both key bias metrics (the fraction of topologies rejected under ascertainment but accepted on all sites, and the fraction of topologies accepted under ascertainment but rejected on all sites), the 1240K and archaic ascertainments are out-performed by many ascertainment schemes, and most notably by the following: 1) the union of Human Origins panels 4 and 5; 2) the 2200K panel, which combines various kinds of ascertainment such as the 1240K panel, ascertainment on two Yoruba individuals, and on the Altai Neanderthal (Fu et al. 2015); and 3) restricting to variants that are common in the African meta-population in SGDP (Suppl. Table 1), optionally followed by removal of all mutation classes except for A/T and G/C (Table 1). R^2 of a linear trend for admixture graph WR is a metric that in some cases is informative in a way that the fractions of rejected/accepted topologies are not. As illustrated in Fig. 1, R^2 may differ substantially across ascertainment schemes while the fractions of topologies rejected under ascertainment but accepted on all sites or *vice versa* stay nearly constant across most ascertainment schemes (Table 1, Suppl. Table 3). Considering R^2 for admixture graph WR, restricting to variants that are common in a diverse set of African groups (we term this ascertainment scheme "pan-African" or "African MAF" for convenience) emerges as the

287 least biased form of ascertainment (Suppl. Table 4). We note that conclusions of this sort
288 are not quantitative since our collection of 12 population quintuplets, although diverse, is
289 just a small sample from the vast set of all possible population combinations. However,
290 exploring all possible combinations is infeasible, and we consider our approach to be useful
291 as a practical guide for assessing the performance of ascertainment schemes when
292 admixture graphs including archaic humans, Africans, and non-Africans are fitted to genetic
293 data.
294

295 2. Simulation studies confirm the qualitative patterns from exploration of empirical data

296 A major limitation of our empirical analyses of ascertainment bias is that fitting a model
297 with two admixture events is almost certainly inadequate for the histories relating various
298 sets of populations being analyzed. Thus, it is almost certain that all fitted models will be
299 wrong. When we fit wrong models, we have no guarantee that the (incorrect) admixture
300 graph fit to the data will give the same signal of deviation for different SNP ascertainments.
301 Different SNP ascertainments including random ascertainments will simply be sensitive to
302 different aspects of the deviations between the wrong model and the true history. Thus,
303 while the poor correlation between model fits on all sites and under different SNP
304 ascertainment schemes for combinations of archaic humans, sub-Saharan Africans, and
305 non-Africans is a potential signal of bias in analyses, it is valuable to analyze data where the
306 truth is known, as is the case for simulations, to provide clear evidence that typical
307 ascertainment schemes can cause false-positive inferences about history.
308 Using *msprime v.1.1.1* (Baumdicker et al. 2022), we simulated genetic data (a diploid
309 genome composed of three 100 Mb chromosomes with recombination) that reproduce the
310 F_{ST} values (Suppl. Fig. 7a) observed when comparing AMH groups, AMH and archaic
311 humans, and AMH and chimpanzee (Fischer et al. 2006). Ten independent simulations were
312 performed under the same parameters and under a topology (Fig. 2a) that in most
313 important aspects conforms to commonly discussed models of the relationships between
314 anatomically modern and archaic humans (Prüfer et al. 2014, Durvasula and Sankararaman
315 2020). We either simulated or omitted the Neanderthal gene flow to the ancestors of non-
316 Africans (via an unsampled proxy group).

We tested several non-outgroup ascertainment schemes: 1) ascertainment on heterozygous sites in a randomly selected individual from the "African 2" group (Fig. 2a, this ascertainment follows the scheme used for generating some of the 12 panels of sites comprising the Human Origins SNP array (Patterson et al. 2012)); 2) ascertainment on heterozygous sites in four randomly selected individuals, one per each "AMH" group (we consider the resulting SNP set to be qualitatively similar to the whole Human Origins SNP set); 3) archaic ascertainment (sites polymorphic in a group composed of one "Denisovan" individual and one individual per each "Neanderthal" group; the same individuals were subsequently used for calculating *f*-statistics); 4) "African MAF ascertainment", that is restricting to sites with MAF > 5% in the union of two "African" groups; 5) similar MAF-based ascertainment on two "non-African" groups or 6) on all four "AMH" groups.

First, we fitted the correct admixture graph as often practiced in the literature (e.g., Lipson et al. 2020, 2022): including the outgroup, one "archaic" individual, and all "AMH" groups. Human Origins-like ascertainment (one panel) always leads to rejection of the correct model, both in the absence and in the presence of the Neanderthal gene flow to non-Africans, with WR ranging from 3.4 to 8.8 SE (Fig. 2b). Another form of Human Origins-like ascertainment (four panels) is less problematic but led to rejection of the correct model (WR > 3 SE) in 9 of 30 cases (in the presence of the Neanderthal gene flow to non-Africans), with WR up to 4.6 SE. Only the archaic and African MAF non-outgroup ascertainments (in the presence of the Neanderthal gene flow to non-Africans) did not lead to rejection of these simplified graph topologies, known to be correct since we simulated them. However, when the full simulated model (with the Neanderthal gene flow to non-Africans) including the outgroup and three "archaic" lineages is fitted to the data, all non-outgroup ascertainment schemes become problematic, except for African MAF ascertainment (Fig. 2b).

We also investigated effects of ascertainment on model ranking using the same approach as that applied to real data. All possible graph topologies with two admixture events (32,745) were fitted to population quintuplets of the following composition: "d or n1 or n2", "a1", "a2", "na1", "na2". The fractions of topologies rejected/accepted under ascertainment but accepted/rejected on all sites (and the bias classifier) were then used to reveal simulation iterations and ascertained schemes that demonstrated biased model fits (Suppl. Table 6). When no Neanderthal/non-African gene flow was simulated, only the non-

African MAF ascertainment emerged as problematic (at least half of simulation iterations for at least one population quintuplet were classified as affected by bias) according to the fraction of topologies rejected under ascertainment but accepted on all sites (Suppl. Table 6). When the Neanderthal to non-African gene flow was simulated, all ascertainment schemes, except for the Human Origins-like ascertainment (four panels), emerged as problematic according to the same metric (Suppl. Table 6). Summarizing these results on model ranking and on fits of the true model, we note that the Human Origins-like ascertainment (four panels) is relatively problem-free, unlike archaic ascertainment, MAF-based ascertainments, and Human Origins-like ascertainment (one panel), but it still led to rejection of the true model more often than on all sites or on random site subsamples (Fig. 2b).

Next, we explored non-outgroup ascertainment schemes that are similar to those presented in Fig. 2b but are based on randomly chosen groups (see Methods for details) and were applied to SNP sets resulting from simulated genetic histories in the form of random admixture graphs. Graphs of four complexity classes including 9 or 10 populations and 4 or 5 admixture events were simulated using *msprime v.1.1.1*. Only simulations where pairwise F_{ST} for groups were in the range characteristic for anatomically modern and archaic humans were selected for further analysis, resulting in 20 random topologies per graph complexity class, each including an outgroup (see examples of the simulated histories and F_{ST} distributions in Suppl. Fig. 8). Fits of the true admixture graph (WR) including an outgroup were compared on all sites and on ascertained SNP sets for each topology and ascertainment iteration (Fig. 2c). We note that our simulation setup generated groups sampled at different dates in the past (from 0 to ca. 40,000 generations) or, in other words, groups that have experienced widely different levels of genetic drift with respect to the root (Fig. 2d).

As illustrated by distributions of true admixture graph WRs in Fig. 2c, 'blindly' ascertaining on individuals or sets of groups randomly sampled across the graph almost guarantees rejecting the true historical model by a wide margin. Ascertainment on sites polymorphic in randomly composed sets of three individuals (one individual per group) and restricting to variants common (MAF > 5%) in randomly composed sets of four populations are two forms of ascertainment that are especially problematic (Fig. 2c). Human Origins-like ascertainments (one or four panels) often yield acceptable fits of the simulated graph (WR <

3 SE), although median WR of the true graphs equals 4.6 SE for these ascertainment schemes across all graph topologies and all (non-root and non-outgroup) populations used for ascertainment (Fig. 2c).

An illuminating result is that F_{ST} between the population used for Human Origins-like non-outgroup ascertainment and the root influences WR of the true graph: all ascertainments with the corresponding $F_{ST} < 0.12$ produce relatively unbiased fits of true graphs (WR < 4 SE, see Fig. 2d). In other words, ascertainment on heterozygous sites in a single individual taken from a population that is not an outgroup and is co-analyzed with other populations, but is genetically close to the root of the simulation, is relatively unbiased, unlike ascertainment on a single individual from a more drifted "present-day" population. We directly illustrate this effect by comparing results on outgroups co-analyzed with other populations that are more or less drifted with respect to the root (with effective population sizes differing by two orders of magnitude) (Fig. 2c,d). Co-analyzing the individual used for ascertainment with other groups does not produce biased results if that individual is a part of a wider population of 10 individuals. However, if that individual is the only representative of its group for model fitting, WRs are inflated drastically (Fig. 2c). We also illustrate the difference between true outgroup ascertainment, when an outgroup is not co-modelled with the other groups, and ascertainment on an outgroup that is included in the fitted model (Fig. 2f), which the kind of ascertainment shown in Fig. 2c and elsewhere. The former form of ascertainment is expected to be unbiased even for a highly drifted outgroup (Fig. 2f), while the latter is not (Patterson et al. 2012, Wang and Nielsen 2012).

Genetic distance from the root is a noisy proxy for the similarity of derived allele frequency (DAF) spectra of the root population and the population used for ascertainment. In Fig. 2e we show DAF spectra across all simulated populations used for Human-Origins-like non-outgroup ascertainment, but only for variants that are polymorphic in the respective root population sample of 10 diploid individuals. These results demonstrate that for the correct admixture graph to fit ascertained data well (WR < 4 SE, Fig. 2e), non-outgroup ascertainment should be based on a population where the derived allele frequency spectrum of sites that were polymorphic at the root is preserved relatively well (see a more extensive version of this plot in Suppl. Fig. 9). We note that some ascertainments may be unbiased with respect to the true graph but may have low power due to the paucity of sites

with high MAF in "present-day" populations. Indeed, ascertainments on the root itself or on groups genetically close to the root (such as outgroups with a large effective size) are unbiased (Fig. 2c,d), but on average demonstrated lower power to reject incorrect admixture graph models as compared to single-panel Human Origins-like ascertainments on more drifted groups (Suppl. Fig. 10).

These results on randomized ascertainment schemes (not to be confused with random site sampling) and simulated histories in the form of random admixture graphs show that ascertainment on groups that are highly drifted with respect to the root of the groups being co-analyzed is problematic. Thus, if proper outgroup ascertainment is impractical (if an outgroup shares few polymorphisms with the other populations analyzed, or if an outgroup is needed for constraining the analysis), unascertained or randomly sampled sets of sites should be treated as a gold standard for admixture graph inference. The 1240K ascertainment is much more complex (Fu et al. 2015, Mathieson et al. 2015) than the ascertainment schemes explored on simulated data, but its effects are possibly intermediate between the effects of a MAF-based ascertainment (since all common SNP panels are more or less depleted for rare variants) and ascertainment on heterozygous sites in single individuals from several groups (since approximately half of the 1240K sites are derived from the Human Origins SNP array, Fu et al. 2015). Thus, we expect an accurate admixture graph including at least one archaic human, at least two African groups, and at least one non-African group (Fig. 2a) to fit the data poorly under the 1240K ascertainment.

Finally, we checked if non-outgroup ascertainment could bias the simplest cladality tests in the absence of gene flow. A tree of four groups conforming to the f_4-statistic (A, B; C, O) was simulated using *msprime v.1.1.1*, with a tree depth of 4,000 or 40,000 generations (Fig. 3a). All the groups had a uniform effective population size of 100,000 diploid individuals, except for a 10x to 10,000x size reduction immediately after the A-B divergence (1,999 or 9,995 generations in the past). While the dramatic drop in the effective population size of group A yields a complex shape of the derived allele frequency spectrum in {A, B} (Martin and Amos 2021), two of three ascertainment schemes explored here (Human Origins-like ascertainment on one group and MAF-based ascertainment, but not removal of the derived end of the allele frequency spectrum; see Methods) increase the noise in the f_4-statistic (A, B; C, O), but do not shift the statistic away from its expectation at 0 (Suppl. Fig. 11). These results confirm an observation by Patterson et al. (2012) that in the case of perfect trees

