

<sup>1</sup> Assessing the relationship of ancient and modern populations

<sup>2</sup> Joshua G. Schraiber

<sup>3</sup> Started on October 22, 2016. Compiled on July 21, 2017

<sup>4</sup> **Abstract**

<sup>5</sup> Genetic material sequenced from ancient samples is revolutionizing our understand-  
<sup>6</sup> ing of the recent evolutionary past. However, ancient DNA is often degraded, resulting  
<sup>7</sup> in low coverage, error-prone sequencing. Several solutions exist to this problem, rang-  
<sup>8</sup> ing from simple approach such as selecting a read at random for each site to more  
<sup>9</sup> complicated approaches involving genotype likelihoods. In this work, we present a  
<sup>10</sup> novel method for assessing the relationship of an ancient sample with a modern pop-  
<sup>11</sup> ulation while accounting for sequencing error by analyzing raw read from multiple  
<sup>12</sup> ancient individuals simultaneously. We show that when analyzing SNP data, it is bet-  
<sup>13</sup> ter to sequencing more ancient samples to low coverage: two samples sequenced to  
<sup>14</sup> 0.5x coverage provide better resolution than a single sample sequenced to 2x coverage.  
<sup>15</sup> We also examined the power to detect whether an ancient sample is directly ancestral  
<sup>16</sup> to a modern population, finding that with even a few high coverage individuals, even  
<sup>17</sup> ancient samples that are very slightly diverged from the modern population can be de-  
<sup>18</sup> tected with ease. When we applied our approach to European samples, we found that  
<sup>19</sup> no ancient samples represent direct ancestors of modern Europeans. We also found  
<sup>20</sup> that, as shown previously, the most ancient Europeans appear to have had the small-  
<sup>21</sup> est effective population sizes, indicating a role for agriculture in modern population  
<sup>22</sup> growth.

## 23 1 Introduction

24 Ancient DNA (aDNA) is now ubiquitous in population genetics. Advances in DNA  
25 isolation [Dabney et al., 2013], library preparation [Meyer et al., 2012], bone sampling  
26 [Pinhasi et al., 2015], and sequence capture Haak et al. [2015] make it possible to obtain  
27 genome-wide data from hundreds of samples [Haak et al., 2015, Mathieson et al., 2015,  
28 Allentoft et al., 2015, Fu et al., 2016]. Analysis of these data can provide new insight  
29 into recent evolutionary processes which leave faint signatures in modern genomes,  
30 including natural selection [Schraiber et al., 2016, Jewett et al., 2016] and population  
31 replacement [Sjödin et al., 2014, Lazaridis et al., 2014].

32 One of the most powerful uses of ancient DNA is to assess the continuity of an-  
33 cient and modern populations. In many cases, it is unclear whether populations that  
34 occupied an area in the past are the direct ancestors of the current inhabitants of that  
35 area. However, this can be next to impossible to assess using only modern genomes.  
36 Questions of population continuity and replacement have particular relevance for the  
37 spread of cultures and technology in humans [Lazaridis et al., 2016]. For instance, re-  
38 cent work showed that modern South Americans are descended from people associated  
39 with Clovis culture that inhabited North America over 10,000 years ago, providing further  
40 evidence toward our understanding of the peopling of the Americas [Rasmussen et al.,  
41 2014].

42 Despite its utility in addressing difficult-to-answer questions in evolutionary biology,  
43 aDNA also has several limitations. Most strikingly, DNA decays rapidly following the  
44 death of an organism, resulting in highly fragmented, degraded starting material when  
45 sequencing [Sawyer et al., 2012]. Thus, ancient data is frequently sequenced to low  
46 coverage and has a significantly higher base-calling error rate than modern samples.  
47 When working with diploid data, as in aDNA extracted from plants and animals, the  
48 low coverage prevents genotypes from being called with confidence.

49 Several strategies are commonly used to address the low-coverage data. One of the  
50 most common approaches is to sample a random read from each covered site and use

51 that as a haploid genotype call [Haak et al., 2015, Mathieson et al., 2015, Allentoft  
52 et al., 2015, Fu et al., 2016, Lazaridis et al., 2016]. Many common approaches to the  
53 analyses of ancient DNA, such as the usage of F-statistics [Patterson et al., 2012], are  
54 designed with this kind of dataset in mind. As shown by Peter [2016], F-statistics can  
55 be interpreted as linear combinations of simpler summary statistics and can often be  
56 understood in terms of testing a tree-like structure relating populations. Nonetheless,  
57 despite the simplicity and appeal of this approach, it has several drawbacks. Primarily,  
58 it throws away reads from sites that are covered more than once, resulting in a potential  
59 loss of information from expensive, difficult-to-acquire data. These approach are also  
60 strongly impacted by sequencing error and contamination.

61 On the other hand, several approaches exist to either work with genotype likelihoods  
62 or the raw read data. Genotype likelihoods are the probabilities of the read data at a  
63 site given each of the three possible diploid genotypes at that site. They can be used  
64 in calculation of population genetic statistics or likelihood functions to average over  
65 uncertainty in the genotype [Korneliussen et al., 2014]. However, many such approaches  
66 assume that genotype likelihoods are fixed by the SNP calling algorithm. However, with  
67 low coverage data, an increase in accuracy is expected if genotype likelihoods are co-  
68 estimated with other parameters of interest, due to the covariation between processes  
69 that influence read quality and genetic diversity, such as contamination.

70 A recent method that coestimates demographic parameters along with error and  
71 contamination rates by using genotype likelihoods showed that there can be significant  
72 power to assess the relationship of a single ancient sample to a modern population  
73 [Racimo et al., 2016]. Nonetheless, they found that for very low coverage data, inferences  
74 were not reliable. Thus, they were unable to apply their method to the large  
75 number of extremely low coverage ( $< 1x$ ) genomes that are available. Moreover, they  
76 were unable to explore the tradeoffs that come with a limited budget: can we learn  
77 more by sequencing fewer individuals to high coverage, or more individuals at lower  
78 coverage?

79 Here, we develop a novel maximum likelihood approach for analyzing low coverage

80 ancient DNA in relation to a modern population. We work directly with raw read data  
81 and explicitly model base-calling errors. Crucially, our approach incorporates data  
82 from multiple individuals that belong to the same ancient population, which we show  
83 substantially increases power and reduces error in parameter estimates. We then apply  
84 our new methodology to ancient human data, and show that we can perform accurate  
85 demographic inference even from very low coverage samples by analyzing them jointly.

## 86 2 Methods

### 87 2.1 Sampling alleles in ancient populations

88 We assume a scenario in which allele frequencies are known with high accuracy in a  
89 modern population. Suppose that an allele is known to be at frequency  $x \in (0, 1)$  in  
90 the modern population, and we wish to compute the probability of obtaining  $k$  copies  
91 of that allele in a sample of  $n$  ( $0 \leq k \leq n$ ) chromosomes from an ancient population.  
92 Conditioning on the frequency of the allele in the modern population minimizes the  
93 impact of ascertainment, and allows this approach to be used for SNP capture data.

To calculate the sampling probability, we assume a simple demographic model in which the ancient individual belongs to a population that split off from the modern population  $\tau_1$  generations ago, and subsequently existed as an isolated population for  $\tau_2$  generations. Further, we assume that the modern population has effective size  $N_e^{(1)}$  and that the ancient population has effective size  $N_e^{(2)}$ , and measure time in diffusion units,  $t_i = \tau_i / (2N_e^{(i)})$ . If we know the conditional probability that an allele is at frequency  $y$  in the ancient sample, given that it is at frequency  $x$ , denoted  $f(y; x, t_1, t_2)$ , then the

sampling probability is simply an integral,

$$\begin{aligned}
P_{n,k}(x) &= \int_0^1 \binom{n}{k} y^k (1-y)^{n-k} f(y; x, t_1, t_2) dy \\
&= \binom{n}{k} \mathbb{E}_x (Y^k (1-Y)^{n-k}; t_1, t_2) \\
&\equiv \binom{n}{k} p_{n,k}(t_1, t_2)
\end{aligned} \tag{1}$$

94 Thus, we must compute the binomial moments of the allele frequency distribution in  
95 the ancient population. In the Appendix, we show that this can be computed using  
96 matrix exponentiation,

$$p_{n,k}(t_1, t_2) = \left( e^{Qt_2} e^{Q^\downarrow t_1} \mathbf{h}_n \right)_i, \tag{2}$$

97 where  $(\mathbf{v})_i$  indicates the  $i$ th element of the vector  $\mathbf{v}$ ,  $\mathbf{h}_n = ((1-x)^n, x(1-x)^{n-1}, \dots, x^n)^T$   
98 and  $Q$  and  $Q^\downarrow$  are the sparse matrices

$$Q_{ij} = \begin{cases} \frac{1}{2}i(i-1) & \text{if } j = i-1 \\ -i(n-i) & \text{if } j = i \\ \frac{1}{2}(n-i)(n-i-1) & \text{if } j = i+1 \\ 0 & \text{else} \end{cases}$$

99 and

$$Q_{ij}^\downarrow = \begin{cases} \frac{1}{2}i(i-1) & \text{if } j = i-1 \\ -i(n-i+1) & \text{if } j = i \\ \frac{1}{2}(n-i+1)(n-i) & \text{if } j = i+1 \\ 0 & \text{else.} \end{cases}$$

100 This result has an interesting interpretation: the matrix  $Q^\downarrow$  can be thought of as  
101 evolving the allele frequencies back in time from the modern population to the common  
102 ancestor of the ancient and modern populations, while  $Q$  evolves the allele frequencies  
103 forward in time from the common ancestor to the ancient population (Fig 1).

