Deep coalescent history of the hominin lineage

Trevor Cousins[1], Richard Durbin[1], Regev Schweiger[1,2,3,*]

[1] Department of Genetics, University of Cambridge, Cambridge, UK
[2] Department of Human Genetics and Computational Medicine, Gray Faculty of Medical and Health Sciences, Tel Aviv University, Tel Aviv, Israel
[3] Blavatnik School of Computer Science and AI, Tel Aviv University, Tel Aviv, Israel

[*] Corresponding author. Email: schweiger@tauex.tau.ac.il

April 9, 2026

1 Abstract

Coalescent-based methods are widely used to infer population size histories, but existing analyses have limited resolution at deep time scales (>2 million years ago). Here, using new high-quality telomere-to-telomere genome assemblies, we extend the scope of such inference and reanalyse an ancient peak in human effective population size around 3–6 million years ago, showing that coalescent-based inference can reach much further into the past than previously thought. Furthermore, peaks at approximately the same time as in humans are also observed in chimpanzee and bonobo, but not in gorilla or orangutans. We show that this pattern is unlikely to be an artefact of model violations and discuss its potential implications for understanding hominin evolutionary history. It is known that such peaks can arise from ancestral genetic structure, and consequently we suggest that a parsimonious explanation for the contemporaneous peaks in humans, chimpanzees and bonobos would be a period of complex speciation before final separation of *Homo* from *Pan*, in contrast to a scenario where these lineages diverge cleanly from a panmictic common ancestor.

2 Introduction

The study of human population size over time provides a valuable lens through which to explore key questions about our evolutionary past [1]. Effective population size history, $N(t)$, taken here as the inverse of the coalescence rate at time t, can be inferred from present-day genome sequences by examining patterns of genetic variation within and between populations. Various methods have been developed for such inference, using information such as the site frequency spectrum [2, 3], linkage disequilibrium [4, 5], or the genomic distribution of heterozygosity [6, 7, 8]. These approaches have provided important insights into the evolutionary history of humans, including the magnitude of exponential growth over the last 20 thousand years (kya), the timing and severity of the out-of-Africa bottleneck around 50-60 kya, and divergence times among human subpopulations as well as between humans, Neanderthals and Denisovans [9, 10]. However, methods to infer human $N(t)$ have typically focused on evolution within the past 1–2 million years (Mya), even though the mean coalescence time in humans is ∼1 Mya and present-day genomes should therefore contain information about substantially older events. Extending inference to these periods could help illuminate the emergence of the genera *Homo* (∼3 Mya [11, 12]) and *Australopithecus* (∼4.5 Mya [13]), as well as

the ancestral population of humans, chimpanzees, and bonobos (∼6 Mya [14]), providing insight into deep hominin evolutionary history.

One class of approaches for studying these timescales analyses divergence among species by comparing genome sequences from humans, chimpanzees, and gorillas. Methods such as CoalHMM [15] and its successors [14, 16] employ coalescent hidden Markov models in which the hidden states represent the ordering of coalescent events, and in some formulations also ancestral coalescence times. While these methods typically estimate divergence times and ancestral population sizes, they focus on a limited set of parameters, rather than reconstructing a continuous population size history through time. A separate class of methods infers population size histories within species, most notably the pairwise sequentially Markovian coalescent (PSMC) [6] and its successors [7, 17, 18, 19, 20], based on the sequentially Markovian coalescent (SMC) model [21, 22]. In principle, estimates of $N(t)$ from these methods may extend beyond 1-2 Mya, but inference at such depths has typically not been discussed, partly because the robustness of the SMC approximation under model violations at these timescales remains unclear. As a result, population dynamics during deep hominin evolutionary history remain poorly understood.

Another key limitation in inferring deep hominin evolutionary history has been the quality of available genome assemblies. Most human demographic analyses have relied on short-read data aligned to reference genomes derived from incomplete assemblies that collapse or omit difficult genomic regions [15, 23, 24, 14, 16]. In addition, assemblies for non-human apes are typically aligned to the human reference genome, or incorporated into multiple genome alignments, despite millions of years of divergence between these species. This evolutionary distance can introduce reference bias in variant calling [25], which may in turn affect downstream demographic inference. Recently, the advent of accurate long-read sequencing [26, 27] has enabled telomere-to-telomere (T2T) assemblies that provide highly accurate, haplotype-resolved reference genomes for humans [27, 28, 29] as well as for a chimpanzee, bonobo, gorilla, Bornean orangutan, and Sumatran orangutan [30]. These assemblies expand the fraction of the genome that can be analysed while also substantially reducing sequencing and assembly errors. Because demographic inference is highly sensitive to false mutation calls and missing sequence, the increased completeness and accuracy of T2T genomes should enable more reliable inference at deeper evolutionary timescales.

Here we extend coalescent-based population size inference further into deep timescales. We examine an ancient peak in inferred effective population size around 3–6 Mya observed in PSMC analyses of human, chimpanzee, and bonobo genomes, but not in gorilla or orangutan genomes, a signal whose traces have appeared in several earlier studies but has received little attention [6, 9, 10, 24, 31, 32]. Notably, this interval overlaps with the estimated divergence of the *Homo* and *Pan* lineages (henceforth the *Homo–Pan* (HP) divergence, ∼6 Mya [14, 16, 30]), a transition debated as either being a clean split or having a period of post-divergence gene flow [33, 34, 14, 35, 36, 37]. Using T2T genome assemblies, we reanalyse this signal and assess potential model violations affecting these inferences. Our results are consistent with gene exchange between the ancestral human and chimpanzee–bonobo lineages prior to their final separation, highlighting the potential of coalescent-based inference to provide new insight into deep hominin evolutionary history.

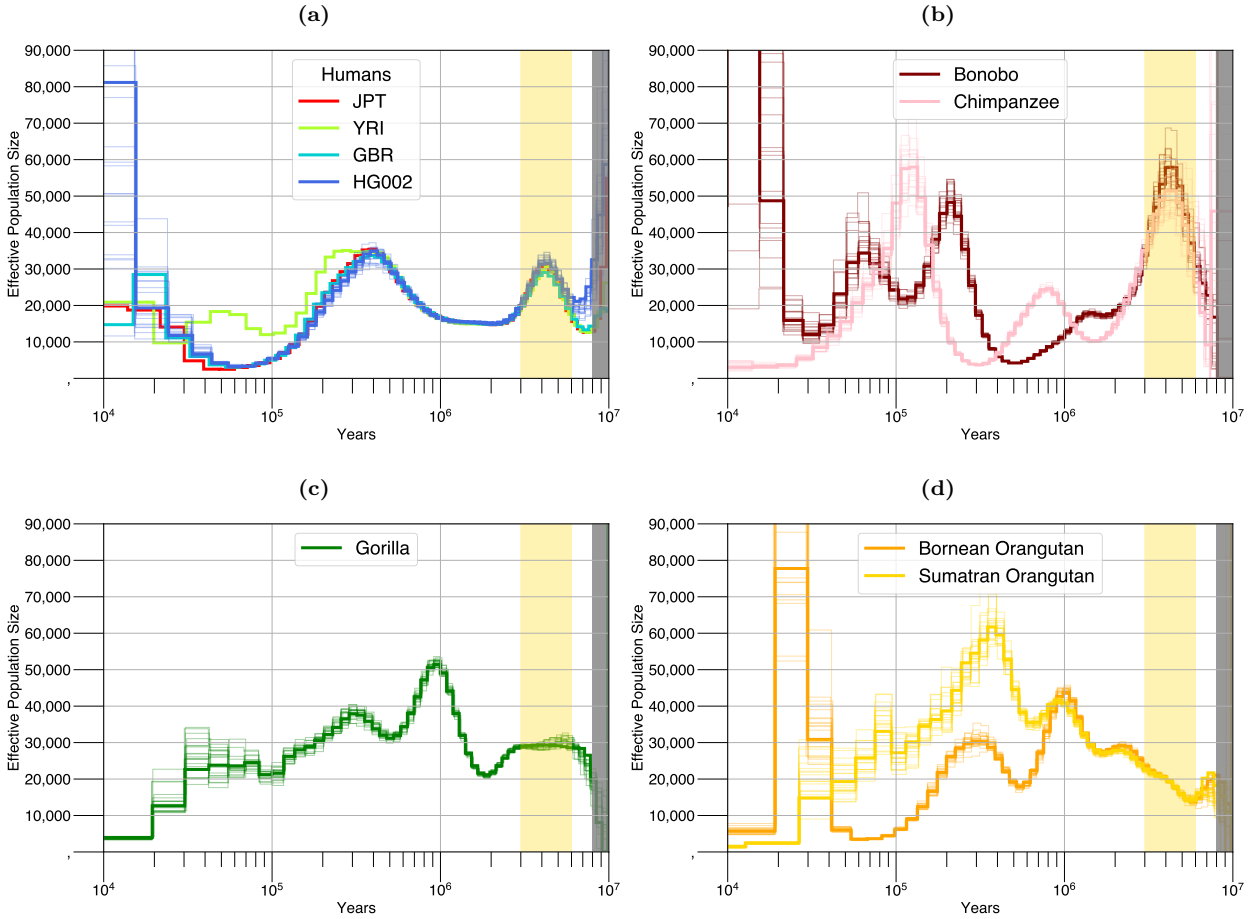

Figure 1: Inference from PSMC in humans, chimpanzees and bonobos, gorillas, and orangutans. **(a)** Estimates of population size history in humans, using HG002 and 3 individuals each from distinct populations in the 1000 Genomes project. **(b)** Estimates of population size history in chimpanzee and bonobo, using mPanTro3 and mPanPan1. **(c)** Estimates of population size history in gorilla, using mGorGor1. **(d)** Estimates of the population size history in Bornean and Sumatran orangutans, using mPonAbe1 and mPonPyg2, respectively. The faint lines in each plot show 20 iterations of non-parametric block bootstrapping.