445 SNP non-outgroup ascertainment does not lead to false rejection of cladality. However, as
446 demonstrated in Fig. 2, non-outgroup ascertainment is generally problematic in the case of
447 complex demographic histories with multiple admixture events.
448
449 3. An overview of f_4-statistic biases caused by non-outgroup ascertainment
450 We explored various classes of f_4-statistics exhaustively to obtain a "bird's-eye view" on
451 ascertainment biases that was previously difficult to do due to technical challenges in
452 calculating millions of *f*-statistics (Bergström *et al*. 2020). Another motivation for this
453 analysis was the fact that it is unfeasible to explore fits of large collections of admixture
454 graphs on thousands of population sets, ascertainment schemes and random site
455 subsamples. However, if an exhaustively sampled class of *f*-statistics is demonstrated to be
456 unbiased, all admixture graph fits based on those statistics are expected to be unbiased too.
457 We used the standard deviation of residuals from a linear trend for f_4-statistic Z-scores
458 on all sites and under ascertainment (termed "residual SE" for brevity and expressed in the
459 same units as Z-scores) as a single statistic for summarizing results, and we note that it
460 reflects both bias introduced by ascertainment and variance generated by random site
461 sampling. In Suppl. Table 7 and Suppl. Figs. 12 and 13 we show residual SE values for a
462 collection of 27 exhaustively sampled f_4-statistic classes and for the large collection of
463 ascertainment schemes that was used for exploring effects of ascertainment on admixture
464 graph fits in section 1. The f_4-statistic classes explored can be described concisely as
465 African(all SGDP populations)$_x$;archaic$_y$;chimpanzee$_1$, African(unadmixed with West
466 Eurasians)$_x$;archaic$_y$;Mediterranean/Middle Eastern (abbreviated as Med/ME)$_z$,
467 African(unadmixed with West Eurasians)$_x$;East Asian$_y$(y>0), American$_x$;European$_y$;Papuan$_z$.
468 Here, *x*, *y*, and *z* stand for the number of groups in the population quadruplet; thus,
469 "African$_3$;East Asian$_1$" would mean three Africans and one East Asian. All possible distinct f_4-
470 statistics composed of those "ingredients" were considered.
471 The effect of ascertainment schemes varies dramatically across the classes of f_4-
472 statistics, but ascertainment schemes based on one or two African individuals (Human
473 Origins sub-panels 4, 5, 13, 4 and 5 combined), on the three archaic individuals (either all
474 sites or transversions only), and components of the 1240K panel such as Illumina650Y
475 emerged as the worst-performing when results across all the f_4-statistic classes were
476 considered (Suppl. Table 7, Suppl. Fig. 12). Ascertainment schemes based on a global MAF

threshold or on a MAF threshold in a single non-African continental meta-population, and the 1000K and 2200K panels are similar in their effects to the 1240K ascertainment (Suppl. Table 7, Suppl. Fig. 12). We recognize that there is a continuum between unbiased and biased ascertainment schemes, and that for nearly all schemes and f_4-statistic classes a majority of statistics remain unaffected by ascertainment, but for describing our results in a concise way and for partially factoring out effects of SNP panel size, we applied the criterion similar to that employed above for admixture graph fits: residual SE for an f_4-statistic class is higher than the 97.5th percentile across 200 randomly thinned datasets matching the 1240K panel in size. According to this criterion, the 1240K ascertainment is problematic in the case of the following nine f_4-statistic classes (Suppl. Table 7): 1) African$_4$, 2) African$_3$;Med/ME$_1$, 3) African$_3$;East Asian$_1$, 4) African$_3$;archaic$_1$, 5) African$_3$;chimpanzee$_1$, 6) African$_2$;archaic$_1$;Med/ME$_1$, 7) African$_2$;archaic$_1$;chimpanzee$_1$, 8) archaic$_3$;Med/ME$_1$, 9) archaic$_3$;African$_1$, and unproblematic for the remaining 18 classes exhaustively explored in this analysis. Unlike all the other classes explored here (Suppl. Table 7, Suppl. Fig. 12), the class of statistics African$_2$;archaic$_1$;Med/ME$_1$, specifically statistics of the form f_4(African X, archaic; African Y, non-African), is substantially biased under all non-random types of ascertainments (Fig. 4a). The classes African$_3$;X are problematic under all ascertainment schemes except for the pan-African ascertainment (Suppl. Table 7, see an example in Fig. 4b), and the class African$_4$ is problematic under all ascertainment schemes except for the 1000K, 2200K, and pan-African ascertainment (Suppl. Table 7). Scatterplots underlying these residual SE estimates are also shown in Fig. 5 (for some of the most problematic classes highlighted above) and in Suppl. Figs. 14-16 (for all classes). Importantly, the classes of statistics most affected by ascertainment (African$_2$;archaic$_1$;Med/ME$_1$, African$_2$;archaic$_1$;chimpanzee$_1$, African$_3$;X, and African$_4$) are often relevant for fitting admixture graph models of African population history (see Suppl. Text 1). However, for most classes that were classified as problematic, except for African$_2$;archaic$_1$;Med/ME$_1$, African$_2$;archaic$_1$;chimpanzee$_1$, and African$_3$;X, residuals remain below 1 SE for a great majority of f_4-statistics (Suppl. Table 7 and Suppl. Fig. 12), and thus these statistics are probably not problematic in practice.

Results very similar to those presented above were obtained with a different metric: R^2 of a linear model for f_4-statistics themselves (Suppl. Table 8), instead of residual SE of a linear model for f_4-statistic Z-scores (Suppl. Table 7). For additional details on f_4-statistic

classes see Suppl. Text 2 (and Suppl. Figs. 13-16), and for a dissection of effects of ascertainment on few selected f_4-statistics see Suppl. Text 3 (and Suppl. Tables 9-12).

In contrast, statistics including non-Africans only, or one or two African groups and non-Africans (see an example in Figs. 4c and 5), are unproblematic under the 1240K, 2200K, pan-African and other MAF-based ascertainments (but demonstrate increased variance due to paucity of sites with high MAF under some other ascertainment types such as Human Origins panels 4&5 and archaic ascertainment, Suppl. Table 7).

Pan-African ascertainment (restricting to variants common across 68 African individuals unadmixed with West Eurasians, or across 94 African individuals, Suppl. Table 1), emerged as the best-performing non-outgroup ascertainment scheme. Unlike the other ascertainment schemes explored in this study, this type of ascertainment demonstrates a bias only in the case of the (African X, archaic; African Y, non-African) class of f_4-statistics (when only statistics with $|Z| < 15$ SE on all sites were considered, Suppl. Table 7, Suppl. Fig. 12). Another class of f_4-statistics is biased under this ascertainment scheme when all statistics are considered: f_4(non-African X, archaic; African, non-African Y) (Suppl. Fig. 12), and pan-African ascertainment is unbiased in the case of the other 25 classes of f_4-statistics explored in this study (Suppl. Table 7, Suppl. Fig. 12), which also translates into downstream analyses such as fits of admixture graph models (Fig. 1, Table 1, Suppl. Fig. 3, Suppl. Tables 3-5).

Discussion

f-statistics (Patterson et al. 2012) form a foundation for a range of methods (*qpWave, qpAdm, qpGraph*) that are used widely for studying population genetic history of humans and other species (see, for instance, Bergström et al. 2020b, Librado et al. 2021). Here, we focused on f_4-statistics, which are used as standalone tests for cladality (Reich et al. 2009, Patterson et al. 2012) and underlie the *qpAdm* method for fitting admixture models (Haak et al. 2015, Harney et al. 2021). The other *f*-statistics (f_2 and f_3) can be defined as special cases of f_4-statistics [$f_2(A, B) = f_4(A, B; A, B)$ and $f_3(A; B, C) = f_4(A, B; A, C)$], and are subject to the same kinds of biases. The existence of bias in the case of non-outgroup ascertainment was recognized in a publication introducing a suite of methods relying on *f*-statistics (Patterson et al. 2012), but its effects on large collections of f_4-statistics or on fits of diverse admixture

graph models were not explored in that study and in subsequent studies. Since usage of archaic or African outgroups is often unavoidable for calculation of f_4-and *D*-statistics and for construction of admixture graph or *qpAdm* models (e.g., Skolgund et al. 2017, Lipson et al. 2020, 2022, Hajdinjak et al. 2021, Kılınç et al. 2021, Yaka et al. 2021), unbiased ascertainment on an outgroup that is not co-analyzed with other populations (as illustrated on simulated data in Fig. 2f) is uncommon in practice. And frequently used SNP panels such as 1240K were built using very complex forms of non-outgroup ascertainment. Therefore, in this study we focused on practical rather than theoretical aspects of the ascertainment bias problem and considered forms of non-outgroup ascertainment that are common in the literature on archaeogenetics of humans, including ascertainment on a phylogenetic outgroup co-analyzed with other populations.

The present analysis showed that f_4-statistics of specific types are affected by ascertainment bias. The most striking example we found is a class of statistics f_4(African X, archaic; African Y, non-African). All statistics in this class are strongly biased in the same direction under the 1240K ascertainment (Suppl. Fig. 15) and under all other non-random ascertainment schemes explored on real (Suppl. Table 7, Suppl. Fig. 12d) and simulated data (Suppl. Fig. 7b). In contrast, all f_4-statistic classes we explored including one or two African groups and non-Africans, or non-Africans only, turned out to be unbiased under the 1240K ascertainment (Suppl. Table 7, Suppl. Fig. 12). Thus, numerous studies relying on fitting *qpAdm* and/or admixture graph models including one African group and various non-Africans are probably minimally affected by ascertainment bias, as we also demonstrated on exhaustive collections of simple admixture graphs for few population quintuplets (Fig. 1, Table 1, Suppl. Fig. 1, Suppl. Tables 3-5). When these classes of methods are applied to African population history, the situation is different, however. As we demonstrated, the 1240K panel emerges as biased when fits of simple admixture graphs including five African groups or one or two archaic and three or four African groups are considered (Fig. 1, Table 1, Suppl. Fig. 1, Suppl. Tables 3-5). We also demonstrated that the 1240K ascertainment affects fits of more complex admixture graphs including in all cases chimpanzee and Altai Neanderthal, and also four or six African groups and one or two groups with substantial non-African ancestry (Supp. Text 1, Suppl. Figs. 5 and 6). We expect fits of many other admixture graphs for Africans beyond those tested in this study to be affected by the 1240K ascertainment since the f_4-statistic classes African$_2$;archaic$_1$;non-African$_1$,

572 African$_2$;archaic$_1$;chimpanzee$_1$, and African$_3$;X are substantially biased under this
573 ascertainment (Suppl. Table 7). These effects were reproduced on simulated data when
574 accurate graphs including “chimpanzee”, one “archaic” lineage, and several “African” and
575 “non-African” lineages were fitted to the data ascertained in various ways (Fig. 2b).
576 In line with theoretical expectations, f_4-statistics including AMH groups only are largely
577 unbiased under archaic ascertainment (Skoglund et al. 2017, Bergström et al. 2020,
578 technical note published on the Daicel Arbor Biosciences product page). However, as
579 compared to other SNP panels of similar size, archaic ascertainment increases variance in
580 nearly all f_4-statistic classes of the types non-African$_3$;X and non-African$_4$ (Suppl. Table 7,
581 Suppl. Figs. 12-14). Increased variance in these cases can be explained by the low
582 information content of an archaically ascertained panel: unlike the other non-random
583 ascertainment schemes we tested, archaic ascertainment preserves most sites with nearly
584 fixed ancestral variants and leads to just a moderate enrichment for common variants (DAF
585 between 5% and 95%), especially if DAF is based on non-Africans (Suppl. Text 3, Suppl. Fig.
586 7d, Suppl. Table 10). Thus, the archaically ascertained panel includes a relatively small
587 number of variants that are common in AMH and especially in non-Africans (Suppl. Table
588 10), and that increases the noise level. This elevated noise level in *f*-statistics under archaic
589 ascertainment translates to reduced power to reject admixture graph models based on
590 these *f*-statistics (Fig. 1, Table 1, Suppl. Fig. 1, Suppl. Tables 3-5). This effect was also
591 reproduced on simulated data (Suppl. Fig. 10). If archaic humans are included in an *f*-
592 statistic or an admixture graph, archaic ascertainment is no longer guaranteed to be
593 unbiased (see Fig. 2 c,d), and indeed due to the existence of the Neanderthal to non-African
594 gene flow it fails to fix the bias affecting the most problematic class of statistics f_4(African X,
595 archaic; African Y, non-African), as demonstrated on simulated data in Suppl. Figs. 7b and
596 17.
597 Many ascertainment schemes such as the 1240K, 2000K, Illumina 650Y panels and MAF-
598 based ascertainment on non-Africans skew average DAF across populations in the
599 quadruplet since these panels are enriched for derived variants common in non-Africans vs.
600 Africans and in AMH vs. archaic humans (Suppl. Text 3, Suppl. Table 10). Overrepresentation
601 of derived variants in certain groups of the quadruplet skews f_4-statistics. We conclude that
602 two ascertainment schemes most often used for studies of African population history
603 (1240K and archaic ascertainment) are not optimal for various reasons: overrepresentation

of derived variants common in non-Africans in the former case and a small number of variants common in AMH in the latter case.

We found that there exists a non-outgroup ascertainment scheme that is less biased than the other schemes we tested: restricting to variants that are common in a diverse collection of African groups. This scheme demonstrated a bias only in the case of the f_4(African X, archaic; African Y, non-African) and f_4(non-African X, archaic; African, non-African Y) classes of f_4-statistics among the 27 classes investigated (Suppl. Table 7, Suppl. Figs. 12, 14, and 17). This scheme does not favor derived variants common in non-Africans and supplies many variants common in both Africans and non-Africans (Suppl. Table 10). While for many f_4-statitic classes and admixture graphs, the difference in performance of the pan-African and archaic ascertainment schemes is small (Table 1, Suppl. Figs. 3, 4, and 12, Suppl. Tables 3-5 and 7), the pan-African scheme is applicable when Neanderthals and Denisovans are co-analyzed (Suppl. Figs. 3 and 4), while archaic ascertainment generates extreme shifts in f_4-statistics in this case (Fig. 2b, Suppl. Fig. 18). The pan-African scheme is also effective for analyses focused on non-Africans, demonstrating no elevated noise level typical for archaic ascertainment (Suppl. Table 7, Suppl. Table 3). Thus, pan-African ascertainment is the most widely applicable scheme among those explored in this study. According to our results on collections of admixture graphs (Table 1) and on f_4-statistic classes (Suppl. Table 7), a similar form of ascertainment, namely combining sites heterozygous in a single San and a single Yoruba individual (Human Origins panels 4 & 5) is also largely unbiased, with the exception of statistics of the form f_4(African X, archaic; African Y, non-African). However, this ascertainment is also more noisy due to the low number of sites available.

As we demonstrated on simulated data, for a non-outgroup ascertainment to be unbiased it should be based on a population where the derived allele frequency spectrum of sites that were polymorphic at the root is preserved relatively well (Fig. 2e) (however, such an ascertainment usually has relatively low statistical power for rejection of incorrect admixture graph models, see Suppl. Fig. 10). We note that the group where ascertainment was performed was co-modelled with the other groups, as is often done in practice. In the light of these results, archaic ascertainment's sub-optimal performance as a non-outgroup ascertainment is due to the fact that Denisovans and Neanderthals have had a low long-term effective population size (Mafessoni et al. 2020), and thus are highly drifted with

respect to the root. Moreover, it is often unavoidable that individuals used for archaic ascertainment are used as sole representatives of the respective groups analyzed, and that is also problematic (Fig. 2c). Africans, in contrast, have had much higher effective population sizes (Mafessoni et al. 2020), and we propose that restricting to variants common in a diverse set of African genomes is much more reliable (than archaic ascertainment or ascertainment on a single African population or individual) for preserving the spectrum of variants that existed at the root of archaic and anatomically modern humans. At the same time, pan-African ascertainment supplies enough variants that are common in non-Africans, making it also relatively powerful statistically for analyses focused on non-Africans.