104

[Figure 1 about here.]

105

Because of the fragmentation and degradation of DNA that is inherent in obtaining sequence data from ancient individuals, it is difficult to obtain the high coverage data necessary to make high quality genotype calls from ancient individuals. To address this, we instead work directly with raw read data, and average over all the possible genotypes weighted by their probability of producing the data. Specifically, we follow Nielsen et al. [2012] in modeling the probability of the read data in the ancient population, given the allele frequency at site  $l$  as

$$\mathbb{P}(R_l|k) = \sum_{g_{1,l}=0}^2 \dots \sum_{g_{n,l}=0}^2 \mathbb{I}\left(\sum_{i=1}^m g_{i,l} = k\right) \prod_{i=1}^n \binom{2}{g_{i,l}} \mathbb{P}(R_{i,l}|g_{i,l}),$$

112

where  $R_{i,l} = (a_{i,l}, d_{i,l})$  are the counts of ancestral and derived reads in individual  $i$  at site  $l$ ,  $g_{i,l} \in \{0, 1, 2\}$  indicates the possible genotype of individual  $i$  at site  $l$  (i.e. 0 = homozygous ancestral, 1 = heterozygous, 2 = homozygous derived), and  $\mathbb{P}(R_{i,l}|g_{i,l})$  is the probability of the read data at site  $l$  for individual  $i$ , assuming that the individual truly has genotype  $g_{i,l}$ . We use a binomial sampling with error model, in which the probability that a truly derived site appears ancestral (and vice versa) is given by  $\epsilon$ . We emphasize that the parameter  $\epsilon$  will capture both sequencing error as well as post-mortem damage (c.f. ? who found that adding an additional parameter to specifically model post-mortem damage does not improve inferences). Thus,

$$\mathbb{P}(R|g) = \binom{a+d}{d} p_g^d (1-p_g)^a$$

with

$$\begin{aligned} p_0 &= \epsilon \\ p_1 &= \frac{1}{2} \\ p_2 &= 1 - \epsilon \end{aligned}$$

121 .

122 Combining these two aspects together by summing over possible allele frequencies  
123 weighted by their probabilities, we obtain our likelihood of the ancient data,

$$L(D) = \prod_{l=1}^L \sum_{k=0}^n \mathbb{P}(R_l|k) p_{n,k}(x_l). \quad (3)$$

124

### 3 Results

125

#### 3.1 Impact of coverage and number of samples on inferences

126

To explore the tradeoff of sequencing more individuals at lower depth compared to fewer  
127 individuals at higher coverage, we performed simulations using `msprime` [Kelleher et al.,  
128 2016] combined with custom scripts to simulate base calling error and low coverage  
129 data. First, we examined the impact of coverage and number of samples on the ability  
130 to recover the drift times in the modern and the ancient populations. Figure 2 shows  
131 results for data simulated with  $t_1 = 0.02$  and  $t_2 = 0.05$ , corresponding to an ancient  
132 individual who died 300 generations ago from population of effective size 1000. The  
133 populations split 400 generations ago, and the modern population has an effective  
134 size of 10000. We simulated approximately 180000 SNPs by simulating 100000 500  
135 base pair fragments. Inferences of  $t_1$  can be relatively accurate even with only one  
136 low coverage ancient sample (Figure 2A). However, inferences of  $t_2$  benefit much more  
137 from increasing the number of ancient samples, as opposed to coverage (Figure 2B). In

138 particular, two individuals sequenced to 0.5x coverage have a much lower error than a  
139 single individual sequenced to 2x coverage. To explore this effect further, we derived the  
140 sampling probability of alleles covered by exactly one sequencing read (see Appendix).  
141 We found that sites covered only once have no information about  $t_2$ , suggesting that  
142 evidence of heterozygosity is very important for inferences about  $t_2$ .

143 [Figure 2 about here.]

144 We next examined the impact of coverage and sampling on the power to reject  
145 the hypothesis that the ancient individuals came from a population that is directly  
146 ancestral to the modern population. We analyzed both low coverage (0.5x) and higher  
147 coverage (4x) datasets consisting of 1 (for both low and high coverage samples) or 5  
148 individuals (only for low coverage). We simulated data with parameters identical to  
149 the previous experiment, except we now examined the impact of varying the age of  
150 the ancient sample from 0 generations ago through to the split time with the modern  
151 population. We then performed a likelihood ratio test comparing the null model of con-  
152 tinuity, in which  $t_2 = 0$ , to a model in which the ancient population is not continuous.  
153 Figure 3 shows the power of the likelihood ratio test. For a single individual sequenced  
154 to low coverage, we see that the test only has power for very recently sampled ancient  
155 individuals (i.e. samples that are highly diverged from the modern population). How-  
156 ever, the power increases dramatically as the number of individuals or the coverage per  
157 individual is increased; sequencing 5 individuals to 0.5x coverage results in essentially  
158 perfect power to reject continuity. Nonetheless, for samples that are very close to the  
159 divergence time, it will be difficult to determine if they are ancestral to the modern  
160 population or not, because differentiation is incomplete.

161 [Figure 3 about here.]

162        **3.2 Impact of admixture**

163        We examined two possible violations of the model to assess their impact on inference.  
164        In many situations, there may have been secondary contact between the population  
165        from which the ancient sample is derived and the modern population used as a refer-  
166        ence. We performed simulations of this situation by modifying the previous simulations  
167        to include subsequent admixture from the ancient population to the modern popula-  
168        tion 200 generations ago (NB: this admixture occurred *more recently* than the ancient  
169        sample). In Figure 4, we show the results for admixture proportions ranging from 0  
170        to 50%. Counterintuitively, estimates of  $t_1$  initially *decrease* before again increasing.  
171        This is likely a result of the increased heterozygosity caused by admixture, which acts  
172        to artificially inflate the effective size of the modern population, and thus decrease  $t_1$ .  
173        As expected,  $t_2$  is estimated to be smaller the more admixture there is; indeed, for an  
174        admixture rate of 100%, the modern and ancient samples are continuous. The impact  
175        on  $t_2$  appears to be linear, and is well approximated by  $(1 - f)t_2$  if the admixture  
176        fraction is  $f$ .

177        [Figure 4 about here.]

178        In other situations, there may be admixture from an unsampled “ghost” population  
179        into the modern population. If the ghost admixture is of a high enough proportion, it  
180        is likely to cause a sample that is in fact a member of a directly ancestral population to  
181        not appear to be ancestral. We explored this situation by augmenting our simulations  
182        in which the ancient sample is continuous with an outgroup population diverged from  
183        the modern population 0.04 time units ago (corresponding to 800 generations ago)  
184        and contributed genes to the modern population 0.01 time units ago (corresponding to  
185        200 generations ago). We then assessed the impact on rejecting continuity using the  
186        likelihood ratio test (Figure 5). As expected, we see that low-power sampling strategies  
187        (such as a single individual sequenced to low coverage) are relatively unimpacted by  
188        ghost admixture. However, even relatively powerful sampling strategies can be robust  
189        to ghost admixture up to approximately 10%.

190

[Figure 5 about here.]