3 Results

3.1 An ancient peak in human, chimpanzee, and bonobo population size history

We used the T2T haplotype-resolved HG002 human genome reference [27, 28, 29] to align the alternate assembly to the primary assembly and call variants (see Methods). We additionally aligned 26 high-coverage genomes from various populations from the 1000 Genomes Project (1000GP) [38, 39, 32, 40] to the HG002 primary assembly and performed variant calling (see Methods). For each of these genomes, we used MSMC2 [41] to infer an effective population size ($N(t)$) curve over time. Relative to previous similar analyses, we extended the inference by applying more stringent variant calling and by fitting the model with more time bins and iterations, and fewer parameter constraints [42] (see Methods), allowing greater resolution at timescales beyond 2 Mya. The resulting $N(t)$ estimates span from 10 kya to 10 Mya, with consistent bootstrap results between 20 kya and 8 Mya (Figure 1a), extending substantially deeper into the past than has been achieved in

previous studies [6, 9, 10, 31]. The inference reveals two peaks, one approximately 400 kya and a second older one, starting around 3 Mya and extending further back in time, reaching a maximum around 4.5 Mya with a population size of ∼32,500, then decreasing until around 6 Mya (shaded in yellow in Figure 1). This pattern is stable across 20 bootstrap replicates and consistent across each of the 26 1000GP populations (Figure A1).

As the ancient peak occurs around the estimated time at which humans split from the common ancestor of chimpanzees and bonobos, it is natural to ask if a similar signal exists in chimpanzee and bonobo population size history, or in other great apes in general. Using T2T assemblies from the T2T-Primates project [30], we therefore applied MSMC2 to chimpanzee (mPanTro3), bonobo (mPanPan1), gorilla (mGorGor1), Sumatran orangutan (mPonAbe1) and Bornean orangutan (mPonPyg2), to infer their population size histories. First, we recovered an ancient peak in the chimpanzee and bonobo, occurring at the same time as the peak seen in the human analysis (Figure 1b). As with humans, this extends further back in time a signal present in previous analyses but not discussed [24]. We see the ancient peak more clearly than previously because we avoid grouping old time intervals and due to the lower error rates of the T2T assemblies, as for human. The fact that there are peaks in all three of humans, chimpanzee and bonobo that occur within the time frame of the estimated HP divergence prompts the question of whether these peaks signify shared demographic events before the lineages fully separated into distinct species. We note that the effective population size of the chimpanzee and bonobo ancient peak is significantly higher than that of humans; we show in Figure A2 that such height differences can occur from shared ancestral structure where gene flow rates are asymmetric, while acknowledging that there are many other possible explanations for such shared patterns - we return to this in the Discussion.

For the gorilla (Figure 1c) and the orangutans (Figure 1d), we do not detect a discernible peak in the region between 3-6 Mya. This is consistent with the fact that gorillas and orangutans genetically separated from humans, chimpanzees and bonobos prior to this period, and so would not be expected to share their history at that time (gorillas are estimated to have separated from the ancestral human, chimpanzee and bonobo population around 10 Mya, and their joint ancestral population is estimated to have separated from orangutans around 19 Mya [30]). Moreover, it supports the claim that the ancient peak observed in humans, chimpanzees and bonobos is not an artefact of the method or of the data processing, as that would result in a peak with gorillas and orangutans as well.

3.2 The fraction of uncoalesced genome as a function of time

The power of PSMC to infer $N(t)$ at ∼3-6 Mya relies on there being sufficiently many places in the genome where the two input lineages coalesce more anciently than this time. To assess whether this is reasonable to expect, we note that under a simple panmictic model the observed range of heterozygosity in humans (e.g., ∼0.00069 for PEL to ∼0.001 for MSL in 1000GP) would correspond to a constant effective diploid population size of 13,800-20,000, using a mutation rate of 1.25e-8 per basepair per generation. Given an exponential distribution of coalescence times, and a generation time of 29 years, we expect a mean coalescence time of 800,400-1,160,000 years, and that the coalescence time will be older than 5 Mya in 0.02%-1.3% of the genome. Similarly, given the observed heterozygosity in the chimpanzee sample (∼0.0008 for mPanTro3) and a mutation rate of 1.78e-8 per base pair per generation and generation time of 24, we expect a mean coalescence time of 540,000 years and that the coalescence time will be older than 5 Mya in 0.01% of the genome.

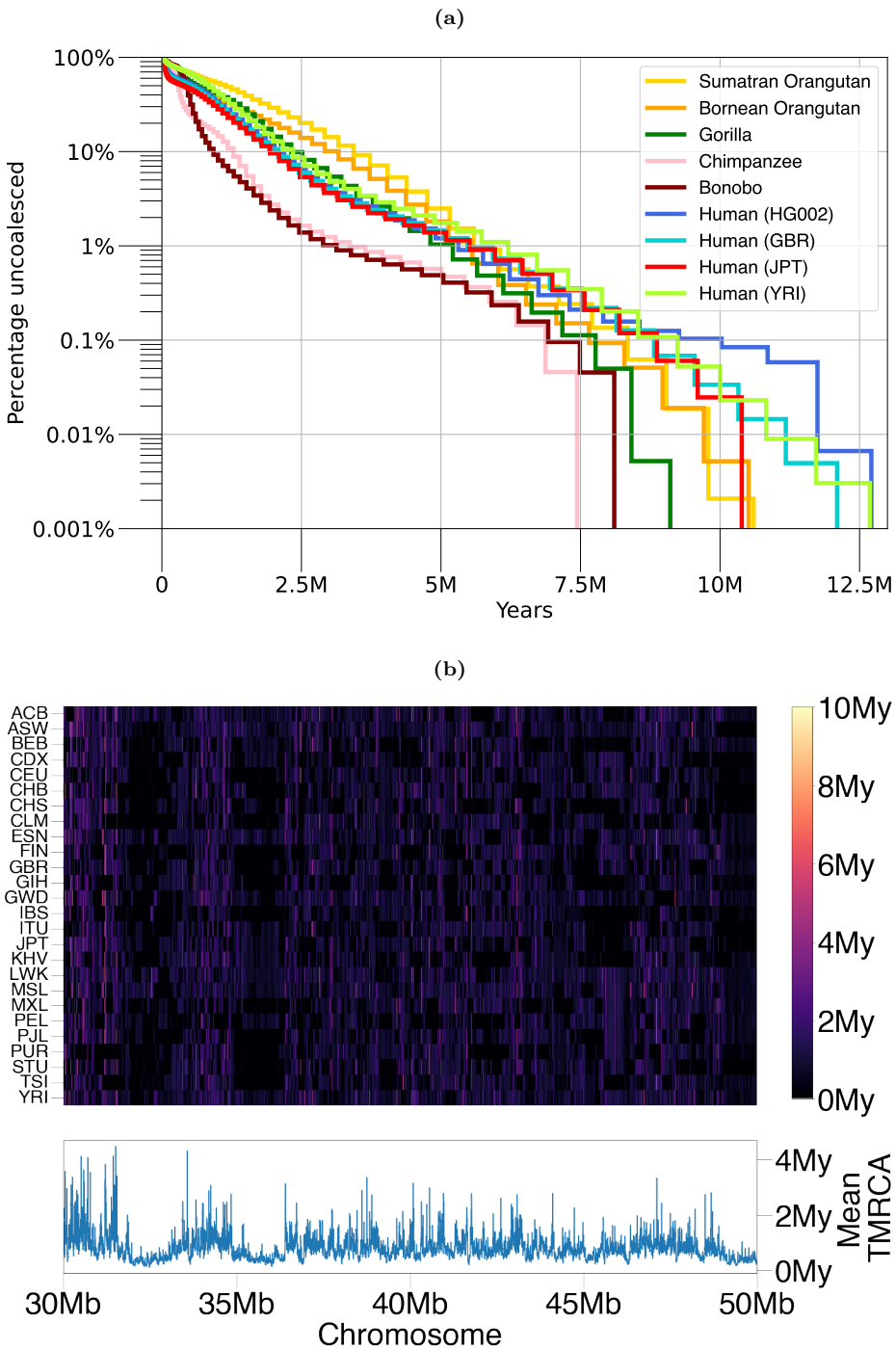

Figure 2: **(a)** Inferred fraction of uncoalesced genome as a function of time, for each sample from Figure 1. **(b)** (Top) Posterior mean TMRCAs from 26 individuals, each from a distinct population in the 1000 Genomes project, for 20Mb on chromosome 20. The bottom panel shows the mean of the top panel.

As more direct confirmation from the data, we inferred the fraction of the genome that has not yet coalesced as a function of time using PSMC decoding (see Methods). At least 1% of the human, gorilla, and orangutan genome is estimated not to have coalesced by 5 Mya, with approximately 0.5% remaining in the chimpanzee and bonobo genome at this time (Figure 2a). All of the non-human primates are estimated to

Population	Sample	1 Mya	2 Mya	3 Mya	4 Mya	5 Mya	6 Mya
ESN	HG03515	54,169	17,213	7,811	3,952	1,841	472
YRI	NA18488	54,486	17,229	7,668	3,841	1,752	426
MSL	HG03212	53,817	17,243	7,661	3,910	1,947	494
GWD	HG02568	54,493	16,893	7,358	3,667	1,721	397
ACB	HG01882	55,368	17,590	7,883	4,129	1,909	457
ASW	NA19625	54,974	17,395	7,798	3,984	1,852	439
LWK	NA19017	54,644	17,709	8,105	4,050	1,874	496

Table 1: The number of ancient non-contiguous segments inferred in different African populations from the 1000GP project. We count segments by looking for regions where the posterior probability of coalescence as old or older than the time given in each column is greater than 0.9.

have fully coalesced between 7.4 Mya and 11 Mya. Notably, however, even at 10 Mya an average of 0.06% of the human genome remains uncoalesced, with full coalescence not being reached until beyond 12.5 Mya in the YRI and HG002 samples.