An enrichment approach is powerful for large-scale ancient DNA research in Africa due to DNA preservation issues in the hot climate (Skoglund et al. 2017). We did not test a range of allele frequency cut-offs or counts of individuals for pan-African ascertainment, and we do not propose a list of sites for a new DNA enrichment panel. However, our results imply that an effective approach for designing such a panel, which would also be useful for human archaeogenetic studies worldwide, would be to combine selection of the A/T and G/C mutation types with depletion of variants rare in Africa. Frequencies of alleles at A/T and G/C loci are not affected by biased gene conversion (its rate depends on population heterozygosity, Pouyet et al. 2018), and these loci are not hypermutable, and are not affected by deamination damage in ancient DNA. As we demonstrated, restricting to A/T and G/C sites does not bias f_4-statistics (Suppl. Table 7, Suppl. Figs. 12 and 16, Suppl. Table 10) or admixture graph fits (Table 1, Suppl. Tables 3-5). Another reason for taking AT/GC sites only is simply reducing the number of sites since enrichment reagents have limited capacity, and this ascertainment scheme with a 5% MAF threshold yields about 1.6 million variable sites on the “SGDP+archaic” dataset.

Methods

1. Simulating genetic data

1.1 Simulating the relationships of AMH and archaic humans with msprime v.0.7.4

Twenty-two chromosomes matching the size of the human chromosomes in the *hg19* assembly were simulated with a flat recombination rate (2×10^{-8} per nt per generation) and a flat mutation rate, 1.25×10^{-8} per nt per generation (Scally & Durbin 2012). The standard

667 coalescent simulation algorithm was used (Kelleher et al. 2016), and diploid genomes were
668 assembled from these independently simulated 22 haploid chromosomes. Although this
669 approach does not recapitulate the linkage disequilibrium pattern in real human genomes,
670 it does not make a difference for simulating allele frequencies in deeply divergent groups
671 since chromosome histories become quickly independent in the past (Nelson et al. 2020).
672 The following groups were simulated: chimpanzee ("Chimp", one individual sampled at
673 the end of the simulation), the Vindija Neanderthal ("Neanderthal", one individual sampled
674 2,000 generations or 50,000 years in the past, considering a generation time of 25 years),
675 the high-coverage Denisovan "Denisova 3" ("Denisovan", one individual sampled 2,000
676 generations in the past), five African groups (10 individuals per group sampled at "present")
677 and three non-African groups (10 individuals per group sampled at "present"). Five classes
678 of simulated topologies are shown in Suppl. Fig. 17b; for a full list of simulation parameters
679 and their values see Suppl. Table 13. Only one simulation iteration was performed for each
680 combination of parameters.
681 We applied archaic ascertainment to the simulated data: restricting to sites polymorphic
682 in the group composed of two "archaic" individuals, "Denisovan" and "Neanderthal" (this
683 scheme reproduces the archaic ascertainment applied to real data, the "SGDP+archaic"
684 dataset, in Suppl. Figs. 17a and 18a). For calculating *f*-statistics on unascertained and
685 ascertained SNP sets, the software package *ADMIXTOOLS 2* (Maier et al. 2022 preprint) was
686 used. Since there was no missing data and all individuals were diploid, we first calculated all
687 possible f_2-statistics for 4 Mbp-sized genome blocks (with the "*maxmiss=0*",
688 "*adjust_pseudohaploid=FALSE*", and "*minac2=FALSE*" settings), and then used them for
689 calculating f_4-statistics as linear combinations of f_2-statistics. This protocol was used for
690 generating the results shown in Suppl. Figs. 17c and 18c.
691
692 *1.2 Simulating the relationships of AMH and archaic humans with msprime v.1.1.1*
693 More realistic simulations were performed with *msprime v.1.1.1* which allows accurate
694 simulation of recombination and of multi-chromosome diploid genomes relying on the
695 Wright-Fisher model (Nelson et al. 2020, Baumdicker et al. 2022). We simulated three
696 chromosomes (each 100 Mb long) in a diploid genome by specifying a flat recombination
697 rate (2×10^{-8} per nt per generation) along the chromosome and a much higher rate at the
698 chromosome boundaries ($\log_e 2$ or ~0.693 per nt per generation, see

699 https://tskit.dev/msprime/docs/stable/ancestry.html#multiple-chromosomes). A flat
700 mutation rate, 1.25 x 10^{-8} per nt per generation (Scally & Durbin 2012), and the binary
701 mutation model were used. To maintain the correct correlation between chromosomes, the
702 discrete time Wright-Fischer model was used for 25 generations into the past, and deeper in
703 the past the standard coalescent simulation algorithm was used (as recommended by
704 Nelson et al. 2020).

705 The following groups were simulated: chimpanzee ("c", one individual sampled at the
706 end of the simulation), the Altai Neanderthal ("n1", one individual sampled 3,790
707 generations in the past), the Vindija Neanderthal ("n2", one individual sampled 1,700
708 generations in the past), the high-coverage Denisovan "Denisova 3" ("d", one individual
709 sampled 1,700 generations in the past), two African groups ("a1" and "a2", 10 individuals
710 per group sampled at "present") and two non-African groups ("na1" and "na2", 10
711 individuals per group sampled at "present"). The topology, dates and some effective
712 population sizes are shown in Fig. 2a; for a full list of simulation parameters see Suppl. Table
713 13. Ten simulation iterations were performed for each combination of parameters, and two
714 combinations were tested: with or without the Neanderthal to non-African gene flow.

715 Upon assessing genetic distances across the simulated groups using F_{ST}, the following
716 ascertainment schemes were applied:

717 1. restricting to sites that are heterozygous in a randomly selected individual from the
718 "a2" group (this scheme simulates the generation of one Human Origins SNP panel,
719 Patterson et al. 2012);
720 2. taking heterozygous sites from one randomly selected individual per "AMH"
721 population ("a1", "a2", "na1", "na2") and merging these SNP sets (this scheme
722 simulates the generation of the whole Human Origins SNP array, Patterson et al.
723 2012);
724 3. restricting to sites having high minor allele frequency (> 5%) in the union of "African"
725 groups "a1" and "a2" (this scheme simulates the MAF ascertainment on Africans);
726 4. restricting to sites having high minor allele frequency (> 5%) in the union of "non-
727 African" groups "na1" and "na2" (this scheme simulates the MAF ascertainment on a
728 non-African continental meta-population);

5. restricting to sites having high minor allele frequency (> 5%) in the union of all "AMH" groups "a1", "a2", "na1", and "na2" (this scheme simulates the global MAF ascertainment);
6. restricting to sites polymorphic in the group composed of three "archaic" individuals, "d", "n1", and "n2" (this scheme simulates the archaic ascertainment applied to the "SGDP+archaic" dataset throughout this study).

1.3 Simulating random admixture graphs and simple trees with msprime v.1.1.1

Genetic histories in the form of random admixture graphs were simulated using the *msprime v.1.1.1* settings described above. We simulated admixture graphs of four complexity classes: the graphs included 8 or 9 non-outgroup populations, one outgroup (all sampled at leaves), and 4 or 5 pulse-like admixture events. Demographic events were separated by date intervals ranging randomly between 1,500 and 8,000 generations, with an upper bound on the tree depth at 40,000 generations. To be more precise, demographic events were not placed in time entirely randomly, but were tied to one or few other events of the same "topological depth" within the graph, as illustrated by examples of the simulated topologies in Suppl. Fig. 8. The same principle was applied to the sampling dates for non-root groups, which were tied to other demographic events such as divergence and admixture of other populations. The random graph topologies and simulated parameter sets were generated using the *random_sim* function from the *ADMIXTOOLS 2* package: https://uqrmaie1.github.io/admixtools/reference/random_sim.html

Outgroups facilitate automated exploration of graph topology space. Outgroup branches diverged from the other populations at 40,000 generation in the past and had a large constant effective population size of 100,000 diploid individuals. Other effective population sizes were constant along each edge and were picked randomly from the range of 2,000-40,000 diploid individuals. Admixture proportions for all admixture events varied randomly between 10% and 40%. The root of the simulation and the root of all non-outgroup populations were sampled, and the other populations were sampled at branch tips exclusively. This setup generates groups sampled at widely different dates in the past (from 0 to ca. 40,000 generations) or, in other words, located at various genetic distances from the root (Fig. 2d). The outgroup population was sampled at the "present" of the simulation. Sample sizes for all populations were identical: 10 diploid individuals with no missing data.

For subsequent analyses we selected only simulations where pairwise F_{ST} for groups were in the range characteristic for anatomically modern and archaic humans (in each simulation there was at least one F_{ST} value below 0.15; see Suppl. Fig. 8). In this way, 20 random topologies were simulated per complexity class. Each topology was simulated only once, and 80 simulations were generated in total (see examples of the topologies and respective F_{ST} distributions in Suppl. Fig. 8). Another set of simulations was prepared with the same topologies and parameters, except for the effective population size on the outgroup branch which was set at 1,000 diploid individuals instead of 100,000.

The following ascertainment schemes were applied to the outcomes of these randomized simulations: 1) ascertainment on sites heterozygous in a single randomly selected individual (this Human Origins-like ascertainment was repeated for all simulated groups including the outgroup and root groups, generating 920 ascertained datasets); 2) unions of four such Human Origins-like SNP panels, with only one individual per group considered (10 random sets of four groups excluding the outgroup and root groups were explored per topology, generating 800 ascertained datasets); 3) ascertainment on sites polymorphic in a group composed of three randomly selected individuals, with only one individual per group considered (10 random sets of three groups excluding the outgroup and root groups were explored per topology, generating 800 ascertained datasets); and 4) MAF ascertainment, that is restricting to sites having MAF > 5% in random meta-groups (10 random sets of four groups excluding the outgroup and root groups were explored per topology, generating 800 ascertained datasets). Group sets used for each ascertainment were recorded. Genetic distances (F_{ST}) were calculated for all populations (including the outgroup and the last common ancestor of all non-outgroup populations) vs. the root sample (Fig. 2d).

Alternatively, simple trees were simulated using the *msprime v.1.1.1* settings described above. A tree of four groups conforming to the f_4-statistic (A, B; C, O) was simulated using *msprime v.1.1.1*, with a tree depth of 4,000 generations (Suppl. Fig. 11). All the groups had a uniform effective population size of 100,000 diploid individuals, except for a bottleneck happening immediately after the A-B divergence (at 1,999 generations in the past) and lasting until the end of the simulation. The following bottleneck classes were simulated: no bottleneck (control), 10x, 100x, 1,000x, and 10,000x reduction in effective population size. For each bottleneck class, 20 independent simulations were performed. All the samples were drawn at “present”: sample sizes were 25, 25, 25 and 10 for populations A, B, C and O,

793 respectively (except for the 10,000x bottleneck class since group A included 10 individuals
794 only in that case). Three ascertainment schemes were tested for the simulated trees: 1)
795 ascertainment on sites heterozygous in a single randomly selected individual (this Human
796 Origins-like ascertainment was repeated for all simulated groups, including group O); 2)
797 restricting to sites having MAF > 5% (or 10%, or 2.5%) in the union of groups A and B
798 composed of 50 diploid individuals; and 3) removal of sites with *derived* allele frequency >
799 95% (or 90%, or 97.5%) in the union of groups A and B. The latter ascertainment scheme was
800 added since the ascertainment schemes we tested on real data deplete the derived end of
801 the allele frequency spectrum more than the ancestral end (Suppl. Tables 10-12).
802 For calculating *f*-statistics and fitting admixture graphs to unascertained and ascertained
803 SNP sets, the *ADMIXTOOLS 2* (Maier et al. 2022 preprint) software package was used. Since
804 there was no missing data and all individuals were diploid, we first calculated all possible f_2-
805 statistics for 4 Mbp-sized genome blocks (with the "*maxmiss=0*",
806 "*adjust_pseudohaploid=FALSE*", and "*minac2=FALSE*" settings) and then used them for
807 calculating f_4-statistics as linear combinations of f_2-statistics or for fitting admixture graphs
808 (with the "*numstart=100*" and "*diag=0.0001*" settings). This calculation protocol was used
809 for generating the results shown in Figs. 2 and in Suppl. Figs. 7-11. When true admixture
810 graphs were fitted to ascertained data, full population samples of 10 individuals were used
811 by default, and in some cases, as indicated in the figure legends, the individual used for
812 ascertainment was used as the only representative of the respective population. Outgroups
813 were included in the fitted graphs; in other words, they were co-analyzed with the other
814 groups. Outgroups were not co-modelled with the other populations in Fig. 2f only.
815 For assessing the power of ascertained simulated datasets to reject incorrect admixture
816 graph models, we first generated a set of such incorrect graphs per each simulated
817 topology. For that purpose, an algorithm for finding well-fitting topologies (*findGraphs* from
818 the *ADMIXTOOLS 2* package) was started on non-ascertained data 300 times, seeded by
819 random graphs containing either the simulated number of admixture events (*n*, 100 runs),
820 or *n-1* events (100 runs), or *n+1* events (100 runs). For a list of settings for the *findGraphs*
821 algorithm see Maier *et al*. (2022 preprint). Thousands of diverse graphs explored by
822 *findGraphs* in the process of topology optimization were generated in this way for each
823 simulated graph, and 100 poorly fitting graphs were randomly picked from a subset of these

graphs having LL scores between 70 and 300. This subset of graphs was then fitted to all ascertained datasets derived from the same simulated admixture graph.

2. Constructing the set of real data

We used the *cteam-lite* dataset described in Mallick *et al*. 2016, composed of the full SGDP set (300 high-coverage genomes from present-day populations), the chimpanzee genome (pseudo-haploid genotype calls, see http://hgdownload.cse.ucsc.edu/goldenPath/panTro2/bigZips/), and the Altai Neanderthal, "Denisova 3" Denisovan, Ust'-Ishim, WHG Loschbour, and LBK Stuttgart ancient genomes (see SI section 3 in that study). We supplemented *cteam-lite* by 44 present-day African genomes sequenced using the SGDP protocols by Fan *et al*. (2019), the Vindija Neanderthal's genome (Prüfer et al. 2017), and the genome of an ancient African forager individual I10871 sequenced by Lipson *et al*. (2020) (Suppl. Table 1). Sites polymorphic in this set of 352 individuals were extracted from the *cteam-lite* files of the "hetfa" format using the *cpoly* tool (Mallick et al. 2016): alleles were grouped into derived and ancestral (polarized) according to the chimpanzee genome; missing data and heterozygous sites were allowed. For each genome, we used individual base quality masks included in *cteam-lite* or constructed using the same protocol for other genomes (Vindija Neanderthal and Fan et al. 2019): minimum base quality was set by default at 1, as recommended in SI section 3 in Mallick *et al*. 2016, which discarded lowest-quality regions marked as "0", "?", or "N". The individual I10871 was not included in most analyses in this study (except for the complex admixture graphs in Suppl. Text 1) due to its relatively high rate of deamination errors.