191

### 3.3 Application to ancient humans

192

We applied our approach to ancient human data from Mathieson et al. [2015], which is primarily derived from a SNP capture approach that targeted 1.2 million SNPs. Based on sampling location and associated archeological materials, the individuals were grouped into *a priori* panels, which we used to specify population membership when analyzing individuals together. We analyzed all samples for their relationship to the CEU individuals from the 1000 Genomes Project [Consortium, 2015]. Based on our results that suggested that extremely low coverage samples would yield unreliable estimates, we excluded panels that are composed of only a single individual sequenced to less than 2x coverage.

201

We computed maximum likelihood estimates of  $t_1$  and  $t_2$  for individuals as grouped into populations (Figure 6A; Table 1). We observe that  $t_2$  is significantly greater than 0 for all populations. Thus, none of these populations are consistent with directly making up a large proportion of the ancestry of modern CEU individuals. Strikingly, we see that  $t_2 \gg t_1$ , despite the fact that the ancient samples must have existed for fewer generations since the population split than the modern samples. This suggests that all of the ancient populations are characterized by extremely small effective population sizes.

209

[Table 1 about here.]

210

[Figure 6 about here.]

211

We further explored the relationship between the dates of the ancient samples and the parameters of the model by plotting  $t_1$  and  $t_2$  against the mean sample date of all samples in that population (Figure 6B, C). We expected to find that  $t_1$  correlated with sample age, under the assumption that samples were members of relatively short-lived populations that diverged from the “main-stem” of CEU ancestry. Instead, we

see no correlation between  $t_1$  and sample time, suggesting that the relationship of these populations to the CEU is complicated and not summarized well by the age of the samples. On the other hand, we see a strong positive correlation between  $t_2$  and sampling time ( $p < 1 \times 10^{-4}$ ). Because  $t_2$  is a compound parameter, it is difficult to directly interpret this relationship. However, it is consistent with the most ancient samples belonging to populations with the smallest effective sizes, consistent with previous observations [Skoglund et al., 2014].

Finally, we examined the impact of grouping individuals into populations in real data. We see that estimates of  $t_1$  for low coverage samples are typically lower when analyzed individually than when pooled with other individuals of the same panel (Figure 7A), suggesting a slightly downward bias in estimating  $t_1$  for low coverage samples. On the other hand, there is substantial bias toward overestimating  $t_2$  when analyzing samples individually, particularly for very low coverage samples (Figure 7B). This again shows that for estimates that rely on heterozygosity in ancient populations, pooling many low coverage individuals can significantly improve estimates.

[Figure 7 about here.]

## 4 Discussion

Ancient DNA (aDNA) presents unique opportunities to enhance our understanding of demography and selection in recent history. However, it also comes equipped with several challenges, due to postmortem DNA damage [Sawyer et al., 2012]. Several strategies have been developed to deal with the low quality of aDNA data, from relatively simple options like sampling a read at random at every site [Green et al., 2010] to more complicated methods making use of genotype likelihoods [Racimo et al., 2016]. Here, we presented a novel maximum likelihood approach for making inferences about how ancient populations are related to modern populations by analyzing read counts from multiple ancient individuals and explicitly modeling relationship between the two

242 populations. Using this approach, we examined some aspects of sampling strategy for  
243 aDNA analysis and we applied our approach to ancient humans.

244 We found that sequencing many individuals from an ancient population to low cov-  
245 erage (.5-1x) can be a significantly more cost effective strategy than sequencing fewer  
246 individuals to relatively high coverage. For instance, we saw from simulations that far  
247 more accurate estimates of the drift time in an ancient population can be obtained by  
248 pooling 2 individuals at 0.5x coverage than by sequencing a single individual to 2x cov-  
249 erage (Figure 2). We saw this replicated in our analysis of the real data: low coverage  
250 individuals showed a significant amount of variation and bias in estimating the model  
251 parameters that was substantially reduced when individuals were analyzed jointly in a  
252 population (Figure 7). To explore this further, we showed that sites sequenced to 1x  
253 coverage in a single individual retain no information about the drift time in the ancient  
254 population. This can be intuitively understood because the drift time in the ancient  
255 population is strongly related the amount of heterozygosity in the ancient population:  
256 an ancient population with a longer drift time will have lower heterozygosity at sites  
257 shared with a modern population. When a site is only sequenced once in a single indi-  
258 vidual, there is no information about the heterozygosity of that site. We also observed  
259 a pronounced upward bias in estimates of the drift time in the ancient population from  
260 low coverage samples. We speculate that this is due to the presence of few sites covered  
261 more than once being likely to be homozygous, thus deflating the estimate of heterozy-  
262 gosity in the ancient population. Thus, for analysis of SNP data, we recommend that  
263 aDNA sampling be conducted to maximize the number of individuals from each ancient  
264 population that can be sequenced to ~1x, rather than attempting to sequence fewer  
265 individuals to high coverage.

266 Of particular interest in many studies of ancient populations is the question of  
267 direct ancestry: are the ancient samples members of a population that contributed  
268 substantially to a modern population? We emphasize that this does not mean that  
269 the particular samples were direct ancestors of any modern individuals; indeed, this  
270 is exceedingly unlikely for old samples [Rohde et al., 2004, Chang, 1999, Baird et al.,

271 2003, Donnelly, 1983]. Instead, we are asking whether an ancient sample was a member  
272 of a population that is directly continuous with a modern population. Several methods  
273 have been proposed to test this question, but thus far they have been limited to many  
274 individuals sequenced at a single locus [Sjödin et al., 2014] or to a single individual with  
275 genome-wide data [Rasmussen et al., 2014]. Our approach provides a rigorous, maxi-  
276 mum likelihood framework for testing questions of population continuity using multiple  
277 low coverage ancient samples. We saw from simulations (Figure 3) that data from sin-  
278 gle, low coverage individuals result in very little power to reject the null hypothesis of  
279 continuity unless the ancient sample is very recent (i.e. it has been diverged from the  
280 modern population for a long time). Nonetheless, when low coverage individuals are  
281 pooled together, or a single high coverage individual is used, there is substantial power  
282 to reject continuity for all but the most ancient samples (i.e. samples dating from very  
283 near the population split time).

284 When we applied our approach to European history, we made several noteworthy  
285 observations. Primarily, we rejected continuity for all populations that we analyzed.  
286 This is unsurprising, given that European history is extremely complicated and has  
287 been shaped by many periods of admixture [Lazaridis et al., 2014, Haak et al., 2015,  
288 Lazaridis et al., 2016]. Thus, modern Europeans have experienced many periods of  
289 “ghost” admixture (relative to any particular ancient sample). Nonetheless, our results  
290 show that none of these populations are even particularly close to directly ancestral, as  
291 our simulations have shown that rejection of continuity is robust to low levels of ghost  
292 admixture.

293 Secondly, we observed that the drift time in the ancient population was much larger  
294 than the drift time in the modern population. Assuming that the ancient sample were  
295 a contemporary sample, the ratio  $t_1/t_2$  is an estimator of the ratio  $N_e^{(2)}/N_e^{(1)}$ . Thus,  
296  $t_2 > t_1$  implies  $N_e^{(2)} < N_e^{(1)}$ . This supports the observation that ancient Europeans  
297 were often members of small, isolated populations [Skoglund et al., 2014]. We fur-  
298 ther examined the effective sizes of ancient populations through time by looking for a  
299 correlation between the age of the ancient populations and the drift time leading to

300 them (Figure 6C). We saw a strong positive correlation, and although this drift time is  
301 a compound parameter, which complicates interpretations, it appears that the oldest  
302 Europeans were members of the smallest populations, and that effective population  
303 size has grown through time as agriculture spread through Europe.

304 We anticipate the further development of methods that explicitly account for dif-  
305 ferential drift times in ancient and modern samples will become important as aDNA  
306 research becomes even more integrating into population genomics. This is because  
307 many common summary methods, such as the use of Structure [Pritchard et al., 2000]  
308 and Admixture [Alexander et al., 2009], are sensitive to differential amounts of drift  
309 between populations [Falush et al., 2016]. As we've shown in ancient Europeans, an-  
310 cient samples tend to come from isolated subpopulations with a large amount of drift,  
311 thus confounding such summary approaches. Moreover, standard population genetics  
312 theory shows that allele frequencies are expected to be deterministically lower in an-  
313 cient samples, even if they are direct ancestors of a modern population. Intuitively,  
314 this arises because the alleles must have arisen at some point from new mutations, and  
315 thus were at lower frequencies in the past. A potentially fruitful avenue to combine  
316 these approaches moving forward may be to separate regions of the genome based on  
317 ancestry components, and assess the ancestry of ancient samples relative to specific  
318 ancestry components, rather than to genomes as a whole.