In humans from the 1000GP, ancient regions tend to be shared among individuals (Figure 2b); e.g. positions with an inferred posterior mean TMRCA of over 3 Mya show an average correlation of r^2=0.21 across individuals (Figure A3). We note that further back in time, segments of shared ancestry become increasingly shorter because of longer exposure to recombination, and that these are scattered across the genome. For example we estimate that African individuals harbor roughly 4,000 non-contiguous segments whose coalescence time is older than 4 Mya, and roughly 400 segments whose coalescence time is older than 6 Mya (Table 1; see Methods). This number of distinct segments is sufficient to give robust coalescence estimates, consistent with the bootstrap results. We note that previous simulation studies [43, 44] and theoretical analysis [45] have shown that posterior mean estimates of true ancient coalescent times substantially older than the mean tend to be significantly downward biased, which may indicate an even older age for detected ancient regions. We repeated the analysis using an independent method, Relate [46], which does not rely on the PSMC assumptions and uses multiple sequences jointly; this also identifies sufficient coalescence material exists in this time period (see Methods and Figure A4).

3.3 Model violations and their effects

As with all statistical frameworks, the PSMC method relies on specific modelling assumptions that may not hold true in reality. Deviations from these assumptions could have varying impacts on inference across different temporal scales, and could potentially account for some or all of the observed ancient peak. We therefore considered how departures from the modelling assumptions could plausibly affect inference of population size history in the deep past.

To this end, we considered the following possible issues. First, repeats may manifest as an artifactual heterozygous signal, leading to long segments with high heterozygosity, which then leads to an apparent excess of ancient coalescence events. Second, variation in mutation rates across different genomic regions could result in some regions having a higher mutation density, which might be interpreted as an older TMRCA, while others with lower mutation density might be interpreted as having a younger TMRCA. This can lead to systematic biases in the inference of accurate evolutionary histories. Third, the human mutation rate is estimated to have slowed down over time [47, 48, 49, 50], while in PSMC it is assumed constant.

Fourth, variability in recombination rates across the genome (in particular at recombination hotspots) is also a departure from model assumptions [51, 52, 53, 54, 55]. Fifth, at longer timescales there may become a significant probability of recurrent mutations at the same site, which are not modelled in PSMC. Sixth, background selection, which has been estimated to be pervasive in humans [56, 57], has been shown to distort the inference of the coalescence rate [58, 59, 60]. Finally, balancing selection can also manifest as an excess in ancient coalescence events [61].

We conducted an extensive set of simulation studies to analyse how these violations may affect inference, which we describe fully in the Appendix. We concluded that it is unlikely any of these issues or violations could cause a false ancient peak of the type we see in real data. Notably, we observed that if there are large variations in the mutation rate across the genome, then false inflations of inferred $N(t)$ are observed. However, the degree of rate variation required to generate this is larger than the degree of this inferred in humans, and moreover their ancient inferred $N(t)$ has high variance, contrasting the low variance we have shown in humans in the ∼3-6 Mya peak.

4 Discussion

We have demonstrated the applicability of coalescent-based inference to examine deep evolutionary history in the hominin lineage. Specifically, we highlighted and extended the analysis of a peak in ancient population size history inferred by PSMC, which is reliably estimated from human, chimpanzee and bonobo data, but not in gorilla or orangutan data. The human, chimpanzee and bonobo peaks look qualitatively similar and span approximately the same time period, suggesting that they may reflect the same event prior to, or during, *Homo-Pan* (HP) speciation.

We discussed several ways in which the data may violate the model underlying our analysis, and concluded that none are likely to generate a false ancient peak of the type seen. Nevertheless, it may be useful to incorporate some of these as extensions to the PSMC model, modifying the transition and emission probabilities within the underlying HMM accordingly. Indeed, some work has already been done towards this goal [62, 63], and future modification could further address slowdown or genomic variability in mutation rate, recombination hotspots, and/or selection, so as to better understand demographic change and speciation across a variety of species.

A straightforward interpretation of the shared ancient peak in humans, chimpanzees and bonobos is that it reflects true changes in population size. Indeed, for humans, changes in inferred $N(t)$ since ∼80 kya have been associated with the out-of-Africa event followed by more recent recent population size increases [7, 19]. The ancient peak may similarly reflect ancient hominin population expansion and decline in the late Miocene and the Pliocene. To the extent that the same ancient peak exists in chimpanzees and bonobos, it may reflect common environmental conditions affecting similar evolutionary processes in the different species. Even if the ancestral population were structured, shared environmental changes may have contributed to changes in inferred $N(t)$.

However, there are other possible explanations. The separation and reconnection of local populations alters the amount of possible gene flow between them, which affects the inferred coalescence rate

[64, 65, 66, 67, 68]. In a structured population (defined as a population whose evolutionary history was not continuously panmictic) the coalescence rate decreases relative to a panmictic population of the same size, thus the effective population size (taken throughout this paper as the inverse of the coalescence rate) increases. Indeed, in the original PSMC publication Li and Durbin demonstrate that a pulse admixture event can generate a peak in effective population size inference [6], and in a more thorough analysis Mazet et al. show that the inferred human effective population size as inferred by PSMC - including the ancient peak - can be equally well obtained by a set of populations with varying migration rates between them without any changes in size [67]. Therefore, it is possible that the ancient peak is a signature of ancient population structure in some form.

One such possible structured event may be complex speciation between the HP lineages, in which there was some form of gene flow between the *Homo* and *Pan* lineages after their initial separation and prior to their final loss of ability to exchange genetic material. Previous publications have put forward evidence for this type of model [33, 34, 14, 35], although other analysis has favoured a model without any gene flow [36, 37]. Broadly, these studies employ two types of models when testing for historical structure. One class of model has a panmictic ancestral population cleanly split into two distinct populations that later exchange genes in a pulse gene flow event. The other assumes that after initial divergence there is some level of continuous gene flow before full separation. Innan and Watanabe proposed there was no strong support for continuous gene flow after initial divergence, and that a clean split best fit the data [36]. This result was reinforced by Yamamichi et al., whose maximum likelihood pulse model also favoured a clean split [37] (we note that their pulse model was slightly different in that after a period of isolation the two populations merged again into a panmictic population before later cleanly separating). Patterson et al. proposed that a pulse gene flow event did occur, because of differences in the estimated divergence time between the autosomes and X chromosome [33], although this approach has been criticised [69, 70, 37]; specifically it was argued that expected coalescent variance and lineage-specific differences in male mutational bias driven mutation could provide an alternative explanation for the observed signal [70]. Using a likelihood ratio test considering divergence times across the genome, Yang rejected the null hypothesis of an absence of gene flow (noting the their model is compatible with either a pulse or continuous gene flow scenario) [34]. With an HMM that explicitly tests for immediate or prolonged separation with a period of continuous gene flow, Mailund et al. proposed that a model with an extended period of gene flow better fits the data [14]. Galtier introduced Aphid, which distinguishes between discordant coalescent trees (ones that do not match the species trees) generated by ILS or gene flow [35]. Aphid leverages the fact that multispecies coalescent trees affected by gene flow tend to have shorter branches, and coalescent trees affected by incomplete lineage sorting longer branches, than the average coalescent tree. This information provided strong support for a model in which there was ancient gene flow between human, chimpanzee, and even gorilla, after the initial human-chimpanzee split (in this model the flow was via a pulse migration event).

We propose that a parsimonious explanation for the shared ancient peak in humans and chimpanzee-bonobo is that there was substantial shared genetic population structure in their ancestry around the time of HP separation, i.e. that this reflects some form of complex speciation. We do not provide an explicit model for this structure, e.g. pulse versus continuous, or a specific number of demes with specific gene flows, because the experience of the field is that such models are underconstrained. However we contend that there is qualitative evidence to support this proposal, independent of the simple suspicious coincidence of timing.

Notably, CoalHMM [71] and TRAILS [16], which assume a panmictic ancestral human-chimpanzee-bonobo population and a clean split, estimate the ancestral effective population size as 177,368 and 167,400, respectively. Both of these estimates are not well aligned with the PSMC inference in this period (Figure 1), and are more than 10 times the long-term effective population size of humans and chimpanzees (defining "long-term effective population size" as the value inferred from the observed heterozygosity using the equation $\theta = 4\bar{N}\mu$. In humans $\theta = 4\bar{N}_h\mu_h$ is roughly 0.0008, which assuming μ_h=1.25e-8 gives human long-term effective size $\bar{N}_h$=16,000; in chimpanzees, θ is roughly 0.001, which assuming μ_c=1.78e-8 gives chimpanzee long-term effective size $\bar{N}_c \approx$14,000). A structured ancestral population could explain the large CoalHMM and TRAILS estimates, as falsely assuming panmixia can inflate estimates of population size [67]. Moreover, Aphid [35], which provides strong support for human-chimpanzee-gorilla gene flow, estimates the size of the ancestral human-chimpanzee population to be of similar size to present-day human, chimpanzee and bonobo populations.

We note that there are believed to be other speciation/separation events in the human lineage for which we see no PSMC peak, for example divergences among early Homo lineages at 2-3 Mya, including those associated with *Homo habilis*. However, even if there was gene flow associated with these separations, a divergence event with gene flow need not always generate a peak in PSMC analyses. In a recent study, Cousins et al. simulated various structured models in which an ancestral population diverges into two distinct populations that later exchange genes via a pulse migration event [68]. They ran PSMC analyses on these simulations and found that the magnitude of the inferred peak is a function of the amount of admixture, the length of time between separation and reconnection, as well as the density of coalescence events during the structured period (Supplementary Figure 19 in [68]). In particular, with a human-inspired effective population size and mutation rate, if the length of time between the two admixing populations is less than 1 million years and the amount of admixture is less than 5% then there is no noticeable signal in the PSMC analysis. As an illustrative example of this effect in real data, consider the Neanderthal to non-African modern humans introgression: the amount of introgression from Neanderthals into non-Africans is estimated to have been on the order of 2% and to have occurred roughly ∼50 kya [10, 31], with the two populations having diverged ∼550kya. There is not a discernible PSMC peak associated with this gene flow event; rather, the human peak around 300-400 kya in Figure 1 is shared with Africans who are not expected to have had any Neanderthal admixture - that peak is better explained by deep structure in the hominin lineage [72, 68]. Conversely, however, Schur et al. propose that the Neanderthal introgression does subtly perturb the magnitude of the peak at ∼300 kya, though they do not propose the peak is a result of this event [73].