The resulting dataset prior to missing data removal and ascertainment includes 94,691,841 autosomal sites (Suppl. Table 2). To keep the polarity of alleles, all data manipulations and ascertainments were performed using *PLINK v.2.0 alpha* (Chang et al. 2015). For calculating f_4-statistics, sets of continental-level meta-populations were selected (e.g., Africans and East Asians or Africans and archaic humans) and then f_4-statistics were calculated for all possible combinations of populations in the resulting subset of the "SGDP+archaic" dataset, with no missing data (at the population level) allowed *within the selected subset*. This was done to avoid potential biases associated with data missing non-randomly across groups. Alternatively, f_4-statistics were drawn randomly from a certain

class of statistics, and no missing data (at the population level) were allowed in the resulting *population quadruplets*.

3. Influence of ascertainment on fits of admixture graphs to real data

First, we fitted all possible graphs including two admixture events (32,745 distinct topologies with no fixed outgroup) for three combinations of groups: 1) one archaic individual, three African groups, and one African group with substantial West Eurasian-related ancestry (Altai Neanderthal, Ju|'hoan North, Biaka, Yoruba, and Agaw, respectively); 2) five deeply divergent ancient and present-day non-African groups (Ust'-Ishim, Papuan, Onge, LBK Stuttgart, Even); and 3) five deeply divergent present-day non-African groups (Papuan, Onge, Palestinian, Even, Mala). These three sets of simple graphs were fitted to all sites, AT/GC sites, and 1240K sites (no missing data were allowed at the group level within these sets of five populations); 5,000 best-fitting models were selected according to the LL score on all sites and WRs of those models were compared across SNP sets (Suppl. Fig. 1).

Next, we explored the same exhaustive set of admixture graph topologies including five groups and two admixture events on the wider collection of ascertainments. Twelve combinations of five groups including up to two archaic humans, up to five African groups, and up to five non-African groups were tested. To ensure fair comparison across at least a subset of population combinations, as a starting point for generating ascertained site sets we used either 11,706,773 sites (with no missing data at the group level) polymorphic in a set of 48 archaic and African groups composed of a total of 97 individuals; or 10,051,585 such sites in a set of 59 archaic, African, European, and Middle Eastern groups composed of a total of 120 individuals; or 5,296,653 such sites in a set of 51 Papuan, Native American, European, Anatolian, and Caucasian groups composed of a total of 112 individuals (Suppl. Tables 1 and 2).

We examined the fits of these collections of admixture graphs from different perspectives. (1) We considered just 5,000 topologies that are best-fitting on the unascertained site set (Suppl. Figs. 1-3) or all 32,745 topologies tested (Suppl. Fig. 4). (2) We also considered alternative admixture graph fit metrics, LL or WR. LL as a fit metric (see left-hand panels in Suppl. Figs. 3 and 4) is more accurate than WR, but difficult to compare across population sets. Finally, instead of R or R^2 of a linear trend as a measure of correlation of admixture graph fits (Fig. 1, Suppl. Figs. 1-4) we considered the fraction of all

887 possible models of a certain complexity that are rejected under ascertainment (WR > 3 SE)
888 but accepted on all sites (WR < 3 SE), or *vice versa*.
889
890 *4. Automated inference of fitting admixture graphs on real data*
891 The 12-population admixture graph published by Lipson et al. 2020 (and later used as a
892 skeleton graph in Lipson et al. 2022) and simpler 7- and 10-population intermediate graphs
893 presented in the former study were revisited by Maier *et al*. (2022 preprint), and thousands
894 of alternative well-fitting graphs of the same complexity were found using the *find_graphs*
895 function from the *ADMIXTOOLS 2* package
896 (https://uqrmaie1.github.io/admixtools/articles/graphs.html). Maier *et al.* used the 1240K
897 dataset only, and in the current study we re-fitted the admixture graphs found by the
898 algorithm on the 1240K SNP panel to the AT/GC and unascertained datasets derived from
899 "SGDP+archaic" and also repeated automated admixture graph inference on these two
900 additional SNP sets. Advantages and pitfalls of automated admixture graph inference are
901 described in detail in Maier *et al.*, along with justifications for the specific protocol used in
902 that study, and here we used protocols identical to those employed by Maier *et al.* We first
903 calculated all possible f_2-statistics for 4 Mbp-sized genome blocks (with the "*maxmiss=0*",
904 "*adjust_pseudohaploid=FALSE*", and "*minac2=2*" settings, see Maier *et al*. 2022 for details
905 on the settings) and then used them for fitting admixture graphs (with the "*numstart=100*"
906 and "*diag=0.0001*" settings) and for automated admixture graph inference with the
907 *find_graphs* function (see the Methods section in Maier *et al.* for a complete list of
908 arguments for this function). Only one topology constraint was used at the graph space
909 exploration step: chimpanzee was assigned as an outgroup.
910
911

912 ## References

913 1. Albrechtsen A, Nielsen FC, Nielsen R. Ascertainment biases in SNP chips affect measures of
914 population divergence. Mol Biol Evol. 2010 Nov;27(11):2534-47. doi: 10.1093/molbev/msq148.
915 2. Baumdicker F, Bisschop G, Goldstein D, Gower G, Ragsdale AP, Tsambos G, Zhu S, Eldon B,
916 Ellerman EC, Galloway JG, Gladstein AL, Gorjanc G, Guo B, Jeffery B, Kretzschumar WW, Lohse K,
917 Matschiner M, Nelson D, Pope NS, Quinto-Cortés CD, Rodrigues MF, Saunack K, Sellinger T,
918 Thornton K, van Kemenade H, Wohns AW, Wong Y, Gravel S, Kern AD, Koskela J, Ralph PL,
919 Kelleher J. Efficient ancestry and mutation simulation with msprime 1.0. Genetics. 2022 Mar
920 3;220(3):iyab229. doi: 10.1093/genetics/iyab229.

3. Bergström A, McCarthy SA, Hui R, Almarri MA, Ayub Q, Danecek P, Chen Y, Felkel S, Hallast P, Kamm J, Blanché H, Deleuze JF, Cann H, Mallick S, Reich D, Sandhu MS, Skoglund P, Scally A, Xue Y, Durbin R, Tyler-Smith C. Insights into human genetic variation and population history from 929 diverse genomes. Science. 2020 Mar 20;367(6484):eaay5012. doi: 10.1126/science.aay5012.
4. Bergström A, Frantz L, Schmidt R, Ersmark E, Lebrasseur O, Girdland-Flink L, Lin AT, Storå J, Sjögren KG, Anthony D, Antipina E, Amiri S, Bar-Oz G, Bazaliiskii VI, Bulatović J, Brown D, Carmagnini A, Davy T, Fedorov S, Fiore I, Fulton D, Germonpré M, Haile J, Irving-Pease EK, Jamieson A, Janssens L, Kirillova I, Horwitz LK, Kuzmanovic-Cvetković J, Kuzmin Y, Losey RJ, Dizdar DL, Mashkour M, Novak M, Onar V, Orton D, Pasarić M, Radivojević M, Rajković D, Roberts B, Ryan H, Sablin M, Shidlovskiy F, Stojanović I, Tagliacozzo A, Trantalidou K, Ullén I, Villaluenga A, Wapnish P, Dobney K, Götherström A, Linderholm A, Dalén L, Pinhasi R, Larson G, Skoglund P. Origins and genetic legacy of prehistoric dogs. Science. 2020b Oct 30;370(6516):557-564. doi: 10.1126/science.aba9572.
5. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015 Feb 25;4:7. doi: 10.1186/s13742-015-0047-8.
6. Clark AG, Hubisz MJ, Bustamante CD, Williamson SH, Nielsen R. Ascertainment bias in studies of human genome-wide polymorphism. Genome Res. 2005 Nov;15(11):1496-502. doi: 10.1101/gr.4107905.
7. Durand EY, Patterson N, Reich D, Slatkin M. Testing for ancient admixture between closely related populations. Mol Biol Evol. 2011 Aug;28(8):2239-52. doi: 10.1093/molbev/msr048.
8. Durvasula A, Sankararaman S. Recovering signals of ghost archaic introgression in African populations. Sci Adv. 2020 Feb 12;6(7):eaax5097. doi: 10.1126/sciadv.aax5097.
9. Fan S, Kelly DE, Beltrame MH, Hansen MEB, Mallick S, Ranciaro A, Hirbo J, Thompson S, Beggs W, Nyambo T, Omar SA, Meskel DW, Belay G, Froment A, Patterson N, Reich D, Tishkoff SA. African evolutionary history inferred from whole genome sequence data of 44 indigenous African populations. Genome Biol. 2019 Apr 26;20(1):82. doi: 10.1186/s13059-019-1679-2. Erratum in: Genome Biol. 2019 Oct 9;20(1):204.
10. Fischer A, Pollack J, Thalmann O, Nickel B, Pääbo S. Demographic history and genetic differentiation in apes. Curr Biol. 2006 Jun 6;16(11):1133-8. doi: 10.1016/j.cub.2006.04.033.
11. Fu Q, Li H, Moorjani P, Jay F, Slepchenko SM, Bondarev AA, Johnson PL, Aximu-Petri A, Prüfer K, de Filippo C, Meyer M, Zwyns N, Salazar-García DC, Kuzmin YV, Keates SG, Kosintsev PA, Razhev DI, Richards MP, Peristov NV, Lachmann M, Douka K, Higham TF, Slatkin M, Hublin JJ, Reich D, Kelso J, Viola TB, Pääbo S. Genome sequence of a 45,000-year-old modern human from western Siberia. Nature. 2014 Oct 23;514(7523):445-9. doi: 10.1038/nature13810.
12. Fu Q, Hajdinjak M, Moldovan OT, Constantin S, Mallick S, Skoglund P, Patterson N, Rohland N, Lazaridis I, Nickel B, Viola B, Prüfer K, Meyer M, Kelso J, Reich D, Pääbo S. An early modern human from Romania with a recent Neanderthal ancestor. Nature. 2015 Aug 13;524(7564):216-9. doi: 10.1038/nature14558.
13. Geibel J, Reimer C, Weigend S, Weigend A, Pook T, Simianer H. How array design creates SNP ascertainment bias. PLoS One. 2021 Mar 30;16(3):e0245178. doi: 10.1371/journal.pone.0245178.
14. Gokcumen O. Archaic hominin introgression into modern human genomes. Am J Phys Anthropol. 2020 May;171 Suppl 70:60-73. doi: 10.1002/ajpa.23951.
15. Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH, Hansen NF, Durand EY, Malaspinas AS, Jensen JD, Marques-Bonet T, Alkan C, Prüfer K, Meyer M, Burbano HA, Good JM, Schultz R, Aximu-Petri A, Butthof A, Höber B, Höffner B, Siegemund M, Weihmann A, Nusbaum C, Lander ES, Russ C, Novod N, Affourtit J, Egholm M, Verna C, Rudan P, Brajkovic D, Kucan Ž, Gušic I, Doronichev VB, Golovanova LV, Lalueza-Fox C, de la Rasilla M, Fortea J, Rosas A, Schmitz RW, Johnson PLF, Eichler EE, Falush D, Birney E, Mullikin

971 JC, Slatkin M, Nielsen R, Kelso J, Lachmann M, Reich D, Pääbo S. A draft sequence of the
972 Neandertal genome. Science. 2010 May 7;328(5979):710-722. doi: 10.1126/science.1188021.
973 16. Guillot G, Foll M. Correcting for ascertainment bias in the inference of population structure.
974 Bioinformatics. 2009 Feb 15;25(4):552-4. doi: 10.1093/bioinformatics/btn665.
975 17. Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, Brandt G, Nordenfelt S, Harney
976 E, Stewardson K, Fu Q, Mittnik A, Bánffy E, Economou C, Francken M, Friederich S, Pena RG,
977 Hallgren F, Khartanovich V, Khokhlov A, Kunst M, Kuznetsov P, Meller H, Mochalov O, Moiseyev
978 V, Nicklisch N, Pichler SL, Risch R, Rojo Guerra MA, Roth C, Szécsényi-Nagy A, Wahl J, Meyer M,
979 Krause J, Brown D, Anthony D, Cooper A, Alt KW, Reich D. Massive migration from the steppe
980 was a source for Indo-European languages in Europe. Nature. 2015 Jun 11;522(7555):207-11.
981 doi: 10.1038/nature14317.
982 18. Hajdinjak M, Mafessoni F, Skov L, Vernot B, Hübner A, Fu Q, Essel E, Nagel S, Nickel B, Richter J,
983 Moldovan OT, Constantin S, Endarova E, Zahariev N, Spasov R, Welker F, Smith GM, Sinet-
984 Mathiot V, Paskulin L, Fewlass H, Talamo S, Rezek Z, Sirakova S, Sirakov N, McPherron SP,
985 Tsanova T, Hublin JJ, Peter BM, Meyer M, Skoglund P, Kelso J, Pääbo S. Initial Upper Palaeolithic
986 humans in Europe had recent Neanderthal ancestry. Nature. 2021 Apr;592(7853):253-257. doi:
987 10.1038/s41586-021-03335-3.
988 19. Harney É, Patterson N, Reich D, Wakeley J. Assessing the performance of qpAdm: a statistical
989 tool for studying population admixture. Genetics. 2021 Apr 15;217(4):iyaa045. doi:
990 10.1093/genetics/iyaa045.
991 20. Hubisz MJ, Williams AL, Siepel A. Mapping gene flow between ancient hominins through
992 demography-aware inference of the ancestral recombination graph. PLoS Genet. 2020 Aug
993 6;16(8):e1008895. doi: 10.1371/journal.pgen.1008895.
994 21. Jouganous J, Long W, Ragsdale AP, Gravel S. Inferring the Joint Demographic History of Multiple
995 Populations: Beyond the Diffusion Approximation. Genetics. 2017 Jul;206(3):1549-1567. doi:
996 10.1534/genetics.117.200493.
997 22. Kelleher J, Etheridge AM, McVean G. Efficient Coalescent Simulation and Genealogical Analysis
998 for Large Sample Sizes. PLoS Comput Biol. 2016 May 4;12(5):e1004842. doi:
999 10.1371/journal.pcbi.1004842.
1000 23. Kılınç GM, Kashuba N, Koptekin D, Bergfeldt N, Dönertaş HM, Rodríguez-Varela R, Shergin D,
1001 Ivanov G, Kichigin D, Pestereva K, Volkov D, Mandryka P, Kharinskii A, Tishkin A, Ineshin E,
1002 Kovychev E, Stepanov A, Dalén L, Günther T, Kırdök E, Jakobsson M, Somel M, Krzewińska M,
1003 Storå J, Götherström A. Human population dynamics and Yersinia pestis in ancient northeast
1004 Asia. Sci Adv. 2021 Jan 6;7(2):eabc4587. doi: 10.1126/sciadv.abc4587.
1005 24. Kuhlwilm M, Gronau I, Hubisz MJ, de Filippo C, Prado-Martinez J, Kircher M, Fu Q, Burbano HA,
1006 Lalueza-Fox C, de la Rasilla M, Rosas A, Rudan P, Brajkovic D, Kucan Ž, Gušic I, Marques-Bonet T,
1007 Andrés AM, Viola B, Pääbo S, Meyer M, Siepel A, Castellano S. Ancient gene flow from early
1008 modern humans into Eastern Neanderthals. Nature. 2016 Feb 25;530(7591):429-33. doi:
1009 10.1038/nature16544.
1010 25. Lachance J, Tishkoff SA. SNP ascertainment bias in population genetic analyses: why it is
1011 important, and how to correct it. Bioessays. 2013 Sep;35(9):780-6. doi:
1012 10.1002/bies.201300014.
1013 26. Lazaridis I, Patterson N, Mittnik A, Renaud G, Mallick S, Kirsanow K, Sudmant PH, Schraiber JG,
1014 Castellano S, Lipson M, Berger B, Economou C, Bollongino R, Fu Q, Bos KI, Nordenfelt S, Li H, de
1015 Filippo C, Prüfer K, Sawyer S, Posth C, Haak W, Hallgren F, Fornander E, Rohland N, Delsate D,
1016 Francken M, Guinet JM, Wahl J, Ayodo G, Babiker HA, Bailliet G, Balanovska E, Balanovsky O,
1017 Barrantes R, Bedoya G, Ben-Ami H, Bene J, Berrada F, Bravi CM, Brisighelli F, Busby GB, Cali F,
1018 Churnosov M, Cole DE, Corach D, Damba L, van Driem G, Dryomov S, Dugoujon JM, Fedorova SA,
1019 Gallego Romero I, Gubina M, Hammer M, Henn BM, Hervig T, Hodoglugil U, Jha AR, Karachanak-
1020 Yankova S, Khusainova R, Khusnutdinova E, Kittles R, Kivisild T, Klitz W, Kučinskas V,
1021 Kushniarevich A, Laredj L, Litvinov S, Loukidis T, Mahley RW, Melegh B, Metspalu E, Molina J,