319 Our current approach leaves several avenues for improvement. We use a relatively  
320 simple error model that wraps up both post-mortem damage and sequencing error  
321 into a single parameter. While Racimo et al. [2016] shows that adding an additional  
322 parameter for PMD-related error does not significantly change results, future work  
323 could examine more robust error models. Although our method is robust to non-  
324 constant demography because we consider only alleles that are segregating in both the  
325 modern and the ancient population, we are losing information by not modeling new  
326 mutations that arise in the ancient population. Similarly, we only consider a single  
327 ancient population at a time, albeit with multiple samples. Ideally, ancient samples  
328 would be embedded in complex demographic models that include admixture, detailing

their relationships to each other and to modern populations [Patterson et al., 2012, Lipson and Reich, 2017]. However, inference of such complex models is difficult, and though there has been some progress in simplified cases [Lipson et al., 2014, Pickrell and Pritchard, 2012], it remains an open problem due to the difficult of simultaneously inferring a non-tree-like topology along with demographic parameters. Software such as `momi` [Kamm et al., 2016] that can compute the likelihood of SNP data in an admixture graph may be able to be used to integrate over genotype uncertainty in larger settings than considered here.

## 5 Appendix

### 5.1 Computing allele frequency moments in the ancient population

We wish to compute moments of the form

$$\mathbb{E}_x(g(Y); t_1, t_2) = \int_0^1 g(y)f(y; x, t_1, t_2)dy. \quad (4)$$

To do so, we make use of several results from diffusion theory. To ensure that this paper is self contained, we briefly review those results here. The interested reader may find much of this material covered in Ewens [2012], Karlin and Taylor [1981]. Several similar calculations can be found in Griffiths [2003].

Let the probability of an allele going from frequency  $x$  to frequency  $y$  in  $\tau$  generations in a population of size  $N_e$  be  $f(x, y; t)$ , where  $t = \tau/(2N_e)$ . Under a wide variety of models, the change in allele frequencies through time is well approximated by the Wright-Fisher diffusion, which is characterized by its generator,

$$\mathcal{L} = \frac{1}{2}x(1-x)\frac{d^2}{dx^2}.$$

The generator of a diffusion process is useful, because it can be used to define a differ-

350 ential equation for the moments of that process,

$$\frac{d}{dt} \mathbb{E}_x(g(X_t)) = \mathbb{E}_x(\mathcal{L}g(X_t)). \quad (5)$$

351 We will require the *speed measure* of the Wright-Fisher diffusion,  $m(x) = x^{-1}(1 -$   
352  $x)^{-1}$ , which essentially describes how slow a diffusion at position  $x$  is “moving” com-  
353 pared to a Brownian motion at position  $x$ . Note that all diffusions are reversible with  
354 respect to their speed measures, i.e.

$$m(x)f(x, y; t) = m(y)f(y, x; t).$$

355 We additionally require the probability of loss, i.e. the probability that the allele  
356 currently at frequency  $x$  is ultimately lost from the population. This is

$$u_0(x) = 1 - x.$$

357 Note that it is possible to condition the Wright-Fisher diffusion to eventually be lost.  
358 The transition density can be computed as

$$f^\downarrow(x, y; t) = f(x, y; t) \frac{u_0(y)}{u_0(x)}$$

359 by using Bayes theorem. The diffusion conditioned on loss is characterized by its  
360 generator,

$$\mathcal{L}^\downarrow = -x \frac{d}{dx} + \frac{1}{2}x(1-x) \frac{d^2}{dx^2}.$$

361 In an infinite sites model, in which mutations occur at the times of a Poisson  
362 process with rate  $\theta/2$  and then each drift according to the Wright-Fisher diffusion, a  
363 quasi-equilibrium distribution will be reached, known as the frequency spectrum. The  
364 frequency spectrum,  $\phi(x)$ , predicts the number of sites at frequency  $x$ , and can be

365 written in terms of the speed measure and the probability of loss,

$$\phi(x) = \theta m(x) u_0(x).$$

366 To proceed with calculating (4), note that the conditional probability of an allele  
367 being at frequency  $y$  in the ancient population given that it's at frequency  $x$  in the  
368 modern population can be calculated

$$f(y; x, t_1, t_2) = \frac{f(x, y; t_1, t_2)}{\phi(x)}$$

369 where  $f(x, y; t_1, t_2)$  is the joint probability of the allele frequencies in the modern and  
370 ancient populations and  $\phi(x)$  is the frequency spectrum in the modern population.

371 Assuming that the ancestral population of the modern and ancient samples was at  
372 equilibrium, the joint distribution of allele frequencies can be computed by sampling  
373 alleles from the frequency spectrum of the ancestor and evolving them forward in time  
374 via the Wright-Fisher diffusion. This can be written mathematically as

$$f(x, y; t_1, t_2) = \int_0^1 f(z, x; t_1) f(z, y; t_2) \phi(z) dz.$$

375 We now expand the frequency spectrum in terms of the speed measure and the prob-  
376 ability of loss and use reversibility with respect to the speed measure to rewrite the  
377 equation,

$$\begin{aligned} \int_0^1 f(z, x; t_1) f(z, y; t_2) \phi(z) dz &= \theta \int_0^1 f(z, x; t_1) f(z, y; t_2) m(z) u_0(z) dz \\ &= \theta \int_0^1 \frac{m(x)}{m(z)} f(x, z; t_1) f(z, y; t_2) m(z) u_0(z) dz \\ &= \theta m(x) u_0(x) \int_0^1 f(x, z; t_1) \frac{u_0(z)}{u_0(x)} f(z, y; t_2) dz \\ &= \phi(x) \int_0^1 f^\downarrow(x, z; t_1) f(z, y; t_2) dz. \end{aligned}$$

375        The third line follows by multiplying by  $u_0(x)/u_0(x) = 1$ . This equation has the interpretation of sampling an allele from the frequency spectrum in the modern population,  
 376        then evolving it *backward* in time to the common ancestor, before evolving it *forward*  
 377        in time to the ancient population. The interpretation of the diffusion conditioned on  
 378        loss as evolving backward in time arises by considering the fact that alleles arose from  
 379        unique mutations at some point in the past; hence, looking backward, alleles must  
 380        eventually be lost at some point in the past.  
 381

To compute the expectation, we substitute this form for the joint probability into  
 (4),

$$\begin{aligned}
 \int_0^1 g(y) f(y; x, t_1, t_2) dy &= \int_0^1 g(y) \left( \int_0^1 f^\downarrow(x, z; t_1) f(z, y; t_2) dz \right) dy \\
 &= \int_0^1 \left( \int_0^1 g(y) f(z, y; t_2) dy \right) f^\downarrow(x, z; t_1) dz,
 \end{aligned}$$

where the second line follows by rearranging terms and exchanging the order of integration. Note that this formula takes the form of nested expectations. Specifically,

$$\begin{aligned}
 \int_0^1 g(y) f(z, y; t_2) dy &= \mathbb{E}_z(g(Y_{t_2})) \\
 &\equiv h(z)
 \end{aligned}$$

and

$$\begin{aligned}
 \int_0^1 h(z) f^\downarrow(x, z; t_1) dz &= \mathbb{E}_x^\downarrow(h(Z_{t_1})) \\
 &= \mathbb{E}_x(g(Y); t_1, t_2).
 \end{aligned}$$

382        We now use (5) to note that

$$\frac{d}{dt} p_{n,k} = \frac{k(k-1)}{2} p_{n,k-1} - k(n-k) p_{n,k} + \frac{(n-k)(n-k-1)}{2} p_{n,k+1}$$

383

and

$$\frac{d}{dt} p_{n,k}^{\downarrow} = \frac{k(k-1)}{2} p_{n,k-1}^{\downarrow} - k(n-k+1) p_{n,k}^{\downarrow} + \frac{(n-k+1)(n-k)}{2} p_{n,k+1}^{\downarrow}$$

384

with obvious boundary conditions  $p_{n,k}(0; z) = z^k(1-z)^{n-k}$  and  $p_{n,k}^{\downarrow}(0; x) = x^k(1-x)^{n-k}$ .

386

These systems of differential equations can be rewritten as matrix differential equations with coefficient matrices  $Q$  and  $Q^{\downarrow}$  respectively. Because they are linear, first order equations, they can be solved by matrix exponentiation. Because the expectation of a polynomial in the Wright-Fisher diffusion remains a polynomial, the nested expectations can be computed via matrix multiplication of the solutions to these differential equations, yielding the formula (2).