Divergence with gene flow events not necessarily generating PSMC peaks also plausibly explains why other reported admixtures do not have PSMC peaks associated with their reported timing. For example, de Manuel et al. estimate that 1% of the central chimpanzee genome descends from an introgression event less than 500 kya from a bonobo-related lineage [74]. Kuhlwilm et al. estimated that 0.9-4.8% of the bonobo's genome descends from an unsampled lineage that separated ∼3 Mya and introgressed ∼600 kya [75]. Nater et al. estimated that the Sumatran orangutan continously exchanged genes with the Tapanuli orangutan around 100 kya after their separation ∼3 Mya [76]. Pawar et al. estimated that the eastern gorillas derive ∼3% of their genome from an unsampled lineage that split from a common ancestor more than 3 Mya [77]. These estimated admixture events are of relatively small magnitude and generally do not produce discernible

PSMC peaks at the time of introgression, reinforcing that only sufficiently strong or prolonged gene flow is expected to generate detectable features in PSMC trajectories.

Mazet et al. [67], as well as Tournebize and Chikhi [78], have demonstrated that PSMC inference on an individual descended from a structured stem population (a set of sub-populations that are exchanging genes at some rate) can show peaks around the periods associated with gene flow [67, 78]. In light of this framework, the peaks observed in human, chimpanzee, and bonobo PSMC trajectories between roughly 3–6 Mya are consistent with descent from a structured ancestral population around that time, which is around the time of human-chimpanzee lineage separation. Assuming the eventual *Homo* and *Pan* lineages descended from different subpopulations of that structured ancestry, this scenario corresponds to our suggestion of complex speciation, in which the two lineages did not split instantaneously from a single panmictic population but instead emerged from different components of a structured ancestral population connected by migration, before complete reproductive isolation was established. In the long term, we can look forward to a better resolution of what happened in this important time for human evolution based on a combination of improved genetic and fossil evidence.

Acknowledgements

We thank Aylwyn Scally for constructive comments. T.C. was funded by a Wellcome Postgraduate Studentship 108864/B/15/Z, and R.D. and R.S. by Wellcome Investigator Award 207492/Z/17/Z. For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

References

[1] Brian Charlesworth. Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, 10(3):195–205, 2009.

[2] Xiaoming Liu and Yun-Xin Fu. Exploring population size changes using SNP frequency spectra. *Nature Genetics*, 47(5):555–559, 2015.

[3] William S DeWitt, Kameron Decker Harris, Aaron P Ragsdale, and Kelley Harris. Nonparametric coalescent inference of mutation spectrum history and demography. *Proceedings of the National Academy of Sciences*, 118(21):e2013798118, 2021.

[4] Enrique Santiago, Irene Novo, Antonio F Pardiñas, María Saura, Jinliang Wang, and Armando Caballero. Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. *Molecular Biology and Evolution*, 37(12):3642–3653, 2020.

[5] Enrique Santiago, Carlos Köpke, and Armando Caballero. Accounting for population structure and data quality in demographic inference with linkage disequilibrium methods. *Nature Communications*, 16(1):6054, 2025.

[6] Heng Li and Richard Durbin. Inference of human population history from individual whole-genome sequences. *Nature*, 475(7357):493–496, 2011.

[7] Stephan Schiffels and Richard Durbin. Inferring human population size and separation history from multiple genome sequences. *Nature Genetics*, 46(8):919–925, 2014.

[8] Jonathan Terhorst. Accelerated bayesian inference of population size history from recombining sequence data. *Nature Genetics*, 57(10):2570–2577, 2025.

[9] Matthias Meyer, Martin Kircher, Marie-Theres Gansauge, Heng Li, Fernando Racimo, Swapan Mallick, Joshua G Schraiber, Flora Jay, Kay Prüfer, Cesare De Filippo, et al. A high-coverage genome sequence from an archaic Denisovan individual. *Science*, 338(6104):222–226, 2012.

[10] Kay Prüfer, Fernando Racimo, Nick Patterson, Flora Jay, Sriram Sankararaman, Susanna Sawyer, Anja Heinze, Gabriel Renaud, Peter H Sudmant, Cesare De Filippo, et al. The complete genome sequence of a neanderthal from the Altai Mountains. *Nature*, 505(7481):43–49, 2014.

[11] Brian Villmoare, William H Kimbel, Chalachew Seyoum, Christopher J Campisano, Erin N DiMaggio, John Rowan, David R Braun, J Ramón Arrowsmith, and Kaye E Reed. Early Homo at 2.8 ma from ledi-geraru, afar, ethiopia. *Science*, 347(6228):1352–1355, 2015.

[12] Erin N DiMaggio, Christopher J Campisano, John Rowan, Guillaume Dupont-Nivet, Alan L Deino, Faysal Bibi, Margaret E Lewis, Antoine Souron, Dominique Garello, Lars Werdelin, et al. Late pliocene fossiliferous sedimentary record and the environmental context of early Homo from afar, ethiopia. *Science*, 347(6228):1355–1359, 2015.

[13] Jason E Lewis, Carol V Ward, William H Kimbel, Casey L Kidney, Frank H Brown, Rhonda L Quinn, John Rowan, Ignacio A Lazagabaster, William J Sanders, Meave G Leakey, et al. A 4.3-million-year-old australopithecus anamensis mandible from ileret, east turkana, kenya, and its paleoenvironmental context. *Journal of Human Evolution*, 194:103579, 2024.

[14] Thomas Mailund, Anders E Halager, Michael Westergaard, Julien Y Dutheil, Kasper Munch, Lars N Andersen, Gerton Lunter, Kay Prüfer, Aylwyn Scally, Asger Hobolth, et al. A new isolation with migration model along complete genomes infers very different divergence processes among closely related great ape species. *PLoS Genetics*, 8(12):e1003125, 2012.

[15] Asger Hobolth, Ole F Christensen, Thomas Mailund, and Mikkel H Schierup. Genomic relationships and speciation times of human, chimpanzee, and gorilla inferred from a coalescent hidden Markov model. *PLoS Genetics*, 3(2):e7, 2007.

[16] Iker Rivas-González, Mikkel H Schierup, John Wakeley, and Asger Hobolth. TRAILS: Tree reconstruction of ancestry using incomplete lineage sorting. *PLoS Genetics*, 20(2):e1010836, 2024.

[17] Sara Sheehan, Kelley Harris, and Yun S Song. Estimating variable effective population sizes from multiple genomes: a sequentially Markov conditional sampling distribution approach. *Genetics*, 194(3):647–662, 2013.

[18] Matthias Steinrücken, Joshua S Paul, and Yun S Song. A sequentially Markov conditional sampling distribution for structured populations with migration and recombination. *Theoretical Population Biology*, 87:51–61, 2013.

[19] Jack Kamm, Jonathan Terhorst, Richard Durbin, and Yun S Song. Efficiently inferring the demographic history of many populations with allele count data. *Journal of the American Statistical Association*, 115(531):1472–1487, 2020.

[20] Donna Henderson, Sha Zhu, Christopher B Cole, and Gerton Lunter. Demographic inference from multiple whole genomes using a particle filter for continuous Markov jump processes. *PLoS One*, 16(3):e0247647, 2021.

[21] Gilean AT McVean and Niall J Cardin. Approximating the coalescent with recombination. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1459):1387–1393, 2005.

[22] Paul Marjoram and Jeff D Wall. Fast "coalescent" simulation. *BMC genetics*, 7(1):1–9, 2006.

[23] Julien Y Dutheil, Ganesh Ganapathy, Asger Hobolth, Thomas Mailund, Marcy K Uyenoyama, and Mikkel H Schierup. Ancestral population genomics: the coalescent hidden Markov model approach. *Genetics*, 183(1):259–274, 2009.

[24] Javier Prado-Martinez, Peter H Sudmant, Jeffrey M Kidd, Heng Li, Joanna L Kelley, Belen Lorente-Galdos, Krishna R Veeramah, August E Woerner, Timothy D O'connor, Gabriel Santpere, et al. Great ape genetic diversity and population history. *Nature*, 499(7459):471–475, 2013.

[25] Marius F Maurstad, Siv Nam Khang Hoff, José Cerca, Mark Ravinet, Ian Bradbury, Kjetill S Jakobsen, Kim Præbel, and Sissel Jentoft. Reference genome bias in light of species-specific chromosomal reorganization and translocations. *Genome Biology*, 26(1):1–26, 2025.

[26] Aaron M Wenger, Paul Peluso, William J Rowell, Pi-Chuan Chang, Richard J Hall, Gregory T Concepcion, Jana Ebler, Arkarachai Fungtammasan, Alexey Kolesnikov, Nathan D Olson, et al. Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nature Biotechnology*, 37(10):1155–1162, 2019.

[27] Mikko Rautiainen, Sergey Nurk, Brian P Walenz, Glennis A Logsdon, David Porubsky, Arang Rhie, Evan E Eichler, Adam M Phillippy, and Sergey Koren. Telomere-to-telomere assembly of diploid chromosomes with Verkko. *Nature Biotechnology*, 41(10):1474–1482, 2023.

[28] Justin M Zook, David Catoe, Jennifer McDaniel, Lindsay Vang, Noah Spies, Arend Sidow, Ziming Weng, Yuling Liu, Christopher E Mason, Noah Alexander, et al. Extensive sequencing of seven human genomes to characterize benchmark reference materials. *Scientific Data*, 3(1):160025, 2016.

[29] Ting Wang, Lucinda Antonacci-Fulton, Kerstin Howe, Heather A Lawson, Julian K Lucas, Adam M Phillippy, Alice B Popejoy, Mobin Asri, Caryn Carson, Mark JP Chaisson, et al. The Human Pangenome Project: a global resource to map genomic diversity. *Nature*, 604(7906):437–446, 2022.

[30] DongAhn Yoo, Arang Rhie, Prajna Hebbar, Francesca Antonacci, Glennis A Logsdon, Steven J Solar, Dmitry Antipov, Brandon D Pickett, Yana Safonova, Francesco Montinaro, et al. Complete sequencing of ape genomes. *Nature*, 641(8062):401–418, 2025.