Mountain J, Näkkäläjärvi K, Nesheva D, Nyambo T, Osipova L, Parik J, Platonov F, Posukh O, Romano V, Rothhammer F, Rudan I, Ruizbakiev R, Sahakyan H, Sajantila A, Salas A, Starikovskaya EB, Tarekegn A, Toncheva D, Turdikulova S, Uktveryte I, Utevska O, Vasquez R, Villena M, Voevoda M, Winkler CA, Yepiskoposyan L, Zalloua P, Zemunik T, Cooper A, Capelli C, Thomas MG, Ruiz-Linares A, Tishkoff SA, Singh L, Thangaraj K, Villems R, Comas D, Sukernik R, Metspalu M, Meyer M, Eichler EE, Burger J, Slatkin M, Pääbo S, Kelso J, Reich D, Krause J. Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature. 2014 Sep 18;513(7518):409-13. doi: 10.1038/nature13673.

27. Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, Cann HM, Barsh GS, Feldman M, Cavalli-Sforza LL, Myers RM. Worldwide human relationships inferred from genome-wide patterns of variation. Science. 2008 Feb 22;319(5866):1100-4. doi: 10.1126/science.1153717.
28. Librado P, Khan N, Fages A, Kusliy MA, Suchan T, Tonasso-Calvière L, Schiavinato S, Alioglu D, Fromentier A, Perdereau A, Aury JM, Gaunitz C, Chauvey L, Seguin-Orlando A, Der Sarkissian C, Southon J, Shapiro B, Tishkin AA, Kovalev AA, Alquraishi S, Alfarhan AH, Al-Rasheid KAS, Seregély T, Klassen L, Iversen R, Bignon-Lau O, Bodu P, Olive M, Castel JC, Boudadi-Maligne M, Alvarez N, Germonpré M, Moskal-Del Hoyo M, Wilczyński J, Pospuła S, Lasota-Kuś A, Tunia K, Nowak M, Rannamäe E, Saarma U, Boeskorov G, Lōugas L, Kyselý R, Peške L, Bălăşescu A, Dumitraşcu V, Dobrescu R, Gerber D, Kiss V, Szécsényi-Nagy A, Mende BG, Gallina Z, Somogyi K, Kulcsár G, Gál E, Bendrey R, Allentoft ME, Sirbu G, Dergachev V, Shephard H, Tomadini N, Grouard S, Kasparov A, Basilyan AE, Anisimov MA, Nikolskiy PA, Pavlova EY, Pitulko V, Brem G, Wallner B, Schwall C, Keller M, Kitagawa K, Bessudnov AN, Bessudnov A, Taylor W, Magail J, Gantulga JO, Bayarsaikhan J, Erdenebaatar D, Tabaldiev K, Mijiddorj E, Boldgiv B, Tsagaan T, Pruvost M, Olsen S, Makarewicz CA, Valenzuela Lamas S, Albizuri Canadell S, Nieto Espinet A, Iborra MP, Lira Garrido J, Rodríguez González E, Celestino S, Olària C, Arsuaga JL, Kotova N, Pryor A, Crabtree P, Zhumatayev R, Toleubaev A, Morgunova NL, Kuznetsova T, Lordkipanize D, Marzullo M, Prato O, Bagnasco Gianni G, Tecchiati U, Clavel B, Lepetz S, Davoudi H, Mashkour M, Berezina NY, Stockhammer PW, Krause J, Haak W, Morales-Muñiz A, Benecke N, Hofreiter M, Ludwig A, Graphodatsky AS, Peters J, Kiryushin KY, Iderkhangai TO, Bokovenko NA, Vasiliev SK, Seregin NN, Chugunov KV, Plasteeva NA, Baryshnikov GF, Petrova E, Sablin M, Ananyevskaya E, Logvin A, Shevnina I, Logvin V, Kalieva S, Loman V, Kukushkin I, Merz I, Merz V, Sakenov S, Varfolomeyev V, Usmanova E, Zaibert V, Arbuckle B, Belinskiy AB, Kalmykov A, Reinhold S, Hansen S, Yudin AI, Vybornov AA, Epimakhov A, Berezina NS, Roslyakova N, Kosintsev PA, Kuznetsov PF, Anthony D, Kroonen GJ, Kristiansen K, Wincker P, Outram A, Orlando L. The origins and spread of domestic horses from the Western Eurasian steppes. Nature. 2021 Oct;598(7882):634-640. doi: 10.1038/s41586-021-04018-9.
29. Lipson M, Cheronet O, Mallick S, Rohland N, Oxenham M, Pietrusewsky M, Pryce TO, Willis A, Matsumura H, Buckley H, Domett K, Nguyen GH, Trinh HH, Kyaw AA, Win TT, Pradier B, Broomandkhoshbacht N, Candilio F, Changmai P, Fernandes D, Ferry M, Gamarra B, Harney E, Kampuansai J, Kutanan W, Michel M, Novak M, Oppenheimer J, Sirak K, Stewardson K, Zhang Z, Flegontov P, Pinhasi R, Reich D. Ancient genomes document multiple waves of migration in Southeast Asian prehistory. Science. 2018 Jul 6;361(6397):92-95. doi: 10.1126/science.aat3188.
30. Lipson M, Ribot I, Mallick S, Rohland N, Olalde I, Adamski N, Broomandkhoshbacht N, Lawson AM, López S, Oppenheimer J, Stewardson K, Asombang RN, Bocherens H, Bradman N, Culleton BJ, Cornelissen E, Crevecoeur I, de Maret P, Fomine FLM, Lavachery P, Mindzie CM, Orban R, Sawchuk E, Semal P, Thomas MG, Van Neer W, Veeramah KR, Kennett DJ, Patterson N, Hellenthal G, Lalueza-Fox C, MacEachern S, Prendergast ME, Reich D. Ancient West African foragers in the context of African population history. Nature. 2020 Jan;577(7792):665-670. doi: 10.1038/s41586-020-1929-1.
31. Lipson M, Sawchuk EA, Thompson JC, Oppenheimer J, Tryon CA, Ranhorn KL, de Luna KM, Sirak KA, Olalde I, Ambrose SH, Arthur JW, Arthur KJW, Ayodo G, Bertacchi A, Cerezo-Román JI,

Culleton BJ, Curtis MC, Davis J, Gidna AO, Hanson A, Kaliba P, Katongo M, Kwekason A, Laird MF, Lewis J, Mabulla AZP, Mapemba F, Morris A, Mudenda G, Mwafulirwa R, Mwangomba D, Ndiema E, Ogola C, Schilt F, Willoughby PR, Wright DK, Zipkin A, Pinhasi R, Kennett DJ, Manthi FK, Rohland N, Patterson N, Reich D, Prendergast ME. Ancient DNA and deep population structure in sub-Saharan African foragers. Nature. 2022 Mar;603(7900):290-296. doi: 10.1038/s41586-022-04430-9.

32. Mafessoni F, Grote S, de Filippo C, Slon V, Kolobova KA, Viola B, Markin SV, Chintalapati M, Peyrégne S, Skov L, Skoglund P, Krivoshapkin AI, Derevianko AP, Meyer M, Kelso J, Peter B, Prüfer K, Pääbo S. A high-coverage Neandertal genome from Chagyrskaya Cave. Proc Natl Acad Sci U S A. 2020 Jun 30;117(26):15132-15136. doi: 10.1073/pnas.2004944117.
33. Maier R, Flegontov P, Flegontova O, Changmai P, Reich D. On the limits of fitting complex models of population history to genetic data. bioRxiv 2022.05.08.491072. doi: https://doi.org/10.1101/2022.05.08.491072.
34. Mallick S, Li H, Lipson M, Mathieson I, Gymrek M, Racimo F, Zhao M, Chennagiri N, Nordenfelt S, Tandon A, Skoglund P, Lazaridis I, Sankararaman S, Fu Q, Rohland N, Renaud G, Erlich Y, Willems T, Gallo C, Spence JP, Song YS, Poletti G, Balloux F, van Driem G, de Knijff P, Romero IG, Jha AR, Behar DM, Bravi CM, Capelli C, Hervig T, Moreno-Estrada A, Posukh OL, Balanovska E, Balanovsky O, Karachanak-Yankova S, Sahakyan H, Toncheva D, Yepiskoposyan L, Tyler-Smith C, Xue Y, Abdullah MS, Ruiz-Linares A, Beall CM, Di Rienzo A, Jeong C, Starikovskaya EB, Metspalu E, Parik J, Villems R, Henn BM, Hodoglugil U, Mahley R, Sajantila A, Stamatoyannopoulos G, Wee JT, Khusainova R, Khusnutdinova E, Litvinov S, Ayodo G, Comas D, Hammer MF, Kivisild T, Klitz W, Winkler CA, Labuda D, Bamshad M, Jorde LB, Tishkoff SA, Watkins WS, Metspalu M, Dryomov S, Sukernik R, Singh L, Thangaraj K, Pääbo S, Kelso J, Patterson N, Reich D. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. Nature. 2016 Oct 13;538(7624):201-206. doi: 10.1038/nature18964.
35. Malomane DK, Reimer C, Weigend S, Weigend A, Sharifi AR, Simianer H. Efficiency of different strategies to mitigate ascertainment bias when using SNP panels in diversity studies. BMC Genomics. 2018 Jan 5;19(1):22. doi: 10.1186/s12864-017-4416-9.
36. Martin SH, Amos W. Signatures of introgression across the allele frequency spectrum. Mol Biol Evol. 2021 Jan 23;38(2):716-726. doi: 10.1093/molbev/msaa239.
37. Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson K, Fernandes D, Novak M, Sirak K, Gamba C, Jones ER, Llamas B, Dryomov S, Pickrell J, Arsuaga JL, de Castro JM, Carbonell E, Gerritsen F, Khokhlov A, Kuznetsov P, Lozano M, Meller H, Mochalov O, Moiseyev V, Guerra MA, Roodenberg J, Vergès JM, Krause J, Cooper A, Alt KW, Brown D, Anthony D, Lalueza-Fox C, Haak W, Pinhasi R, Reich D. Genome-wide patterns of selection in 230 ancient Eurasians. Nature. 2015 Dec 24;528(7583):499-503. doi: 10.1038/nature16152.
38. McTavish EJ, Hillis DM. How do SNP ascertainment schemes and population demographics affect inferences about population history? BMC Genomics. 2015 Apr 3;16(1):266. doi: 10.1186/s12864-015-1469-5.
39. Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S, Schraiber JG, Jay F, Prüfer K, de Filippo C, Sudmant PH, Alkan C, Fu Q, Do R, Rohland N, Tandon A, Siebauer M, Green RE, Bryc K, Briggs AW, Stenzel U, Dabney J, Shendure J, Kitzman J, Hammer MF, Shunkov MV, Derevianko AP, Patterson N, Andrés AM, Eichler EE, Slatkin M, Reich D, Kelso J, Pääbo S. A high-coverage genome sequence from an archaic Denisovan individual. Science. 2012 Oct 12;338(6104):222-6. doi: 10.1126/science.1224344.
40. Nelson D, Kelleher J, Ragsdale AP, Moreau C, McVean G, Gravel S. Accounting for long-range correlations in genome-wide simulations of large cohorts. PLoS Genet. 2020 May 5;16(5):e1008619. doi: 10.1371/journal.pgen.1008619.