392

## 5.2 Sites covered exactly once have no information about drift in the ancient population

394

Consider a simplified model in which each site has exactly one read. When we have sequence from only a single individual, we have a set  $l_a$  of sites where the single read is an ancestral allele and a set  $l_d$  of sites where the single read is a derived allele. Thus, we can rewrite (3) as

$$L(D) = \prod_{l \in l_a} \left( (1-\epsilon)P_{2,0}(x_l) + \frac{1}{2}P_{2,1}(x_l) + \epsilon P_{2,2}(x_l) \right) \prod_{l \in l_d} \left( \epsilon P_{2,0}(x_l) + \frac{1}{2}P_{2,1}(x_l) + (1-\epsilon)P_{2,2}(x_l) \right).$$

We can use formulas from Racimo et al. [2016] to compute  $P_{2,k}(x_l)$  for  $k \in \{0, 1, 2\}$ ,

$$\begin{aligned} P_{2,0}(x_l) &= 1 - x_l e^{-t_1} - \frac{1}{2} x_l e^{-(t_1+t_2)} + x_l \left( x_l - \frac{1}{2} \right) e^{-(3t_1+t_2)} \\ P_{2,1}(x_l) &= x_l e^{-(t_1+t_2)} + x_l (1 - 2x_l) e^{-(3t_1+t_2)} \\ P_{2,2}(x_l) &= x_l e^{-t_1} - \frac{1}{2} x_l e^{-(t_1+t_2)} + x_l \left( x_l - \frac{1}{2} \right) e^{-(3t_1-t_2)}. \end{aligned}$$

398 Note then that

$$(1 - \epsilon)P_{2,0}(x_l) + \frac{1}{2}P_{2,1}(x_l) + \epsilon P_{2,2}(x_l) = 1 - \epsilon - x(1 - 2\epsilon)e^{-t_1}$$

399 and

$$\epsilon P_{2,0}(x_l) + \frac{1}{2}P_{2,1}(x_l) + (1 - \epsilon)P_{2,2}(x_l) = \epsilon + x(1 - 2\epsilon)e^{-t_1}.$$

400 Neither of these formulas depend on  $t_2$ ; hence, there is no information about the drift  
401 time in the ancient population from data that is exactly 1x coverage.

## 402 6 Software Availability

403 Python implementations of the described methods are available at [www.github.com/schraiber/continuity/](http://www.github.com/schraiber/continuity/)

## 404 7 Acknowledgments

405 We wish to thank Melinda Yang, Iain Mathieson, Pontus Skoglund, and Fernando  
406 Racimo for several discussions during the conception of this work that greatly improved  
407 its scope and rigor. We are grateful to Fernando Racimo and Benjamin Vernot for  
408 comments on early versions of this work that significantly improved its quality. We  
409 also appreciate comments on the preprint from Alex Kim. We would also like to thank  
410 Tamara Broderick for several stimulating discussions about clustering that ultimately  
411 didn't result in any interesting application to ancient DNA.

## 412 References

413 David H Alexander, John Novembre, and Kenneth Lange. Fast model-based estimation  
414 of ancestry in unrelated individuals. *Genome research*, 19(9):1655–1664, 2009.

415 Morten E Allentoft, Martin Sikora, Karl-Göran Sjögren, Simon Rasmussen, Morten  
416 Rasmussen, Jesper Stenderup, Peter B Damgaard, Hannes Schroeder, Torbjörn

- 417 Ahlström, Lasse Vinner, et al. Population genomics of bronze age eurasia. *Nature*,  
418 522(7555):167–172, 2015.
- 419 SJE Baird, NH Barton, and AM Etheridge. The distribution of surviving blocks of an  
420 ancestral genome. *Theoretical population biology*, 64(4):451–471, 2003.
- 421 Joseph T Chang. Recent common ancestors of all present-day individuals. *Advances*  
422 *in Applied Probability*, 31(4):1002–1026, 1999.
- 423 1000 Genomes Project Consortium. A global reference for human genetic variation.  
424 *Nature*, 526(7571):68–74, 2015.
- 425 Jesse Dabney, Matthias Meyer, and Svante Pääbo. Ancient dna damage. *Cold Spring*  
426 *Harbor perspectives in biology*, 5(7):a012567, 2013.
- 427 Kevin P Donnelly. The probability that related individuals share some section of  
428 genome identical by descent. *Theoretical population biology*, 23(1):34–63, 1983.
- 429 Warren J Ewens. *Mathematical population genetics 1: theoretical introduction*, vol-  
430 ume 27. Springer Science & Business Media, 2012.
- 431 Daniel Falush, Lucy van Dorp, and Daniel Lawson. A tutorial on how (not) to over-  
432 interpret structure/admixture bar plots. *bioRxiv*, page 066431, 2016.
- 433 Qiaomei Fu, Cosimo Posth, Mateja Hajdinjak, Martin Petr, Swapan Mallick, Daniel  
434 Fernandes, Anja Furtwängler, Wolfgang Haak, Matthias Meyer, Alissa Mittnik, et al.  
435 The genetic history of ice age europe. *Nature*, 2016.
- 436 Richard E Green, Johannes Krause, Adrian W Briggs, Tomislav Maricic, Udo Stenzel,  
437 Martin Kircher, Nick Patterson, Heng Li, Weiwei Zhai, Markus Hsi-Yang Fritz, et al.  
438 A draft sequence of the neandertal genome. *science*, 328(5979):710–722, 2010.
- 439 RC Griffiths. The frequency spectrum of a mutation, and its age, in a general diffusion  
440 model. *Theoretical population biology*, 64(2):241–251, 2003.

- 441 Wolfgang Haak, Iosif Lazaridis, Nick Patterson, Nadin Rohland, Swapan Mallick,  
442 Bastien Llamas, Guido Brandt, Susanne Nordenfelt, Eadaoin Harney, Kristin Stew-  
443 ardson, et al. Massive migration from the steppe was a source for indo-european  
444 languages in europe. *Nature*, 522(7555):207–211, 2015.
- 445 Ethan M Jewett, Matthias Steinrücken, and Yun S Song. The effects of population  
446 size histories on estimates of selection coefficients from time-series genetic data.  
447 *Molecular Biology and Evolution*, page msw173, 2016.
- 448 John A Kamm, Jonathan Terhorst, and Yun S Song. Efficient computation of the joint  
449 sample frequency spectra for multiple populations. *Journal of Computational and*  
450 *Graphical Statistics*, (just-accepted):1–37, 2016.
- 451 Samuel Karlin and Howard E Taylor. *A second course in stochastic processes*. Elsevier,  
452 1981.
- 453 Jerome Kelleher, Alison M Etheridge, and Gilean McVean. Efficient coalescent sim-  
454 ulation and genealogical analysis for large sample sizes. *PLoS Comput Biol*, 12(5):  
455 e1004842, 2016.
- 456 Thorfinn Sand Korneliussen, Anders Albrechtsen, and Rasmus Nielsen. Angsd: analysis  
457 of next generation sequencing data. *BMC bioinformatics*, 15(1):356, 2014.
- 458 Iosif Lazaridis, Nick Patterson, Alissa Mitnik, Gabriel Renaud, Swapan Mallick,  
459 Karola Kirsanow, Peter H Sudmant, Joshua G Schraiber, Sergi Castellano, Mark  
460 Lipson, et al. Ancient human genomes suggest three ancestral populations for  
461 present-day europeans. *Nature*, 513(7518):409–413, 2014.
- 462 Iosif Lazaridis, Dani Nadel, Gary Rollefson, Deborah C Merrett, Nadin Rohland, Swa-  
463 pan Mallick, Daniel Fernandes, Mario Novak, Beatriz Gamarra, Kendra Sirak, et al.  
464 Genomic insights into the origin of farming in the ancient near east. *Nature*, 536  
465 (7617):419–424, 2016.