[31] Qiaomei Fu, Heng Li, Priya Moorjani, Flora Jay, Sergey M Slepchenko, Aleksei A Bondarev, Philip LF Johnson, Ayinuer Aximu-Petri, Kay Prüfer, Cesare De Filippo, et al. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature*, 514(7523):445–449, 2014.

[32] 1000 Genomes Project Consortium et al. A global reference for human genetic variation. *Nature*, 526(7571):68, 2015.

[33] Nick Patterson, Daniel J Richter, Sante Gnerre, Eric S Lander, and David Reich. Genetic evidence for complex speciation of humans and chimpanzees. *Nature*, 441(7097):1103–1108, 2006.

[34] Ziheng Yang. A likelihood ratio test of speciation with gene flow using genomic sequence data. *Genome biology and evolution*, 2:200–211, 2010.

[35] Nicolas Galtier. An approximate likelihood method reveals ancient gene flow between human, chimpanzee and gorilla. *Peer Community Journal*, 4, 2024.

[36] Hideki Innan and Hidemi Watanabe. The effect of gene flow on the coalescent time in the human-chimpanzee ancestral population. *Molecular Biology and Evolution*, 23(5):1040–1047, 2006.

[37] Masato Yamamichi, Jun Gojobori, and Hideki Innan. An autosomal analysis gives no genetic evidence for complex speciation of humans and chimpanzees. *Molecular Biology and Evolution*, 29(1):145–156, 2012.

[38] 1000 Genomes Project Consortium et al. A map of human genome variation from population scale sequencing. *Nature*, 467(7319):1061, 2010.

[39] 1000 Genomes Project Consortium et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*, 491(7422):56, 2012.

[40] Marta Byrska-Bishop, Uday S Evani, Xuefang Zhao, Anna O Basile, Haley J Abel, Allison A Regier, André Corvelo, Wayne E Clarke, Rajeeva Musunuri, Kshithija Nagulapalli, et al. High-coverage whole-genome sequencing of the expanded 1000 Genomes Project cohort including 602 trios. *Cell*, 185(18):3426–3440, 2022.

[41] Anna-Sapfo Malaspinas, Michael C Westaway, Craig Muller, Vitor C Sousa, Oscar Lao, Isabel Alves, Anders Bergström, Georgios Athanasiadis, Jade Y Cheng, Jacob E Crawford, et al. A genomic history of Aboriginal Australia. *Nature*, 538(7624):207–214, 2016.

[42] Leon Hilgers, Shenglin Liu, Axel Jensen, Thomas Brown, Trevor Cousins, Regev Schweiger, Katerina Guschanski, and Michael Hiller. Avoidable false psmc population size peaks occur across numerous studies. *bioRxiv*, pages 2024–06, 2024.

[43] Débora YC Brandt, Xinzhu Wei, Yun Deng, Andrew H Vaughn, and Rasmus Nielsen. Evaluation of methods for estimating coalescence times using ancestral recombination graphs. *Genetics*, 221(1):iyac044, 2022.

[44] Regev Schweiger and Richard Durbin. Ultrafast genome-wide inference of pairwise coalescence times. *Genome Research*, 33(7):1023–1031, 2023.

[45] Yun S Song. Lecture Notes on Computational and Mathematical Population Genetics. 2021.

[46] Leo Speidel, Marie Forest, Sinan Shi, and Simon R Myers. A method for genome-wide genealogy estimation for thousands of samples. *Nature Genetics*, 51(9):1321–1329, 2019.

[47] Priya Moorjani, Ziyue Gao, and Molly Przeworski. Human germline mutation and the erratic evolutionary clock. *PLoS Biology*, 14(10):e2000744, 2016.

[48] Aylwyn Scally and Richard Durbin. Revising the human mutation rate: implications for understanding human evolution. *Nature Reviews Genetics*, 13(10):745–753, 2012.

[49] V Yi Soojin. Morris Goodman's hominoid rate slowdown: the importance of being neutral. *Molecular Phylogenetics and Evolution*, 66(2):569–574, 2013.

[50] Manjusha Chintalapati and Priya Moorjani. Evolution of the mutation rate across primates. *Current Opinion in Genetics & Development*, 62:58–64, 2020.

[51] S Myers, C C A Spencer, A Auton, L Bottolo, C Freeman, P Donnelly, and G McVean. The distribution and causes of meiotic recombination in the human genome. *Biochem. Soc. Trans.*, 34(Pt 4):526–530, 2006.

[52] International HapMap Consortium. The International HapMap Project. *Nature*, 426(6968):789–796, 2003.

[53] Thomas D Petes. Meiotic recombination hot spots and cold spots. *Nature Reviews Genetics*, 2(5):360–369, 2001.

[54] KT Nishant, Chetan Kumar, and MRS Rao. Humhot: a database of human meiotic recombination hot spots. *Nucleic Acids Research*, 34(suppl_1):D25–D28, 2006.

[55] Laure Ségurel, Ellen Miranda Leffler, and Molly Przeworski. The case of the fickle fingers: how the PRDM9 zinc finger protein specifies meiotic recombination hotspots in humans. *PLoS Biology*, 9(12):e1001211, 2011.

[56] Graham McVicker, David Gordon, Colleen Davis, and Phil Green. Widespread genomic signatures of natural selection in hominid evolution. *PLoS Genetics*, 5(5):e1000471, 2009.

[57] David A Murphy, Eyal Elyashiv, Guy Amster, and Guy Sella. Broad-scale variation in human genetic diversity levels is predicted by purifying selection on coding and non-coding elements. *eLife*, 12:e76065, 2022.

[58] Daniel R Schrider, Alexander G Shanku, and Andrew D Kern. Effects of linked selective sweeps on demographic inference and model selection. *Genetics*, 204(3):1207–1223, 2016.

[59] Parul Johri, Kellen Riall, Hannes Becher, Laurent Excoffier, Brian Charlesworth, and Jeffrey D Jensen. The impact of purifying and background selection on the inference of population history: problems and prospects. *Molecular Biology and Evolution*, 38(7):2986–3003, 2021.

[60] Simon Boitard, Armando Arredondo, Lounès Chikhi, and Olivier Mazet. Heterogeneity in effective size across the genome: effects on the inverse instantaneous coalescence rate (IICR) and implications for demographic inference under linked selection. *Genetics*, 220(3):iyac008, 2022.

[61] Deborah Charlesworth. Balancing selection and its effects on sequences in nearby genome regions. *PLoS Genetics*, 2(4):e64, 2006.

[62] Trevor Cousins, Daniel Tabin, Nick Patterson, David Reich, and Arun Durvasula. Accurate inference of population history in the presence of background selection. *bioRxiv*, 2024.

[63] Gustavo V. Barroso, Nataša Puzović, and Julien Y Dutheil. Inference of recombination maps from a single pair of genomes and its application to ancient samples. *PLoS Genetics*, 15(11):e1008449, 2019.

[64] Sewall Wright. Isolation by distance. *Genetics*, 28(2):114, 1943.

[65] Masatoshi Nei and Naoyuki Takahata. Effective population size, genetic diversity, and coalescence time in subdivided populations. *Journal of Molecular Evolution*, 37:240–244, 1993.

[66] Olivier Mazet, Willy Rodríguez, and Lounès Chikhi. Demographic inference using genetic data from a single individual: Separating population size variation from population structure. *Theoretical Population Biology*, 104:46–58, 2015.

[67] Olivier Mazet, Willy Rodríguez, Simona Grusea, Simon Boitard, and Lounès Chikhi. On the importance of being structured: instantaneous coalescence rates and human evolution—lessons for ancestral population size inference? *Heredity*, 116(4):362–371, 2016.

[68] Trevor Cousins, Aylwyn Scally, and Richard Durbin. A structured coalescent model reveals deep ancestral structure shared by all modern humans. *Nature Genetics*, 57(4):856–864, 2025.

[69] John Wakeley. Complex speciation of humans and chimpanzees. *Nature*, 452(7184):E3–E4, 2008.

[70] Daven C Presgraves and V Yi Soojin. Doubts about complex speciation between humans and chimpanzees. *Trends in ecology & evolution*, 24(10):533–540, 2009.

[71] Iker Rivas-González, Marjolaine Rousselle, Fang Li, Long Zhou, Julien Y Dutheil, Kasper Munch, Yong Shao, Dongdong Wu, Mikkel H Schierup, and Guojie Zhang. Pervasive incomplete lineage sorting illuminates speciation and selection in primates. *Science*, 380(6648):eabn4409, 2023.

[72] Aaron P Ragsdale, Timothy D Weaver, Elizabeth G Atkinson, Eileen G Hoal, Marlo Möller, Brenna M Henn, and Simon Gravel. A weakly structured stem for human origins in Africa. *Nature*, pages 1–9, 2023.

[73] Vladimir Shchur, Débora YC Brandt, Anna Ilina, and Rasmus Nielsen. Estimating population split times and migration rates from historical effective population sizes. *Preprint at biorxiv.org/content/10.1101/2022.06.17.496540v1*, 2022.

[74] Marc De Manuel, Martin Kuhlwilm, Peter Frandsen, Vitor C Sousa, Tariq Desai, Javier Prado-Martinez, Jessica Hernandez-Rodriguez, Isabelle Dupanloup, Oscar Lao, Pille Hallast, et al. Chimpanzee genomic diversity reveals ancient admixture with bonobos. *Science*, 354(6311):477–481, 2016.

[75] Martin Kuhlwilm, Sojung Han, Vitor C Sousa, Laurent Excoffier, and Tomas Marques-Bonet. Ancient admixture from an extinct ape lineage into bonobos. *Nature Ecology & Evolution*, 3(6):957–965, 2019.

[76] Alexander Nater, Maja P Mattle-Greminger, Anton Nurcahyo, Matthew G Nowak, Marc De Manuel, Tariq Desai, Colin Groves, Marc Pybus, Tugce Bilgin Sonay, Christian Roos, et al. Morphometric, behavioral, and genomic evidence for a new orangutan species. *Current Biology*, 27(22):3487–3498, 2017.