41. Nielsen R, Signorovitch J. Correcting for ascertainment biases when analyzing SNP data: applications to the estimation of linkage disequilibrium. Theor Popul Biol. 2003 May;63(3):245-55. doi: 10.1016/s0040-5809(03)00005-4.
42. Nielsen R. Population genetic analysis of ascertained SNP data. Hum Genomics. 2004 Mar;1(3):218-24. doi: 10.1186/1479-7364-1-3-218.
43. Nielsen R, Hubisz MJ, Clark AG. Reconstituting the frequency spectrum of ascertained single-nucleotide polymorphism data. Genetics. 2004 Dec;168(4):2373-82. doi: 10.1534/genetics.104.031039.
44. Olalde I, Posth C. Latest trends in archaeogenetic research of west Eurasians. Curr Opin Genet Dev. 2020 Jun;62:36-43. doi: 10.1016/j.gde.2020.05.021.
45. Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. Ancient admixture in human history. Genetics. 2012 Nov;192(3):1065-93. doi: 10.1534/genetics.112.145037.
46. Posth C, Wißing C, Kitagawa K, Pagani L, van Holstein L, Racimo F, Wehrberger K, Conard NJ, Kind CJ, Bocherens H, Krause J. Deeply divergent archaic mitochondrial genome provides lower time boundary for African gene flow into Neanderthals. Nat Commun. 2017 Jul 4;8:16046. doi: 10.1038/ncomms16046.
47. Pouyet F, Aeschbacher S, Thiéry A, Excoffier L. Background selection and biased gene conversion affect more than 95% of the human genome and bias demographic inferences. Elife. 2018 Aug 23;7:e36317. doi: 10.7554/eLife.36317.
48. Prendergast ME, Lipson M, Sawchuk EA, Olalde I, Ogola CA, Rohland N, Sirak KA, Adamski N, Bernardos R, Broomandkhoshbacht N, Callan K, Culleton BJ, Eccles L, Harper TK, Lawson AM, Mah M, Oppenheimer J, Stewardson K, Zalzala F, Ambrose SH, Ayodo G, Gates HL Jr, Gidna AO, Katongo M, Kwekason A, Mabulla AZP, Mudenda GS, Ndiema EK, Nelson C, Robertshaw P, Kennett DJ, Manthi FK, Reich D. Ancient DNA reveals a multistep spread of the first herders into sub-Saharan Africa. Science. 2019 Jul 5;365(6448):eaaw6275. doi: 10.1126/science.aaw6275.
49. Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, Heinze A, Renaud G, Sudmant PH, de Filippo C, Li H, Mallick S, Dannemann M, Fu Q, Kircher M, Kuhlwilm M, Lachmann M, Meyer M, Ongyerth M, Siebauer M, Theunert C, Tandon A, Moorjani P, Pickrell J, Mullikin JC, Vohr SH, Green RE, Hellmann I, Johnson PL, Blanche H, Cann H, Kitzman JO, Shendure J, Eichler EE, Lein ES, Bakken TE, Golovanova LV, Doronichev VB, Shunkov MV, Derevianko AP, Viola B, Slatkin M, Reich D, Kelso J, Pääbo S. The complete genome sequence of a Neanderthal from the Altai Mountains. Nature. 2014 Jan 2;505(7481):43-9. doi: 10.1038/nature12886.
50. Prüfer K, de Filippo C, Grote S, Mafessoni F, Korlević P, Hajdinjak M, Vernot B, Skov L, Hsieh P, Peyrégne S, Reher D, Hopfe C, Nagel S, Maricic T, Fu Q, Theunert C, Rogers R, Skoglund P, Chintalapati M, Dannemann M, Nelson BJ, Key FM, Rudan P, Kućan Ž, Gušić I, Golovanova LV, Doronichev VB, Patterson N, Reich D, Eichler EE, Slatkin M, Schierup MH, Andrés AM, Kelso J, Meyer M, Pääbo S. A high-coverage Neandertal genome from Vindija Cave in Croatia. Science. 2017 Nov 3;358(6363):655-658. doi: 10.1126/science.aao1887.
51. Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY, Viola B, Briggs AW, Stenzel U, Johnson PL, Maricic T, Good JM, Marques-Bonet T, Alkan C, Fu Q, Mallick S, Li H, Meyer M, Eichler EE, Stoneking M, Richards M, Talamo S, Shunkov MV, Derevianko AP, Hublin JJ, Kelso J, Slatkin M, Pääbo S. Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature. 2010 Dec 23;468(7327):1053-60. doi: 10.1038/nature09710.
52. Reich D, Patterson N, Campbell D, Tandon A, Mazieres S, Ray N, Parra MV, Rojas W, Duque C, Mesa N, García LF, Triana O, Blair S, Maestre A, Dib JC, Bravi CM, Bailliet G, Corach D, Hünemeier T, Bortolini MC, Salzano FM, Petzl-Erler ML, Acuña-Alonzo V, Aguilar-Salinas C, Canizales-Quinteros S, Tusié-Luna T, Riba L, Rodríguez-Cruz M, Lopez-Alarcón M, Coral-Vazquez R, Canto-Cetina T, Silva-Zolezzi I, Fernandez-Lopez JC, Contreras AV, Jimenez-Sanchez G, Gómez-Vázquez MJ, Molina J, Carracedo A, Salas A, Gallo C, Poletti G, Witonsky DB, Alkorta-Aranburu G, Sukernik RI, Osipova L, Fedorova SA, Vasquez R, Villena M, Moreau C, Barrantes R, Pauls D, Excoffier L,

Bedoya G, Rothhammer F, Dugoujon JM, Larrouy G, Klitz W, Labuda D, Kidd J, Kidd K, Di Rienzo A, Freimer NB, Price AL, Ruiz-Linares A. Reconstructing Native American population history. Nature. 2012 Aug 16;488(7411):370-4. doi: 10.1038/nature11258.

53. Rohland N, Mallick S, Mah M, Maier R, Patterson N, Reich D. Three assays for in-solution enrichment of ancient human DNA at more than a million SNPs. Genome Res. 2022 Nov-Dec;32(11-12):2068-2078. doi: 10.1101/gr.276728.122.
54. Scally A, Durbin R. Revising the human mutation rate: implications for understanding human evolution. Nat Rev Genet. 2012 Oct;13(10):745-53. doi: 10.1038/nrg3295.
55. Sirak KA, Fernandes DM, Lipson M, Mallick S, Mah M, Olalde I, Ringbauer H, Rohland N, Hadden CS, Harney É, Adamski N, Bernardos R, Broomandkhoshbacht N, Callan K, Ferry M, Lawson AM, Michel M, Oppenheimer J, Stewardson K, Zalzala F, Patterson N, Pinhasi R, Thompson JC, Van Gerven D, Reich D. Social stratification without genetic differentiation at the site of Kulubnarti in Christian Period Nubia. Nat Commun. 2021 Dec 14;12(1):7283. doi: 10.1038/s41467-021-27356-8.
56. Skoglund P, Thompson JC, Prendergast ME, Mittnik A, Sirak K, Hajdinjak M, Salie T, Rohland N, Mallick S, Peltzer A, Heinze A, Olalde I, Ferry M, Harney E, Michel M, Stewardson K, Cerezo-Román JI, Chiumia C, Crowther A, Gomani-Chindebvu E, Gidna AO, Grillo KM, Helenius IT, Hellenthal G, Helm R, Horton M, López S, Mabulla AZP, Parkington J, Shipton C, Thomas MG, Tibesasa R, Welling M, Hayes VM, Kennett DJ, Ramesar R, Meyer M, Pääbo S, Patterson N, Morris AG, Boivin N, Pinhasi R, Krause J, Reich D. Reconstructing Prehistoric African Population Structure. Cell. 2017 Sep 21;171(1):59-71.e21. doi: 10.1016/j.cell.2017.08.049.
57. Skoglund P, Mathieson I. Ancient Genomics of Modern Humans: The First Decade. Annu Rev Genomics Hum Genet. 2018 Aug 31;19:381-404. doi: 10.1146/annurev-genom-083117-021749.
58. Stoneking M, Arias L, Liu D, Oliveira S, Pugach I, Rodriguez JJRB. Genomic perspectives on human dispersals during the Holocene. Proc Natl Acad Sci U S A. 2023 Jan 24;120(4):e2209475119. doi: 10.1073/pnas.2209475119.
59. van de Loosdrecht M, Bouzouggar A, Humphrey L, Posth C, Barton N, Aximu-Petri A, Nickel B, Nagel S, Talbi EH, El Hajraoui MA, Amzazi S, Hublin JJ, Pääbo S, Schiffels S, Meyer M, Haak W, Jeong C, Krause J. Pleistocene North African genomes link Near Eastern and sub-Saharan African human populations. Science. 2018 May 4;360(6388):548-552. doi: 10.1126/science.aar8380.
60. Wang Y, Nielsen R. Estimating population divergence time and phylogeny from single-nucleotide polymorphisms data with outgroup ascertainment bias. Mol Ecol. 2012 Feb;21(4):974-86. doi: 10.1111/j.1365-294X.2011.05413.x.
61. Wang K, Goldstein S, Bleasdale M, Clist B, Bostoen K, Bakwa-Lufu P, Buck LT, Crowther A, Dème A, McIntosh RJ, Mercader J, Ogola C, Power RC, Sawchuk E, Robertshaw P, Wilmsen EN, Petraglia M, Ndiema E, Manthi FK, Krause J, Roberts P, Boivin N, Schiffels S. Ancient genomes reveal complex patterns of population movement, interaction, and replacement in sub-Saharan Africa. Sci Adv. 2020 Jun 12;6(24):eaaz0183. doi: 10.1126/sciadv.aaz0183.
62. Yaka R, Mapelli I, Kaptan D, Doğu A, Chyleński M, Erdal ÖD, Koptekin D, Vural KB, Bayliss A, Mazzucato C, Fer E, Çokoğlu SS, Lagerholm VK, Krzewińska M, Karamurat C, Gemici HC, Sevkar A, Dağtaş ND, Kılınç GM, Adams D, Munters AR, Sağlıcan E, Milella M, Schotsmans EMJ, Yurtman E, Çetin M, Yorulmaz S, Altınışık NE, Ghalichi A, Juras A, Bilgin CC, Günther T, Storå J, Jakobsson M, de Kleijn M, Mustafaoğlu G, Fairbairn A, Pearson J, Togan İ, Kayacan N, Marciniak A, Larsen CS, Hodder I, Atakuman Ç, Pilloud M, Sürer E, Gerritsen F, Özbal R, Baird D, Erdal YS, Duru G, Özbaşaran M, Haddow SD, Knüsel CJ, Götherström A, Özer F, Somel M. Variable kinship patterns in Neolithic Anatolia revealed by ancient genomes. Curr Biol. 2021 Jun 7;31(11):2455-2468.e18. doi: 10.1016/j.cub.2021.03.050.

Acknowledgements

The authors are grateful to Mark Lipson and Nick Patterson for discussions. P.F., U.I., and P.C. were supported by the Czech Ministry of Education, Youth and Sports (program ERC CZ, project no. LL2103). P.F. and P.C. were supported by the Czech Science Foundation (project no. 21-27624S). P.F. was also supported by a subsidy from the Russian federal budget (project No. 075-15-2019-1879 “From paleogenetics to cultural anthropology: a comprehensive interdisciplinary study of the traditions of the peoples of transboundary regions: migration, intercultural interaction and worldview”). This research was funded by NIH grant HG012287, by the Allen Discovery Center program, a Paul G. Allen Frontiers Group advised program of the Paul G. Allen Family Foundation, by John Templeton Foundation grant 61220, by private gifts from Jean-Francois Clin to D.R. and P.F., and by the Howard Hughes Medical Institute. Computational resources for this work were supplied by the projects "e-Infrastruktura CZ" (e-INFRA CZ LM2018140) and “IT4Innovations National Supercomputing Center – LM2015070” supported by the Ministry of Education, Youth and Sports of the Czech Republic.