- 466 Mark Lipson and David Reich. Working model of the deep relationships of diverse  
467 modern human genetic lineages outside of africa. *Molecular Biology and Evolution*,  
468 page msw293, 2017.
- 469 Mark Lipson, Po-Ru Loh, Nick Patterson, Priya Moorjani, Ying-Chin Ko, Mark  
470 Stoneking, Bonnie Berger, and David Reich. Reconstructing austronesian popu-  
471 lation history in island southeast asia. *Nature communications*, 5, 2014.
- 472 Iain Mathieson, Iosif Lazaridis, Nadin Rohland, Swapan Mallick, Nick Patterson,  
473 Songül Alpaslan Roodenberg, Eadaoin Harney, Kristin Stewardson, Daniel Fernan-  
474 des, Mario Novak, et al. Genome-wide patterns of selection in 230 ancient eurasians.  
475 *Nature*, 528(7583):499–503, 2015.
- 476 Matthias Meyer, Martin Kircher, Marie-Theres Gansauge, Heng Li, Fernando Racimo,  
477 Swapan Mallick, Joshua G Schraiber, Flora Jay, Kay Prüfer, Cesare De Filippo, et al.  
478 A high-coverage genome sequence from an archaic denisovan individual. *Science*, 338  
479 (6104):222–226, 2012.
- 480 Rasmus Nielsen, Thorfinn Korneliussen, Anders Albrechtsen, Yingrui Li, and Jun  
481 Wang. Snp calling, genotype calling, and sample allele frequency estimation from  
482 new-generation sequencing data. *PloS one*, 7(7):e37558, 2012.
- 483 Nick Patterson, Priya Moorjani, Yontao Luo, Swapan Mallick, Nadin Rohland, Yiping  
484 Zhan, Teri Genschoreck, Teresa Webster, and David Reich. Ancient admixture in  
485 human history. *Genetics*, 192(3):1065–1093, 2012.
- 486 Benjamin M Peter. Admixture, population structure, and f-statistics. *Genetics*, 202  
487 (4):1485–1501, 2016.
- 488 Joseph K Pickrell and Jonathan K Pritchard. Inference of population splits and mix-  
489 tures from genome-wide allele frequency data. *PLoS Genet*, 8(11):e1002967, 2012.
- 490 Ron Pinhasi, Daniel Fernandes, Kendra Sirak, Mario Novak, Sarah Connell, Songül  
491 Alpaslan-Roodenberg, Fokke Gerritsen, Vyacheslav Moiseyev, Andrey Gromov, Pál

- 492 Raczky, et al. Optimal ancient dna yields from the inner ear part of the human  
493 petrous bone. *PloS one*, 10(6):e0129102, 2015.
- 494 Jonathan K Pritchard, Matthew Stephens, and Peter Donnelly. Inference of population  
495 structure using multilocus genotype data. *Genetics*, 155(2):945–959, 2000.
- 496 Fernando Racimo, Gabriel Renaud, and Montgomery Slatkin. Joint estimation of  
497 contamination, error and demography for nuclear dna from ancient humans. *PLoS  
498 Genet*, 12(4):e1005972, 2016.
- 499 Morten Rasmussen, Sarah L Anzick, Michael R Waters, Pontus Skoglund, Michael De-  
500 Giorgio, Thomas W Stafford Jr, Simon Rasmussen, Ida Moltke, Anders Albrechtsen,  
501 Shane M Doyle, et al. The genome of a late pleistocene human from a clovis burial  
502 site in western montana. *Nature*, 506(7487):225–229, 2014.
- 503 Douglas LT Rohde, Steve Olson, and Joseph T Chang. Modelling the recent common  
504 ancestry of all living humans. *Nature*, 431(7008):562, 2004.
- 505 Susanna Sawyer, Johannes Krause, Katerina Guschanski, Vincent Savolainen, and  
506 Svante Pääbo. Temporal patterns of nucleotide misincorporations and dna frag-  
507 mentation in ancient dna. *PloS one*, 7(3):e34131, 2012.
- 508 Joshua G Schraiber, Steven N Evans, and Montgomery Slatkin. Bayesian inference of  
509 natural selection from allele frequency time series. *Genetics*, 203(1):493–511, 2016.
- 510 Per Sjödin, Pontus Skoglund, and Mattias Jakobsson. Assessing the maximum con-  
511 tribution from ancient populations. *Molecular biology and evolution*, page msu059,  
512 2014.
- 513 Pontus Skoglund, Helena Malmström, Ayça Omrak, Maanasa Raghavan, Cristina  
514 Valdiosera, Torsten Günther, Per Hall, Kristiina Tambets, Jüri Parik, Karl-Göran  
515 Sjögren, et al. Genomic diversity and admixture differs for stone-age scandinavian  
516 foragers and farmers. *Science*, 344(6185):747–750, 2014.

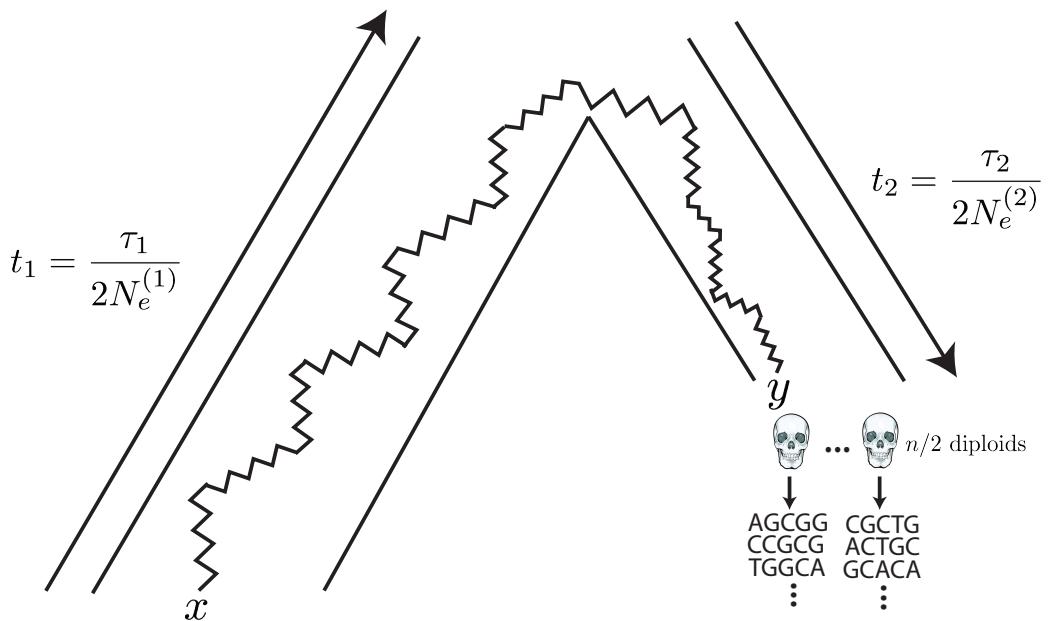


Figure 1: The generative model. Alleles are found at frequency  $x$  in the modern population and are at frequency  $y$  in the ancient population. The modern population has effective size  $N_e^{(1)}$  and has evolved for  $\tau_1$  generations since the common ancestor of the modern and ancient populations, while the ancient population is of size  $N_e^{(2)}$  and has evolved for  $\tau_2$  generations. Ancient diploid samples are taken and sequenced to possibly low coverage, with errors. Arrows indicate that the sampling probability can be calculated by evolving alleles *backward* in time from the modern population and then forward in time to the ancient population.

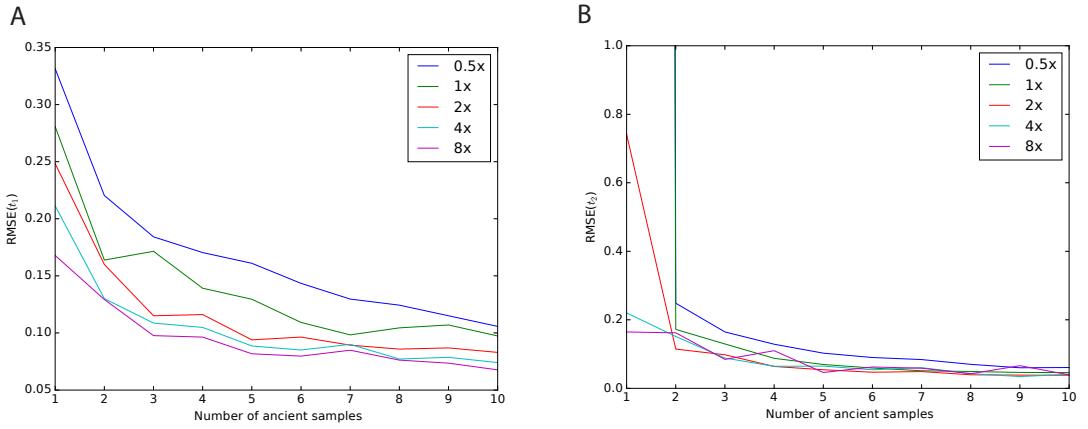


Figure 2: Impact of sampling scheme on parameter estimation error. In each panel, the  $x$  axis represents the number of simulated ancient samples, while the  $y$  axis shows the relative root mean square error for each parameter. Each different line corresponds to individuals sequenced to different depth of coverage. Panel A shows results for  $t_1$  while panel B shows results for  $t_2$ . Simulated parameters are  $t_1 = 0.02$  and  $t_2 = 0.05$ .