[77] Harvinder Pawar, Aigerim Rymbekova, Sebastian Cuadros-Espinoza, Xin Huang, Marc De Manuel, Tom Van der Valk, Irene Lobon, Marina Alvarez-Estape, Marc Haber, Olga Dolgova, et al. Ghost admixture in eastern gorillas. *Nature Ecology & Evolution*, 7(9):1503–1514, 2023.

[78] Rémi Tournebize and Lounès Chikhi. Ignoring population structure in hominin evolutionary models can lead to the inference of spurious admixture events. *Nature Ecology & Evolution*, 9(2):225–236, 2025.

[79] Nancy F Hansen, Nathan Dwarshuis, Hyun Joo Ji, Arang Rhie, Hailey Loucks, Glennis A Logsdon, Mitchell R Vollger, Jessica M Storer, Juhyun Kim, Eleni Adam, et al. A complete diploid human genome benchmark for personalized genomics. *bioRxiv*, 2025.

[80] Heng Li, Jonathan M Bloom, Yossi Farjoun, Mark Fleharty, Laura Gauthier, Benjamin Neale, and Daniel MacArthur. A synthetic-diploid benchmark for accurate variant-calling evaluation. *Nature Methods*, 15(8):595–597, 2018.

[81] Gary Benson. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research*, 27(2):573–580, 1999.

[82] Aleksandr Morgulis, E Michael Gertz, Alejandro A Schäffer, and Richa Agarwala. A fast and symmetric DUST implementation to mask low-complexity DNA sequences. *Journal of Computational Biology*, 13(5):1028–1040, 2006.

[83] Kelley Harris and Rasmus Nielsen. Error-prone polymerase activity causes multinucleotide mutations in humans. *Genome Research*, 24(9):1445–1454, 2014.

[84] Søren Besenbacher, Patrick Sulem, Agnar Helgason, Hannes Helgason, Helgi Kristjansson, Aslaug Jonasdottir, Adalbjorg Jonasdottir, Olafur Th Magnusson, Unnur Thorsteinsdottir, Gisli Masson, et al. Multi-nucleotide de novo mutations in humans. *PLoS Genetics*, 12(11):e1006315, 2016.

[85] Heng Li. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, 34(18):3094–3100, 2018.

[86] Petr Danecek, James K Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin O Pollard, Andrew Whitwham, Thomas Keane, Shane A McCarthy, Robert M Davies, et al. Twelve years of samtools and bcftools. *Gigascience*, 10(2):giab008, 2021.

[87] Heng Li, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, and 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and samtools. *Bioinformatics*, 25(16):2078–2079, 2009.

[88] Heng Li. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21):2987–2993, 2011.

[89] Moisès Coll Macià, Laurits Skov, Benjamin Marco Peter, and Mikkel Heide Schierup. Different historical generation intervals in human populations inferred from neanderthal fragment lengths and mutation signatures. *Nature Communications*, 12(1):5317, 2021.

[90] Jared C Roach, Gustavo Glusman, Arian FA Smit, Chad D Huff, Robert Hubley, Paul T Shannon, Lee Rowen, Krishna P Pant, Nathan Goodman, Michael Bamshad, et al. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science*, 328(5978):636–639, 2010.

[91] Philip Awadalla, Julie Gauthier, Rachel A Myers, Ferran Casals, Fadi F Hamdan, Alexander R Griffing, Mélanie Côté, Edouard Henrion, Dan Spiegelman, Julien Tarabeux, et al. Direct measure of the de novo mutation rate in autism and schizophrenia cohorts. *The American Journal of Human Genetics*, 87(3):316–324, 2010.

[92] Raheleh Rahbari, Arthur Wuster, Sarah J Lindsay, Robert J Hardwick, Ludmil B Alexandrov, Saeed Al Turki, Anna Dominiczak, Andrew Morris, David Porteous, Blair Smith, et al. Timing, rates and spectra of human germline mutation. *Nature Genetics*, 48(2):126–133, 2016.

[93] Augustine Kong, Michael L Frigge, Gisli Masson, Soren Besenbacher, Patrick Sulem, Gisli Magnusson, Sigurjon A Gudjonsson, Asgeir Sigurdsson, Aslaug Jonasdottir, Adalbjorg Jonasdottir, et al. Rate of de novo mutations and the importance of father's age to disease risk. *Nature*, 488(7412):471–475, 2012.

[94] Jacob J Michaelson, Yujian Shi, Madhusudan Gujral, Hancheng Zheng, Dheeraj Malhotra, Xin Jin, Minghan Jian, Guangming Liu, Douglas Greer, Abhishek Bhandari, et al. Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell*, 151(7):1431–1442, 2012.

[95] 1000 Genomes Project Consortium et al. Variation in genome-wide mutation rates within and between human families. *Nature Genetics*, 43(7):712–714, 2011.

[96] Aylwyn Scally. The mutation rate in human evolution and demographic inference. *Current Opinion in Genetics & Development*, 41:36–43, 2016.

[97] Jerome Kelleher, Alison M Etheridge, and Gilean McVean. Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLoS Computational Biology*, 12(5):e1004842, 2016.

[98] Hans Ellegren, Nick GC Smith, and Matthew T Webster. Mutation rate variation in the mammalian genome. *Current opinion in genetics & development*, 13(6):562–568, 2003.

[99] Alan Hodgkinson and Adam Eyre-Walker. Variation in the mutation rate across mammalian genomes. *Nature reviews genetics*, 12(11):756–766, 2011.

[100] Thomas CA Smith, Peter F Arndt, and Adam Eyre-Walker. Large scale variation in the rate of germ-line de novo mutation, base composition, divergence and diversity in humans. *PLoS Genetics*, 14(3):e1007254, 2018.

[101] Laurent C Francioli, Paz P Polak, Amnon Koren, Androniki Menelaou, Sung Chun, Ivo Renkens, Genome of the Netherlands Consortium, Cornelia M van Duijn, Morris Swertz, Cisca Wijmenga, et al. Genome-wide patterns and properties of de novo mutations in humans. *Nature Genetics*, 47(7):822–826, 2015.

[102] Hákon Jónsson, Patrick Sulem, Birte Kehr, Snaedis Kristmundsdottir, Florian Zink, Eirikur Hjartarson, Marteinn T Hardarson, Kristjan E Hjorleifsson, Hannes P Eggertsson, Sigurjon Axel Gudjonsson, et al. Parental influence on human germline de novo mutations in 1,548 trios from iceland. *Nature*, 549(7673):519–522, 2017.

[103] Wendy SW Wong, Benjamin D Solomon, Dale L Bodian, Prachi Kothiyal, Greg Eley, Kathi C Huddleston, Robin Baker, Dzung C Thach, Ramaswamy K Iyer, Joseph G Vockley, et al. New observations on maternal age effect on germline de novo mutations. *Nature Communications*, 7(1):10486, 2016.

[104] Miguel Rodriguez-Galindo, Sònia Casillas, Donate Weghorn, and Antonio Barbadilla. Germline de novo mutation rates on exons versus introns in humans. *Nature Communications*, 11(1):3304, 2020.

[105] Richard R Hudson. Properties of a neutral allele model with intragenic recombination. *Theoretical population biology*, 23(2):183–201, 1983.

[106] S Myers, CCA Spencer, A Auton, L Bottolo, C Freeman, P Donnelly, and et G McVean. The distribution and causes of meiotic recombination in the human genome, 2006.

[107] Richard A Gibbs, John W Belmont, Paul Hardenbol, Thomas D Willis, Fuli L Yu, HM Yang, Lan-Yang Ch'ang, Wei Huang, Bin Liu, Yan Shen, et al. The international hapmap project. 2003.

[108] Simon Myers, Rory Bowden, Afidalina Tumian, Ronald E Bontrop, Colin Freeman, Tammie S MacFie, Gil McVean, and Peter Donnelly. Drive against hotspot motifs in primates implicates the prdm9 gene in meiotic recombination. *Science*, 327(5967):876–879, 2010.

[109] Emil D Parvanov, Petko M Petkov, and Kenneth Paigen. PRDM9 controls activation of mammalian recombination hotspots. *Science*, 327(5967):835–835, 2010.

[110] Kenneth Paigen and Petko M Petkov. PRDM9 and its role in genetic recombination. *Trends in Genetics*, 34(4):291–300, 2018.

[111] Corinne Grey, Frédéric Baudat, and Bernard de Massy. PRDM9, a driver of the genetic map. *PLoS Genetics*, 14(8):e1007479, 2018.

[112] International HapMap Consortium Altshuler David altshuler@ molbio. mgh. harvard. edu Donnelly Peter donnelly@ stats. ox. ac. uk. A haplotype map of the human genome. *Nature*, 437(7063):1299–1320, 2005.

[113] International HapMap Consortium et al. A second generation human haplotype map of over 3.1 million snps. *Nature*, 449(7164):851, 2007.

[114] TH Jukes. Evolution of protein molecules. *Mammalian Protein Metabolism*, 3, 1969.

[115] Brian Charlesworth, MT Morgan, and Deborah Charlesworth. The effect of deleterious mutations on neutral molecular variation. *Genetics*, 134(4):1289–1303, 1993.

[116] Magnus Nordborg, Brian Charlesworth, and Deborah Charlesworth. The effect of recombination on background selection. *Genetics Research*, 67(2):159–174, 1996.

[117] Parul Johri, Susanne P Pfeifer, and Jeffrey D Jensen. Developing an evolutionary baseline model for humans: jointly inferring purifying selection with population history. *Molecular Biology and Evolution*, 40(5):msad100, 2023.

[118] Jacob I Marsh and Parul Johri. Biases in arg-based inference of historical population size in populations experiencing selection. *Molecular Biology and Evolution*, page msae118, 2024.

[119] KL Bubb, D Bovee, D Buckley, E Haugen, M Kibukawa, M Paddock, A Palmieri, S Subramanian, Y Zhou, R Kaul, et al. Scan of human genome reveals no new loci under ancient balancing selection. *Genetics*, 173(4):2165–2177, 2006.