Figures

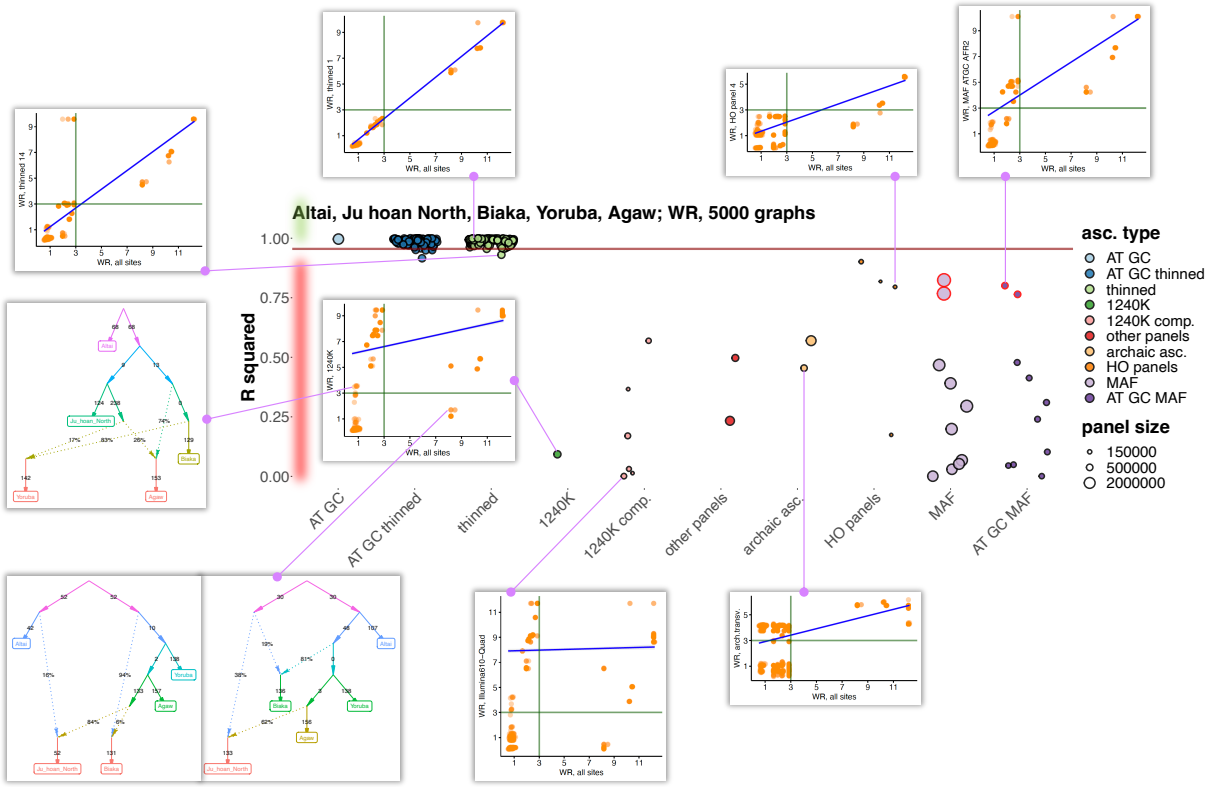
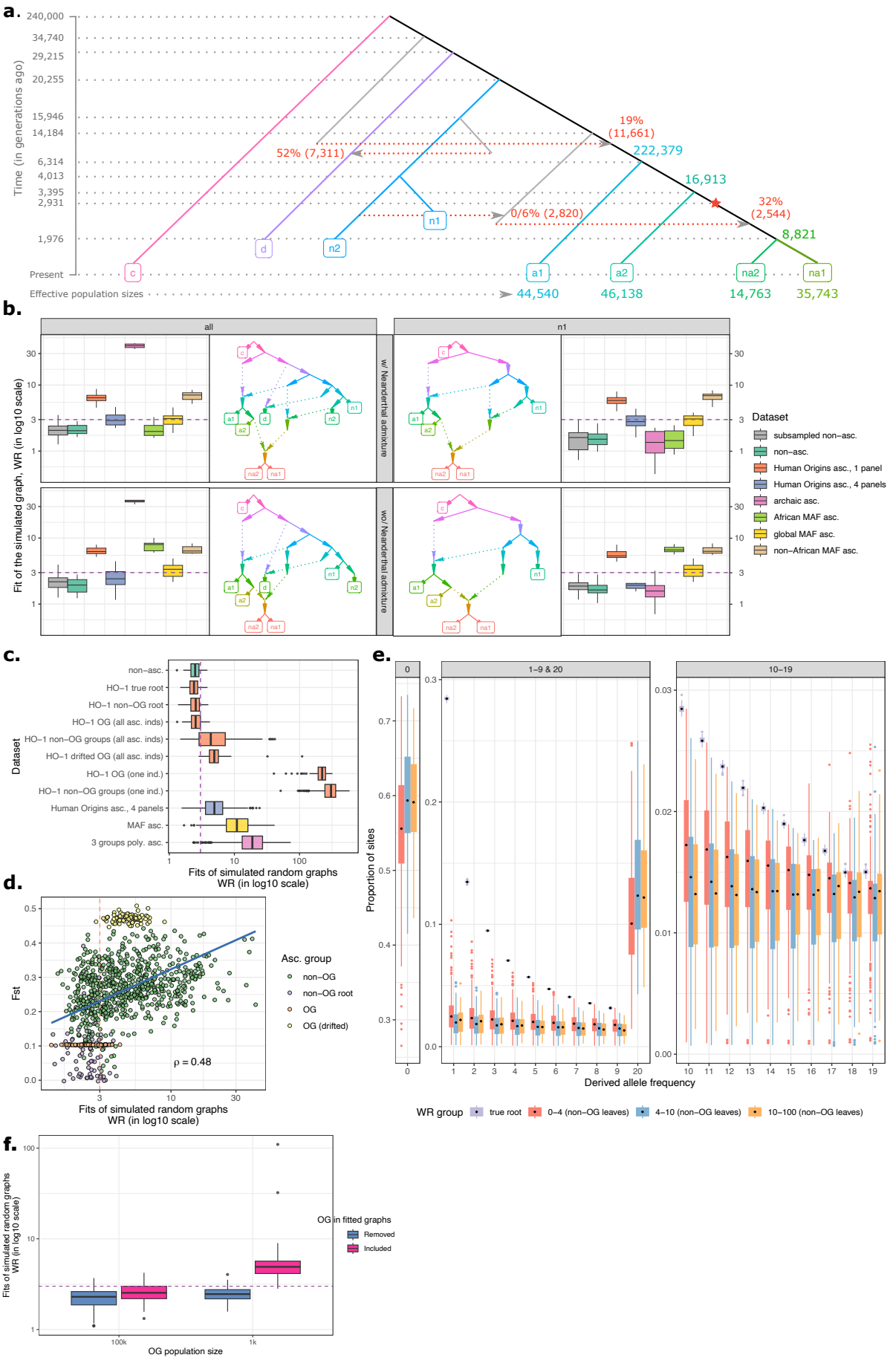


Fig. 1. The effect of ascertainment bias on admixture graph fits illustrated on a population combination "Altai Neanderthal, Ju|'hoan North, Biaka, Yoruba, Agaw". Five thousand best-fitting graphs (according to LL on all sites) of 32,745 possible graphs were selected, and correlation of WR was explored for graphs fitted on all sites and on ascertained sites. Results for ascertainment on variants common in Africans (either those having no detectable West Eurasian ancestry or all Africans in the SGDP dataset) are circled in red. Thirty eight site subsampling schemes were explored: 1) AT/GC mutation classes; 2) random thinning of the AT/GC dataset to the 1240K SNP count for a given combination of groups (no missing data allowed), results for 100 thinned replicates are shown; 3) random thinning of all sites to the 1240K SNP count, results for 100 thinned replicates are shown; 4) the 1240K enrichment panel; 5) major components of the 1240K panel: sites included in the Illumina 650Y and/or Human Origins SNP arrays, sites included exclusively in one of them, and remaining sites; 6) the 1000K and 2200K SNP panels; 7) restricting to sites polymorphic in a group composed of the three high-coverage archaic individuals (either all such sites or transversions only); 8) the largest Human Origins sub-panels (4, 5, 13) or their union (4+5); 9) restricting to common variants based on a global MAF threshold of 5% or on the same threshold in one of nine continental-scale groups; 10) the same procedure repeated on AT/GC sites. The size of the resulting SNP panels is coded by point size, and the ten broad ascertainment types are coded by color according to the legend. R^2 values of a linear trend for admixture graph WRs are plotted (WR for the large collections of admixture graphs were compared on all sites and under a particular ascertainment). The 2.5th WR percentile of all the thinned replicates combined, including those on all sites and AT/GC sites, is marked with the brown line. The area of the plot where ascertainments are considered biased according to this classifier is highlighted in red on the left. Scatterplots illustrating effects of selected ascertainment schemes on WR are shown beside the central plot and are connected to the respective data points (ascertainments) by magenta lines. Dots on these scatterplots correspond to

1266 distinct admixture graph topologies. Few examples of admixture graphs whose fits are affected by
1267 ascertainment are also shown beside the scatterplots.



1268

1269

Fig. 2. Exploring the influence of non-outgroup ascertainment on fits of admixture graphs on simulated data. Results are presented for two topologies (with or without the Neanderthal to non-African gene flow) and for eight types of SNP sets: 1) 10 sets of randomly selected variable sites matching the average size of the "Human Origins, one panel" set, 500K sites; 2) unascertained sites (on average 5.55M polymorphic sites without missing data); 3) Human Origins-like ascertainment, one panel based on the "a2" group (500K sites on average across simulation iterations); 4) Human Origins-like ascertainment, a union of four panels based on randomly selected individuals from four groups ("a1", "a2", "na1", and "na2", 1.34M sites on average); 5) archaic ascertainment (1.05M sites on average); 6) "African MAF ascertainment", that is removal of sites with MAF < 5% in the union of "a1" and "a2" groups (1.85M sites on average); 7) similar MAF ascertainment on the union of "a1", "a2", "na1", "na2" (1.62M sites on average); 8) similar MAF ascertainment on the union of "na1" and "na2" groups (1.48M sites on average). (**a**) The simulated topology, with dates of demographic events and sampling dates (in generations) shown on the y-axis or next to gene flows. The Neanderthal to non-African gene flow was simulated either at 0% (by omitting the "n2" to ghost AMH gene flow) or as shown in the figure. Effective population sizes of archaic groups are omitted for clarity. The out-of-Africa bottleneck is marked with a star. (**b**) Boxplots illustrating the effects of various ascertainment schemes on fits (WR) of the correct admixture graphs. The dashed line on the logarithmic scale marks a threshold often used in the literature for classifying models into fitting and non-fitting ones—3 standard errors—and the observation that common ascertainment schemes consistently produce much higher Z-scores than this threshold provides unambiguous evidence that ascertainment bias can profoundly compromise admixture graph fitting. The topologies fitted to the data are shown beside the boxplots. In the panels on the right, simple graphs including only one archaic lineage are fitted ("n1" used as an example, but very similar results were obtained for the "n2" and "d" groups). In the panels on the left, results for the full simulated model fitted to the data are shown. (**c**) Ascertainment bias was also explored across 80 simulated genetic histories in the form of random admixture graphs. WR of the correct admixture graph was used as a measure of bias. WRs for non-ascertained data and four ascertainment schemes are summarized with boxplots: 1) Human Origins-like ascertainment, one panel; 2) Human Origins-like ascertainment, four panels; 3) MAF-based ascertainment (restricting to common variants) in random sets of four populations; 4) ascertainment on sites polymorphic in random sets of three individuals (one individual sampled per population). Human Origins-like ascertainment on single individuals was performed on the true root or the root of non-outgroup populations, on non-outgroup populations, or on more or less drifted outgroups (having effective population sizes of 100,000 or 1,000 diploid individuals, respectively) co-modelled with the other populations (abbreviated as "OG"). Alternatively, the same individual that was used for ascertainment acted as the only representative of its group for model fitting. (**d**) correct admixture graphs under the Human Origins-like ascertainment (one panel) are guaranteed to be well-fitting (WR < ca. 4 SE) if F_{ST} between the whole population sample used for ascertainment vs. the sample at the root of the simulation is below 0.12. (**e**) Derived allele frequency spectra (derived allele count in a sample of 20 chromosomes vs. proportion of sites) across simulated root and non-outgroup populations grouped according to the level of ascertainment bias. The spectra were calculated for sites polymorphic in the root population sample of 20 chromosomes. Populations are binned by WR of the true graph fitted to sites heterozygous in a single individual randomly drawn from that population (single-panel Human Origins-like ascertainment). The boxplots summarize DAF across all simulated graphs. DAF bins are shown in three separate panels with different y-axis ranges: 0 derived alleles; 1 to 9 and 20 derived alleles; 10 to 19 derived alleles. (**f**) Illustration of the principle that outgroup ascertainment is expected to be unbiased only if an outgroup (abbreviated as "OG") is not co-analyzed with the other populations. Human Origins-like ascertainment (one panel) was performed on more or less drifted outgroups (having effective population sizes of 100,000 or 1,000 diploid individuals, respectively) that were then either included in the fitted true admixture graphs or removed from them. WRs of these graphs are summarized with boxplots on the y-axis. The dashed horizontal line corresponds to WR = 3 SE.

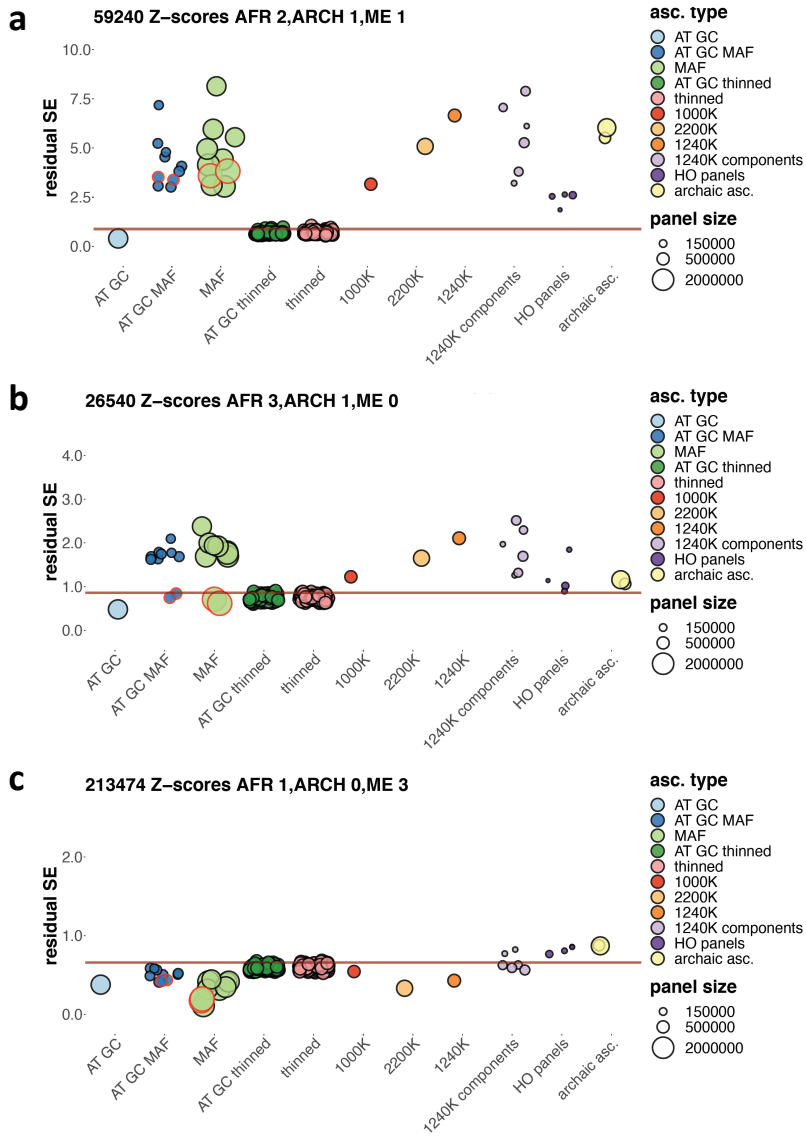


Fig. 4. Variance in f_4-statistic Z-scores resulting from ascertainment and random site subsampling expressed as standard deviations of residuals of a linear model (expressed in the same units as f_4-statistic Z-scores). Results are shown for three classes of f_4-statistics: f_4(African X, archaic; African Y, Mediterranean/Middle Eastern), f_4(African X, African Y; African Z, archaic) (expressed in another notation as African$_3$;archaic$_1$), and f_4(African, Med/ME X; Med/ME Y, Med/ME Z) (expressed in another notation as African$_1$;Mediterranean/Middle Eastern$_3$). Results for ascertainment on variants common in Africans (either those having no detectable West Eurasian ancestry according to Fan et al. (2019) or on all Africans in the SGDP dataset) are circled in red. Residual SE values for f_4-statistic Z-scores lying not far from 0 (absolute Z-scores on all sites < 15) are plotted. The 97.5% percentiles of all the thinned replicates combined, including those on all sites and AT/GC sites, are marked by the brown lines. Size of the SNP panels is coded by point size, and the broad ascertainment types are coded by color according to the legend. Thirty eight site subsampling schemes were explored: 1) AT/GC mutation classes; 2) AT/GC mutation classes and restricting to common variants based on a global MAF threshold of 5% or on the same threshold applied to one of nine continental-scale groups; 3) the same procedure repeated on all sites; 4) random thinning of the AT/GC dataset to the 1240K SNP count for a given combination of groups (no missing data allowed), results for 100 thinning replicates are shown; 5) random thinning of all sites to the 1240K SNP count, results for 100 thinning replicates are shown; 6) the 1000K, 1240K, and 2200K enrichment panels (2200K = 1240K +

1000K); 7) major components of the 1240K panel: sites included in the Illumina 650Y and/or Human Origins SNP arrays, sites included exclusively in one of them, and remaining sites; 8) the largest Human Origins sub-panels (4, 5, 13) or their union (4+5); 9) restricting to sites polymorphic in a group composed of the three high-coverage archaic individuals (either all such sites or transversions only).