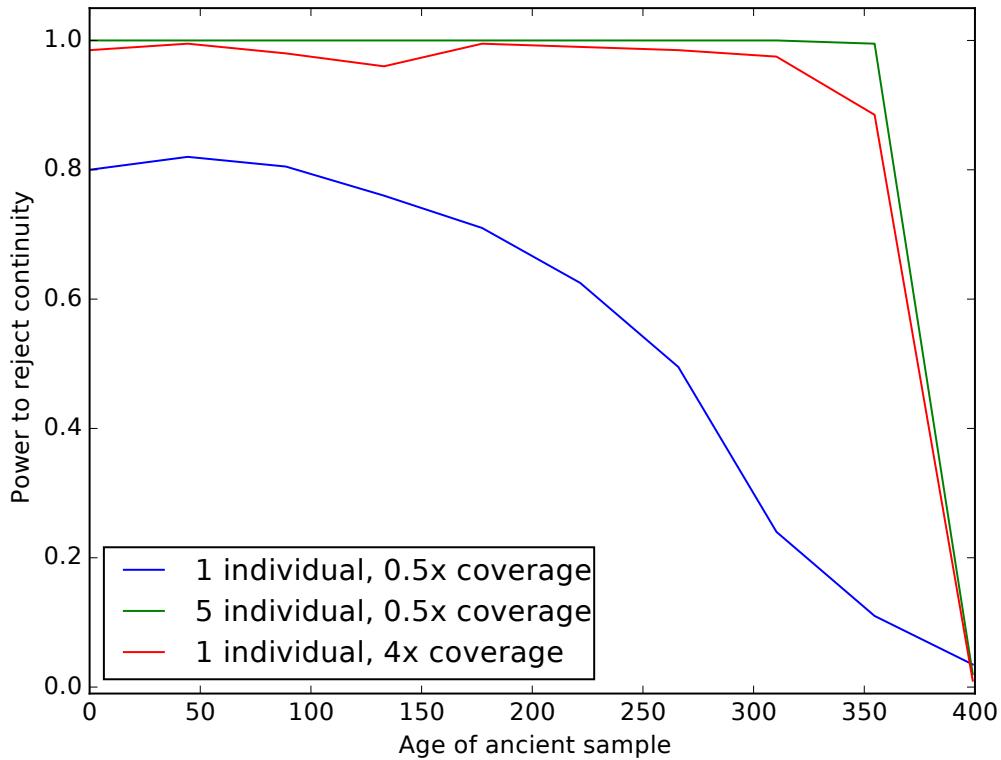


Figure 3: Impact of sampling scheme on rejecting population continuity. The  $x$  axis represents the age of the ancient sample in generations, with 0 indicating a modern sample and 400 indicating a sample from exactly at the split time 400 generations ago. The  $y$  axis shows the proportion of simulations in which we rejected the null hypothesis of population continuity. Each line shows different sampling schemes, as explained in the legend.

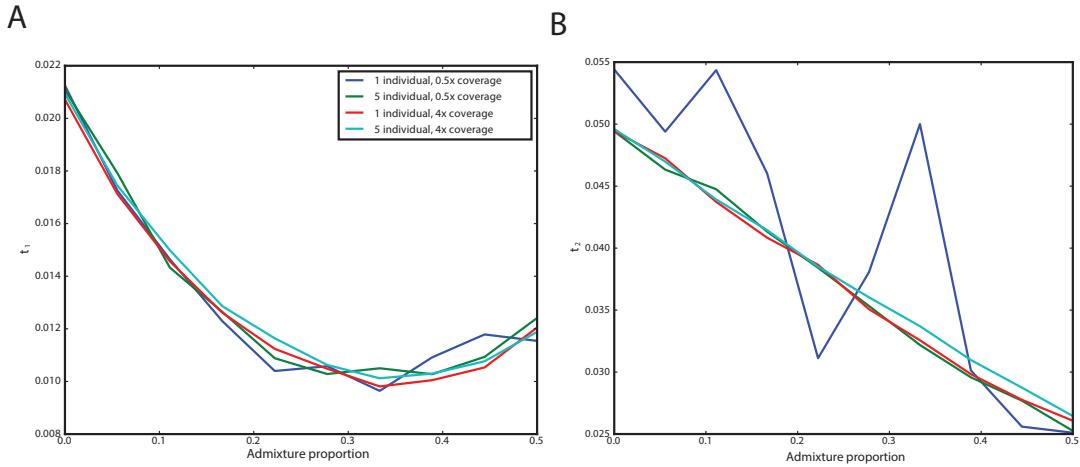


Figure 4: Impact of admixture from the ancient population on inferred parameters. The  $x$  axis shows the admixture proportion and the  $y$  axis shows the average parameter estimate across simulations. Each line corresponds to a different sampling strategy, as indicated in the legend. Panel A shows results for  $t_1$  and Panel B shows results for  $t_2$ .

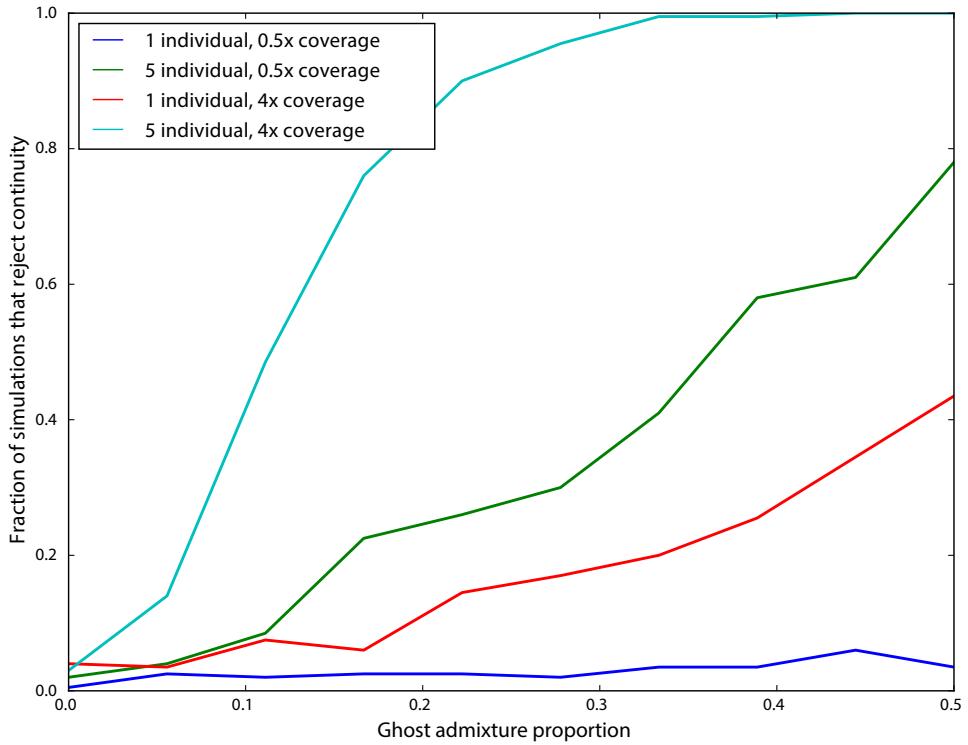


Figure 5: Impact of ghost admixture on falsely rejecting continuity. The  $x$  axis shows the admixture proportion from the ghost population, and the  $y$  axis shows the fraction of simulations in which continuity was rejected. Each line corresponds to a different sampling strategy, as indicated in the legend.