[120] Saurabh Asthana, Steffen Schmidt, and Shamil Sunyaev. A limited role for balancing selection. *Trends in Genetics*, 21(1):30–32, 2005.

[121] Felix M Key, João C Teixeira, Cesare de Filippo, and Aida M Andrés. Advantageous diversity maintained by balancing selection in humans. *Current Opinion in Genetics & Development*, 29:45–51, 2014.

[122] Bora E Baysal, Elizabeth C Lawrence, and Robert E Ferrell. Sequence variation in human succinate dehydrogenase genes: evidence for long-term balancing selection on sdha. *BMC biology*, 5:1–14, 2007.

[123] Aida M Andrés, Melissa J Hubisz, Amit Indap, Dara G Torgerson, Jeremiah D Degenhardt, Adam R Boyko, Ryan N Gutenkunst, Thomas J White, Eric D Green, Carlos D Bustamante, et al. Targets of balancing selection in the human genome. *Molecular Biology and Evolution*, 26(12):2755–2764, 2009.

[124] Ellen M Leffler, Ziyue Gao, Susanne Pfeifer, Laure Ségurel, Adam Auton, Oliver Venn, Rory Bowden, Ronald Bontrop, Jeffrey D Wall, Guy Sella, et al. Multiple instances of ancient balancing selection shared between humans and chimpanzees. *Science*, 339(6127):1578–1582, 2013.

[125] Radoje Drmanac, Andrew B Sparks, Matthew J Callow, Aaron L Halpern, Norman L Burns, Bahram G Kermani, Paolo Carnevali, Igor Nazarenko, Geoffrey B Nilsen, George Yeung, et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*, 327(5961):78–81, 2010.

[126] Matthew D Rasmussen, Melissa J Hubisz, Ilan Gronau, and Adam Siepel. Genome-wide inference of ancestral recombination graphs. *PLoS Genetics*, 10(5):e1004342, 2014.

[127] W James Kent, Charles W Sugnet, Terrence S Furey, Krishna M Roskin, Tom H Pringle, Alan M Zahler, and David Haussler. The human genome browser at ucsc. *Genome Research*, 12(6):996–1006, 2002.

5 Methods

5.1 Processing data

We downloaded the HG002 fasta assemblies from Hansen et al. [79]. We downloaded the chimpanzee (mPanTro3), bonobo (mPanPan1), gorilla (mGorGor1), Bornean orangutan (mPonPyg2) and Sumatran orangutan (mPonAbe1) diploid fasta assemblies from Yoo et al. [30]. For each species, we processed the data through the following pipeline:

1. We used dipcall [80] with default parameters to align the alternate assembly to the primary, call variants, and designate regions of poor quality which were subsequently masked.

2. We masked all indels as well as 5 base pairs upstream and downstream of each.

3. We detected short tandem repeats using Tandem Repeat Finder with default parameters [81] and masked these.

4. We detected regions of low complexity using sdust [82] with $w = 64$ and $t = 20$ and masked these.

5. To account for multi-nucleotide mutation events [83, 84] we masked SNPs if they were within 20 base pairs of another SNP.

We took high-coverage whole-genome-sequence CRAM files for one individual in each of the 26 populations from the 1000 Genomes project [40]. We randomly chose the following individuals from each population: ACB-HG01882, ASW-NA19625, BEB-HG03006, CDX-HG02373, CEU-NA12718, CHB-NA18530, CHS-HG00443, CLM-HG01250, ESN-HG03515, FIN-HG00266, GBR-HG00118, GIH-NA20845, GWD-HG02568, IBS-HG01783, ITU-HG03977, JPT-NA18939, KHV-HG02113, LWK-NA19017, MSL-HG03212, MXL-NA19648, PEL-HG02285, PJL-HG03234, PUR-HG01171, STU-HG03753, TSI-NA20752, YRI-NA18488. We realigned these to the primary HG002 assembly using minimap2 [85]. The CRAM files were converted to BAM and indexed with samtools [86, 87]. The genotype likelihoods were calculated with bcftools mpileup [88] by skipping alignments with mapping quality less than 20 and setting the coefficient for downgrading mapping quality to 50. All indels were excluded. Variants were then designated as uncallable if the minimum mapping quality was less than 20, the minimum consensus quality was less than 20, or the coverage was less than half or more than double the mean coverage. Finally, we masked low complexity regions as detected from sdust as well as short tandem repeats as detected with Tandem Repeat Finder, as above for the genome assemblies. Uncallable positions were labelled as missing data in the input to the MSMC2 HMM.

5.2 Mutation rates and generation times for humans and primates

In humans we set the generation time as 29 years [89] and the mutation rate per generation per base pair as μ=1.25e-8 [90, 91, 92, 93, 94, 95, 96]. The rates we set for the other great apes are taken from [47, 50]. In chimpanzees and bonobos we set μ=1.78e-8 and generation time equal to 24 years. For Gorillas we set μ=1.42e-8 and generation time equal to 19 years. For Orangutans we set μ=2.03e-8 and generation time equal to 27 years.

5.3 PSMC analysis

When we discuss PSMC in the main text, we refer to the implementation of the pairwise SMC' [22] model as in MSMC2 [41]. We used 64 time intervals and enabled the parameters for the effective population size in each to be inferred freely. This contrasts with the default settings of previous implementations of the algorithm that force adjacent intervals to be the same, which has been shown to lead to fitting problems [42]. For the initial value for the ratio of mutation rate to recombination rate we used the default value of 4, and the recombination rate was set to be inferred as part of the EM algorithm, which was iterated for 40 iterations. For all analyses we discretised time into D=64 non-overlapping intervals, where the i-th boundary is defined by

$$\tau_i = \omega \cdot e^{\left(\frac{i}{D}\alpha - 1\right)} \quad \text{for } i \in \{1, \ldots, D\} \tag{1}$$

where $\tau_0 = 0$, $\alpha = \log\left(1 + \frac{T_{\max}}{\omega}\right)$, and ω and $T_{\max}$ are parameters that control the spread of time interval boundaries that are determined by MSMC2.

5.4 PSMC decoding

Using the inferred effective population size parameters, scaled recombination rate, and scaled mutation rate (θ), we used the PSMC decoding to get a posterior probability of coalescence at every position. To calculate the amount of uncoalesced genome over time, we averaged the posterior probability distributions across all

positions and calculated the empirical survival function (Figure 2). The model estimates time in coalescent units (Equation (1)), where one unit corresponds to $2N$ generations for a diploid population. To convert these times to years, we first estimate long-term effective population size $\bar{N}$ from the estimated scaled mutation rate using $\theta = 4\bar{N}\mu$, giving $\bar{N} = \theta/(4\mu)$, which uses the unscaled mutation rate μ. We then convert each inferred time point t by multiplying by $2\bar{N}$ to obtain time in generations, and finally multiply by the generation time g (in years) to obtain time in years. Equivalently, substituting for $\bar{N}$, this yields $t \cdot \frac{\theta}{2\mu} \cdot g$. To get a point estimate from PSMC (Figure 2b), we take the posterior mean at each position using the midpoint of the time interval boundaries.

To calculate the number of ancient non-contiguous segments inferred in different African populations in the 1000GP (Table 1), we looked for regions where the posterior probability of coalescence older than 2 Mya, 3 Mya or 4 Mya, respectively, is greater than 0.9. We then only counted non-contiguous segments, which are all segments that are not adjacent.

5.5 Non-parametric bootstrapping

We can resample the data with replacement to get an estimate of the uncertainty of the parameters inferred by PSMC. For a given individual, for each chromosome, we break the sequence up into segments of 5Mb, then reconstruct a new sequence by uniformly sampling with replacement from all 5Mb segments. The sampled segments are then concatenated together to form a contiguous sequence, on which PSMC inference is performed using 40 iterations. This procedure is repeated 20 times for each T2T sample in Figure 1.

5.6 Relate

Relate [46] is a method that estimates a genome wide ancestral recombination graph for numerous sequences. Rather than using the SMC to probabilistically infer genealogies, for each position it builds a distance matrix to describe the level of similarity between sequences, from which a rooted binary tree is constructed. The branch lengths of these trees are then estimated using an MCMC algorithm. We ran Relate on all samples in the 1000GP for GBR, CHB, and YRI. We randomly selected 10 diploid sequences from each population. To calculate the fraction of uncoalesced genome we discretised time into intervals of 500 generations and computed a histogram based on Relate's ARG. The inferred fraction of uncoalesced genome at 5 Mya was roughly 0.3%, for each sample in each population.

5.7 Simulation of ancient asymmetric gene flow

Using msprime [97] , we simulated 2Gb of sequence for a human and chimpanzee according to the following demographic model. A panmictic ancestral population of size 20,000 splits into proto1 and proto2 at time 6 Mya, each of size 20,000. At 3 Mya, human forms as a 80:20% mix of proto1 and proto2, respectively, and chimpanzee forms as a 50:50% mix of them. Both human and chimpanzee are of constant size 20,000 from 3 Mya to present. We used a generation time of 30 years and mutation rate of 1.25e-08. We ran MSMC2 on these simulated sequences as described in Section 5.3.

5.8 Reproducibility

Key scripts that were used for analysis have been uploaded to `https://github.com/trevorcousins/twin_peaks`.