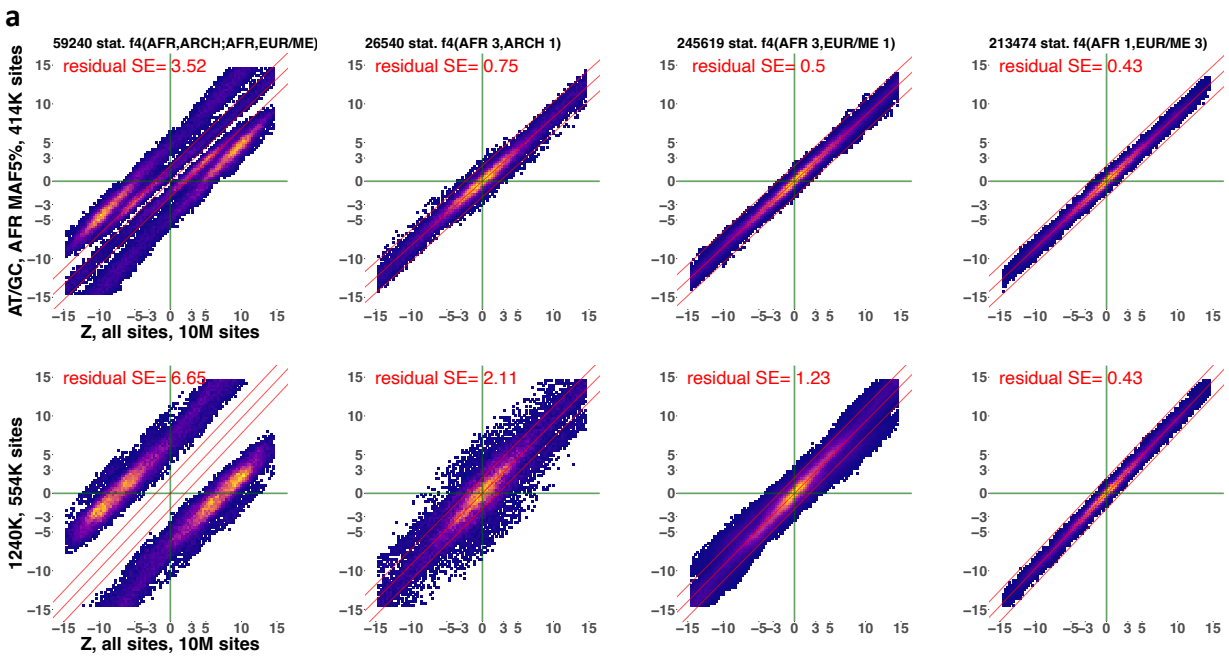


Fig. 5. Scatterplots illustrating the effects of two ascertainment schemes on Z-scores of f_4-statistics of four classes including African and/or archaic and/or Mediterranean/Middle Eastern groups. Only statistics of the form f_4(African X, archaic; African Y, non-African) were considered in the class African$_2$;archaic$_1$;Med/ME$_1$. The f_4-statistic classes were selected to represent severe ascertainment bias (leftmost panels), moderate level of bias (two middle panels) and no bias (rightmost panels). The ascertainments selected are 1240K (the most widely used SNP enrichment panel) and the new "pan-African" scheme proposed in this study to mitigate ascertainment bias for nearly all f_4-statistic classes. For results on other f_4-statistic classes see Suppl. Fig. 14, and results for a wider range of ascertainment schemes are summarized in Suppl. Figs. 12 and 13. Class labels and numbers of statistics plotted are shown above the panels. Instead of individual points, heatmaps illustrating point density are shown. Z-scores on all sites (10 million sites, as indicated on the x-axes) are compared to Z-scores on ascertained datasets on the y-axes. Ascertainment types and site counts are shown on the y-axes. All plots include only statistics with absolute Z-scores below 15 on all sites. A linear model fitted to the data and lines representing ± 2 SE are shown in red. Residual SE values for those linear trends are shown in each plot in red.

percentage of models rejected on asc. but accepted on all sites				*arch 2, afr 3*	*arch 1, afr 4*	*arch 1, afr 3, nafr 1*	*arch 1, afr 2, nafr 2*	*afr 5*	*afr 4, nafr 1*	*afr 4, nafr 1*	*afr 3, nafr 2*	*afr 1, nafr 4*	*nafr 5*	*nafr 5*	*nafr 5*		number of biased pop. sets, both metrics
ascertainment type	**further details on the ascertainment**	**min. size of the SNP panel***	**max. size of the SNP panel***	**Denisovan, Altai, Yoruba, Dinka, Bulala**	**Denisovan, Khomani San, Mbuti, Dinka, Mursi**	**Altai, Ju hoan North, Biaka, Yoruba, Agaw**	**Altai, Ju hoan North, Luhya, Palestinian, Spanish**	**Mbuti, Baka, Laka, Fulani, Bantu Tswana**	**Khomani San, Bakola, Igbo, Mursi, Aari**	**Bedzan, Cameroon SMA, Esan, Mozabite, Masai**	**Mbuti, Biaka, Ngumba, LBK, Iranian**	**Luo, Bedouin B, Jordanian, Abkhasian, Sardinian**	**Australian, Quechua, Mayan, Lezgin, French**	**Papuan, Chipewyan, Eskimo Naukan, Finnish, Sardinian**	**Karitiana, Cree, Eskimo Sireniki, Hungarian, Icelandic**	**number of biased pop. sets**	
A<>T and G<>C mutations		805,042	1,757,840	0.000%	0.009%	0.000%	0.003%	0.000%	0.000%	0.000%	0.311%	0.000%	0.000%	0.000%	0.000%	0	0
1240K panel		501,429	663,239	0.015%	**1.017%**	**7.974%**	**5.528%**	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	3	4
components of the 1240K panel	Illumina610-Quad sites	256,277	304,292	0.012%	0.015%	**7.992%**	**5.528%**	0.000%	0.000%	0.000%	0.000%	0.021%	0.000%	0.000%	0.000%	2	3
	sites exclusive to Illumina610-Quad	183,680	216,478	0.000%	0.012%	**7.974%**	**5.528%**	0.000%	0.000%	0.000%	0.000%	**0.061%**	0.000%	0.000%	0.000%	3	4
	sites included in both Illumina610-Quad and HumanOrigins	72,597	87,814	0.018%	0.000%	**7.971%**	**5.528%**	0.000%	0.000%	0.003%	0.000%	0.000%	0.000%	0.000%	0.000%	2	6
	HumanOrigins sites	244,922	354,460	0.000%	0.000%	**7.952%**	0.000%	0.000%	0.000%	0.000%	0.009%	0.000%	0.000%	0.000%	0.000%	1	3
	sites exclusive to HumanOrigins	171,249	266,646	0.003%	0.000%	**6.883%**	0.000%	0.000%	0.000%	0.000%	0.003%	0.000%	0.000%	0.000%	0.000%	1	2
	1240K, other sites	67,096	92,301	0.000%	0.003%	**7.952%**	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	1	4
the largest panels included in the HumanOrigins array	panel 13 based on a San individual and Denisovan	67,557	89,655	0.003%	0.168%	**6.993%**	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	1	6
	panel 4 based on a San individual	52,862	94,493	0.000%	0.000%	0.000%	0.000%	0.003%	0.000%	0.000%	0.012%	0.000%	0.000%	0.000%	0.000%	0	4
	panel 5 based on a Yoruba individual	44,674	73,180	**0.724%**	0.000%	**5.250%**	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	2	8
	panels 4 and 5	46,701	157,126	0.015%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.006%	0.000%	0.000%	0.000%	0.000%	0	1
other enrichment panels	1000K: transversions in 2 Yoruba ind. and in Altai Neand.	364,079	590,775	0.000%	0.000%	**6.890%**	0.000%	0.000%	0.000%	0.000%	0.012%	0.000%	0.000%	0.000%	0.000%	1	2
	2200K = 1000K + 1240K	814,915	1,190,758	0.015%	0.000%	**7.952%**	0.000%	0.000%	0.000%	0.000%	0.012%	0.000%	0.000%	0.000%	0.000%	1	1
archaic ascertainment	transitions and transversions	525,014	1,555,781	**10.484%**	0.000%	**6.899%**	0.009%	0.000%	0.000%	0.003%	0.021%	0.000%	0.000%	0.000%	0.000%	2	8
	transversions	165,249	484,675	**10.484%**	0.000%	**6.880%**	0.003%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	2	8
MAF ascertainment	global, >5% MAF	2,129,201	2,511,335	0.000%	0.000%	**7.952%**	0.000%	0.000%	0.000%	0.000%	0.003%	0.000%	0.000%	0.000%	0.000%	1	2
	>5% MAF in Africans unadmixed with non-Africans	2,045,769	3,231,875	**1.032%**	0.003%	**6.932%**	0.000%	0.000%	0.000%	0.015%	0.006%	0.000%	0.000%	0.000%	0.000%	2	2
	>5% MAF in all Africans	2,109,808	3,120,326	0.037%	0.003%	**6.929%**	0.000%	0.000%	0.000%	0.000%	0.009%	0.000%	0.000%	0.000%	0.000%	1	1
	>5% MAF in Native Americans	1,513,207	1,764,715	0.015%	0.000%	**7.992%**	**5.195%**	0.000%	0.000%	0.003%	0.000%	0.000%	0.000%	0.000%	0.000%	2	3
	>5% MAF in Central Asians and Siberians	1,843,262	2,150,675	0.018%	0.000%	**7.974%**	0.000%	0.000%	0.000%	0.244%	0.006%	0.000%	0.000%	0.000%	0.000%	1	2
	>5% MAF in East Asians	1,723,831	2,020,860	0.000%	0.000%	**7.992%**	**5.528%**	0.000%	0.000%	0.003%	0.003%	0.000%	0.000%	0.000%	0.000%	2	3
	>5% MAF in Europeans	1,885,336	2,192,571	0.000%	0.000%	**7.952%**	0.000%	0.000%	0.000%	0.000%	0.003%	0.000%	0.000%	0.000%	0.000%	1	2
	>5% MAF in Middle Eastern groups	2,018,884	2,306,319	0.006%	0.000%	**7.952%**	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	1	1
	>5% MAF in Papuans and Aboriginal Australians	1,515,022	1,791,390	**1.032%**	0.000%	**7.995%**	**5.528%**	0.000%	0.000%	0.244%	0.311%	0.000%	0.000%	0.000%	0.000%	3	4
	>5% MAF in South Asians	1,908,459	2,235,024	0.000%	0.000%	**7.968%**	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	1	2
AT/GC mutation types + MAF ascertainment	global, >>5% MAF	323,296	378,287	0.000%	0.000%	**7.952%**	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	1	2
	>5% MAF in Africans unadmixed with non-Africans	309,172	486,906	0.003%	0.000%	**6.929%**	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	1	1
	>5% MAF in all Africans	319,053	470,070	0.003%	0.006%	**6.929%**	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	1	1
	>5% MAF in Native Americans	229,939	266,113	0.000%	0.000%	**7.974%**	**6.114%**	0.000%	0.000%	0.049%	0.318%	0.000%	0.000%	0.000%	0.000%	2	3
	>5% MAF in Central Asians and Siberians	280,103	324,245	0.000%	0.000%	**7.974%**	0.000%	0.000%	0.000%	**0.253%**	0.049%	0.000%	0.000%	0.000%	0.000%	2	3
	>5% MAF in East Asians	261,857	304,567	0.000%	0.000%	**7.992%**	0.000%	0.000%	0.000%	0.003%	0.324%	0.000%	0.000%	0.000%	0.000%	1	2
	>5% MAF in Europeans	285,723	330,244	0.000%	0.000%	**7.952%**	0.000%	0.000%	0.000%	0.000%	0.000%	0.003%	0.000%	0.000%	0.000%	1	3
	>5% MAF in Middle Eastern groups	306,450	347,536	0.003%	0.000%	**7.952%**	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	1	3
	>5% MAF in Papuans and Aboriginal Australians	230,124	272,093	**1.029%**	0.000%	**8.001%**	**5.528%**	0.000%	0.000%	0.244%	0.324%	0.000%	0.000%	0.000%	0.000%	3	4
	>5% MAF in South Asians	289,739	336,996	0.000%	0.000%	**6.948%**	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	1	3
* the SNP counts correspond to sites polymorphic in larger collections of groups from which the analyzed population quintuplets were taken, see Suppl. Table X.		**number of biased asc. =>**		6	1	33	9	0	0	1	0	1	0	0	0		
		number of biased asc., both metrics =>		6	2	61	9	0	6	10	0	5	0	9	3		

Table 1. Performance of ascertainment schemes explored across 12 population quintuplets and assessed as the fraction of all possible admixture graph topologies that are rejected under ascertainment (WR > 3 SE) but accepted on all sites (WR < 3 SE). We also applied the binary classifier to determine if the ascertainment produces unbiased or biased results (the latter highlighted in bold and underlined text). The numbers of population quintuplets or ascertainment schemes affected by bias (according to the fraction of topologies that are rejected under ascertainment but accepted on all sites, or according to this metric and the fraction of topologies that are accepted under ascertainment but rejected on all sites) are shown in the two rightmost

columns and in two bottom rows, respectively. The composition of the population sets is shown above the table in an abbreviated way: *arch*, archaic humans, followed by the number of archaic groups; *afr*, Africans; *nafr*, non-Africans or Africans with substantial non-African admixture (Fan et al. 2019).