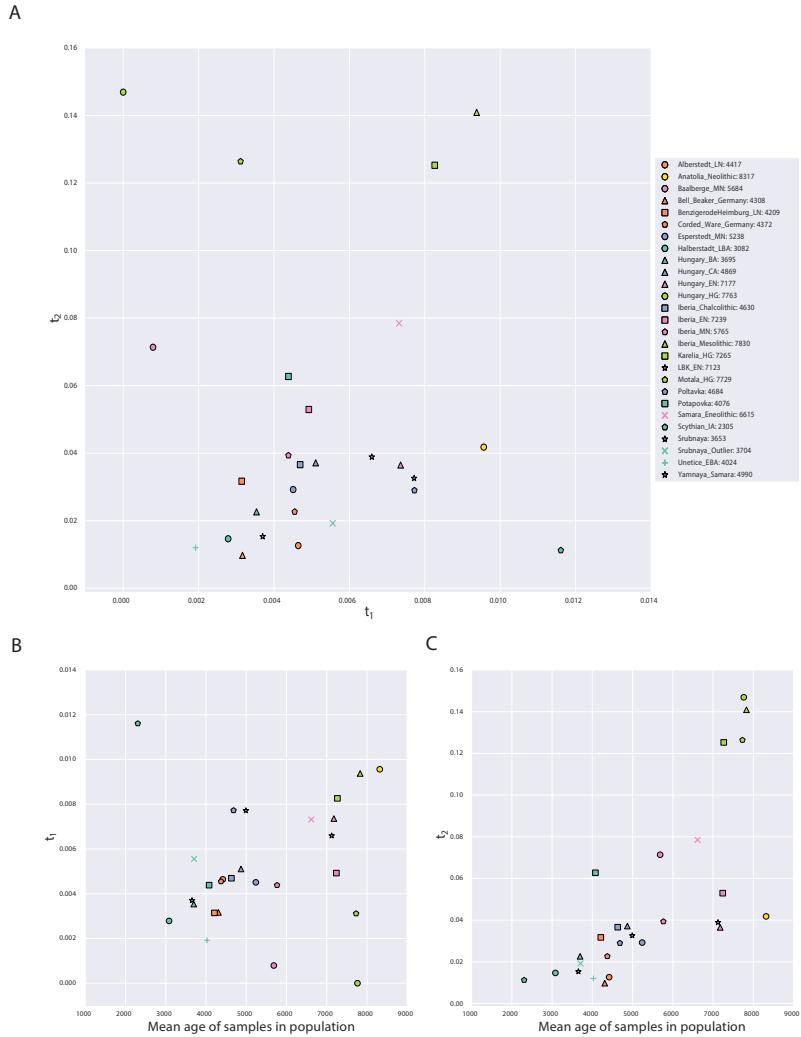


Figure 6: Parameters of the model inferred from ancient West Eurasian samples. Panel A shows  $t_1$  on the x-axis and  $t_2$  on the y-axis, with each point corresponding to a population as indicated in the legend. Numbers in the legend correspond to the mean date of all samples in the population. Panels B and C show scatterplots of the mean age of the samples in the population (x-axis) against  $t_1$  and  $t_2$ , respectively. Points are described by the same legend as Panel A.

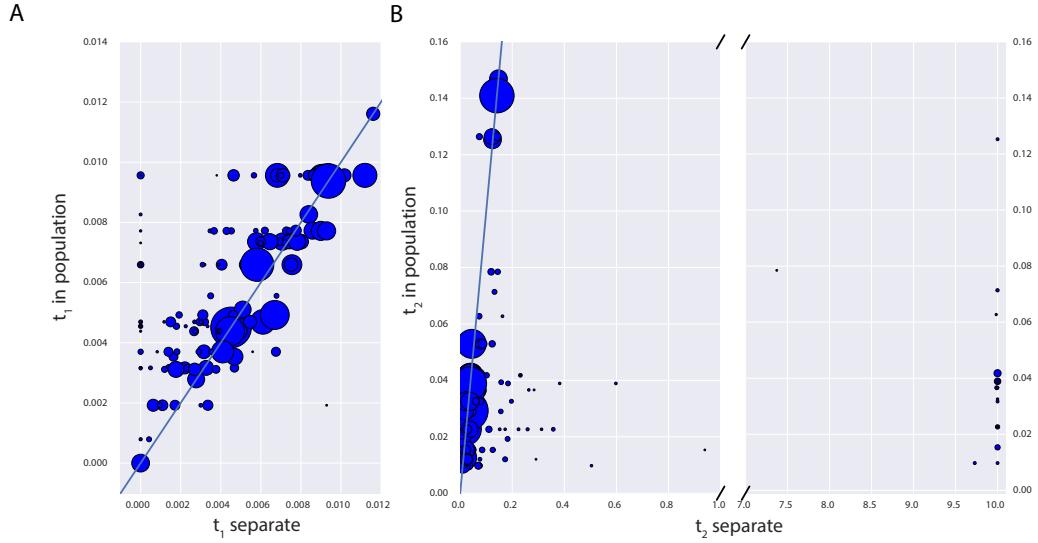


Figure 7: Impact of pooling individuals into populations when estimating model parameters from real data. In both panels, the x-axis indicates the parameter estimate when individuals are analyzed separately, while the y-axis indicates the parameter estimate when individuals are grouped into populations. Size of points is proportional to the coverage of each individual. Panel A reports the impact on estimation of  $t_1$ , while Panel B reports the impact on  $t_2$ . Note that Panel B has a broken x-axis. Solid lines in each figure indicate  $y = x$ .

pop	cov	date	$t_1$	$t_2$	lnL	$t_1$ (cont)	lnL (cont)
Alberstedt_LN	12.606	4417.000	0.005	0.013	-779411.494	0.006	-779440.143
Anatolia_Neolithic	3.551	8317.500	0.010	0.042	-9096440.714	0.044	-9106156.877
Baalberge_MN	0.244	5684.333	0.001	0.071	-201575.306	0.007	-201750.419
Bell_Beaker_Germany	1.161	4308.444	0.003	0.010	-1834486.744	0.008	-1834652.858
BenzigerodeHeimburg_LN	0.798	4209.750	0.003	0.032	-346061.545	0.007	-346134.356
Corded_Ware_Germany	2.250	4372.833	0.005	0.023	-2139002.723	0.017	-2139858.192
Esperstedt_MN	30.410	5238.000	0.005	0.029	-975890.329	0.009	-976047.889
Halberstadt_LBA	5.322	3082.000	0.003	0.015	-558966.522	0.004	-558993.078
Hungary_BA	3.401	3695.750	0.004	0.023	-789754.969	0.010	-789939.889
Hungary_CA	5.169	4869.500	0.005	0.037	-504413.094	0.010	-504549.603
Hungary_EN	4.033	7177.000	0.007	0.036	-3478429.262	0.033	-3481855.461
Hungary_HG	5.807	7763.000	0.000	0.147	-469887.471	0.015	-471652.083
Iberia_Chalcolithic	1.686	4630.625	0.005	0.037	-2351769.869	0.028	-2354249.543
Iberia_EN	4.875	7239.500	0.005	0.053	-1483274.628	0.030	-1485675.934
Iberia_MN	5.458	5765.000	0.004	0.039	-1491407.962	0.023	-1492793.179
Iberia_Mesolithic	21.838	7830.000	0.009	0.141	-720759.133	0.030	-723091.935
Karelia_HG	2.953	7265.000	0.008	0.125	-652952.676	0.033	-655352.439
LBK_EN	2.894	7123.429	0.007	0.039	-3656617.954	0.033	-3660838.639
Motala_HG	2.207	7729.500	0.003	0.126	-1477338.076	0.068	-1489573.895
Poltavka	2.211	4684.500	0.008	0.029	-1334662.071	0.020	-1335358.630
Potapovka	0.267	4076.500	0.004	0.063	-220112.816	0.011	-220251.379
Samara_Eneolithic	0.463	6615.000	0.007	0.078	-362161.674	0.020	-362689.209
Scythian_IA	3.217	2305.000	0.012	0.011	-492961.306	0.013	-492973.694
Srubnaya	1.662	3653.273	0.004	0.015	-2578065.957	0.013	-2578645.731
Srubnaya_Outlier	0.542	3704.500	0.006	0.019	-285828.766	0.008	-285851.523
Unetice_EBA	1.320	4024.786	0.002	0.012	-1676798.610	0.008	-1677026.310
Yamnaya_Samara	1.937	4990.500	0.008	0.033	-2440183.354	0.028	-2442192.801

Table 1: Details of populations included in analysis. “pop” is population name, “cov” is mean coverage of individuals in the population, “date” is mean date of individuals in the population, “ $t_1$ ” is the maximum likelihood estimate of  $t_1$  in the full model, “ $t_2$ ” is the maximum likelihood estimate of  $t_2$  in the full model, “LnL” is the maximum likelihood value in the full model, “ $t_1$  (cont)” is the maximum likelihood estimate of  $t_1$  in the model where  $t_2 = 0$ , “LnL” is the maximum likelihood value in the model where  $t_2 = 0$ .