6 Appendix

6.1 Figures

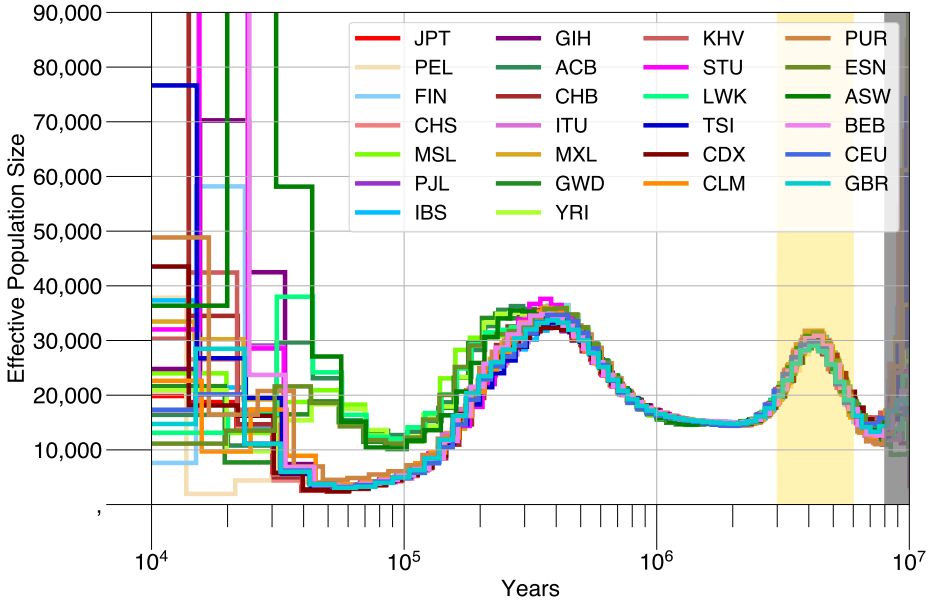

Figure A1: Inference from PSMC on 26 individuals from the 1000 Genomes project, each from a distinct population.

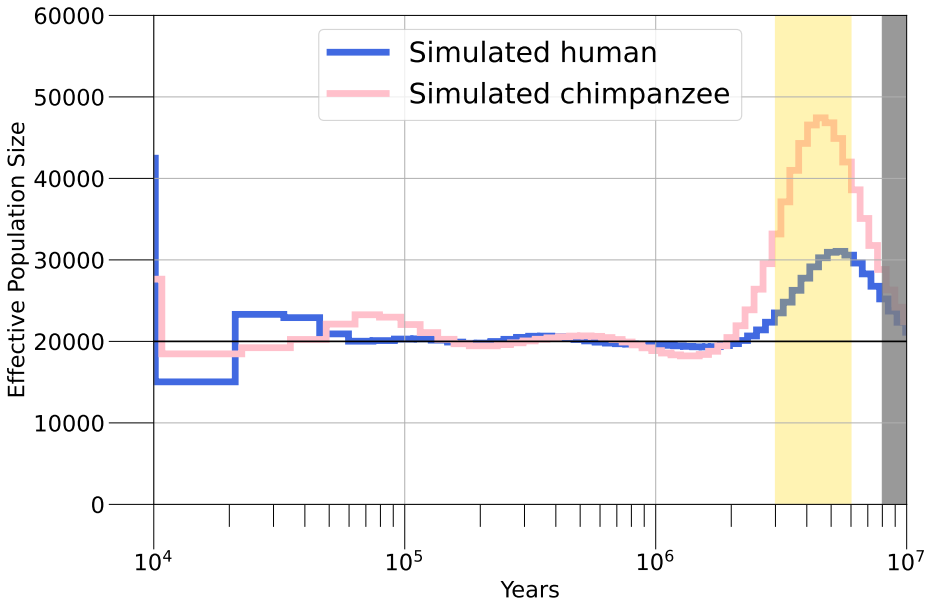

Figure A2: PSMC inference on 2Gb of simulated human and chimpanzee sequence. In this simulation, at 6 Mya a panmictic ancestral population of size 20k splits into proto1 and proto2 (each of size 20k). At 3 Mya, human forms as a 80:20 mix of proto1 and proto2, respectively, and chimpanzee forms as a 50:50 mix of proto1 and proto2.

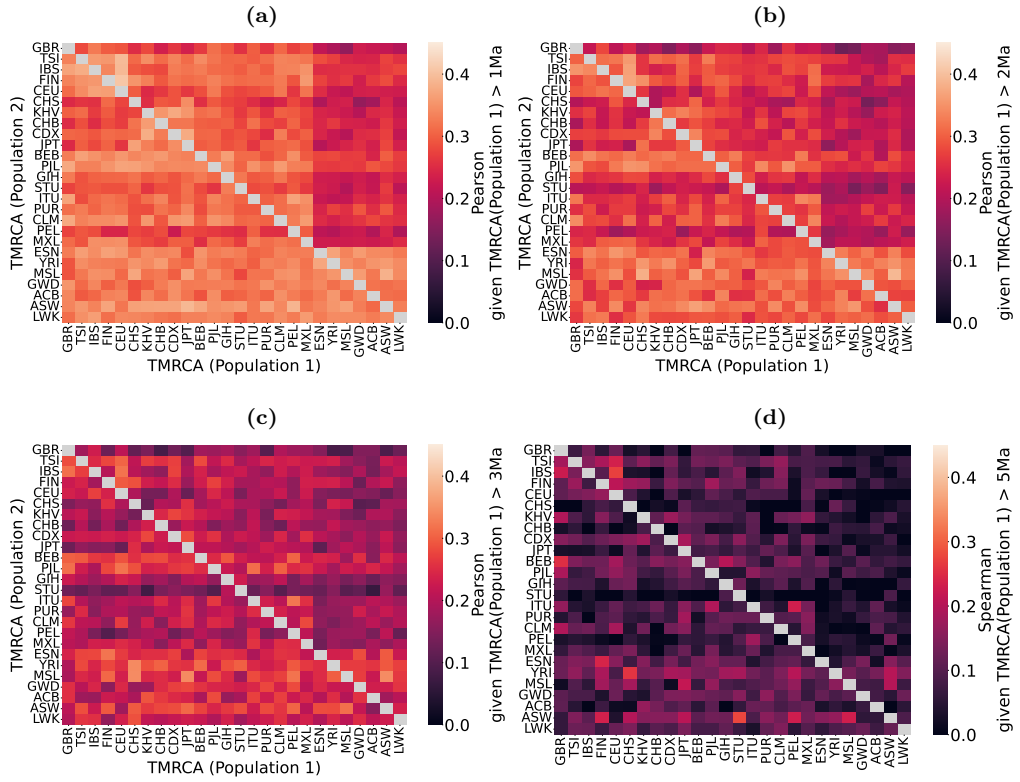

Figure A3: The Pearson correlation between inferred TMRCAs, for all pairwise combinations of two populations from the 1000GP. We calculated these for chromosome 1. To obtain point estimates at each position, we used the posterior mean. For each possible population pair, we look at positions where population 1 (x-axis) is larger than t and correlate these with the TMRCAs from population 2 (y-axis; note these are not necessarily larger than t, so the matrices are not symmetric). The value of t for **(a)**, **(b)**, **(c)**, and **(d)** is 1 Mya, 2 Mya, 3 Mya and 5 Mya, respectively. The average Pearson r^2 for (a), (b), (c), and (d) is 0.3, 0.26, 0.21 and 0.11, respectively, and all correlations have p-value<0.05 except for 72 combinations from (d).

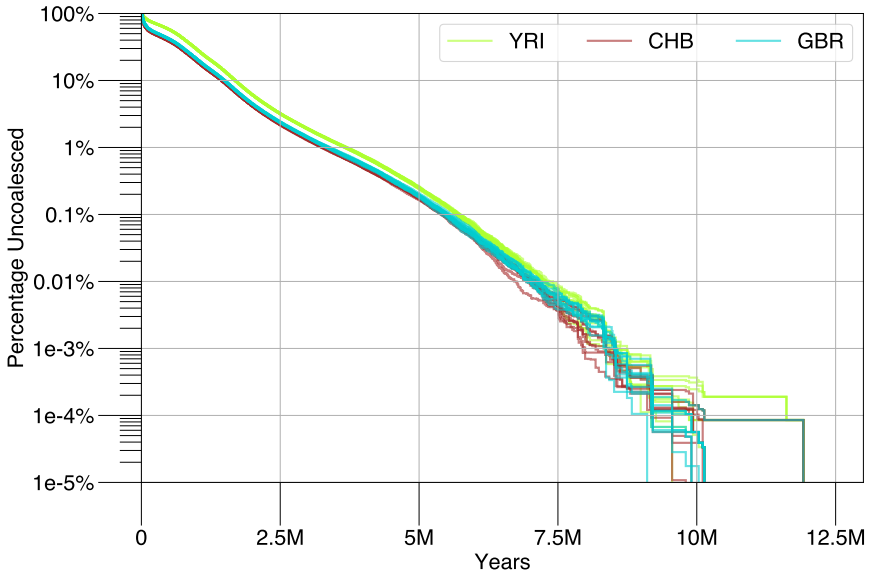

Figure A4: The Relate inferred fraction of uncoalesced genome as a function of time, for 10 diploid samples in the YRI, GBR and CHB populations from 1000GP.

6.2 Analysis of PSMC model violations

In this section, we analyse the effect of various model mispecifications on the inference from PSMC. We also analyse how the inferred coalescence times across the genome associate with various annotations, including for example repeat content, distance to coding sequence, strength of background selection, and recombination rate.

6.2.1 Repeats as a cause of artifactual heterozygous signal

As noted in previous work [6], false heterozygotes caused by repeated regions or segmental duplications may lead to excessively long segments with high heterozygosity, which may lead to an excess of ancient coalescent events. To examine whether this might artefactually be creating the observed signal, we analysed the fraction of the genome in repeated regions in the 1000GP data, and stratified this by inferred TMRCA. We observed that the repeat content increases as a function of TMRCA (regression slope $\sim 2\%$), and that the effect is significant (Pearson's $r^2 \approx 0.63$, p-value<0.005; see Methods and Figure A5a). A similar analysis with segmental duplications revealed no significant correlation (Figure A5b). We re-ran our analysis with repeated regions and segmental duplications masked out, and observed that the ancient peak is still present, although when repeats are removed it exhibits less prominence (Figure A6, left and middle panel, respectively). Thus we conclude that repeated regions or segmental duplications are unlikely to be the cause of the ancient peak. We further note that the ancient peak exists in the T2T data for which repeats have been filtered explicitly (see Methods).

6.2.2 Variable mutation rate across the genome

The PSMC model assumes a constant mutation rate throughout the genome, despite ample evidence that the mutation rate varies significantly by genomic location [98, 99]. Unfortunately, variation in mutation rates across different genomic regions can lead to systematic biases in the estimation of TMRCAs. This is because

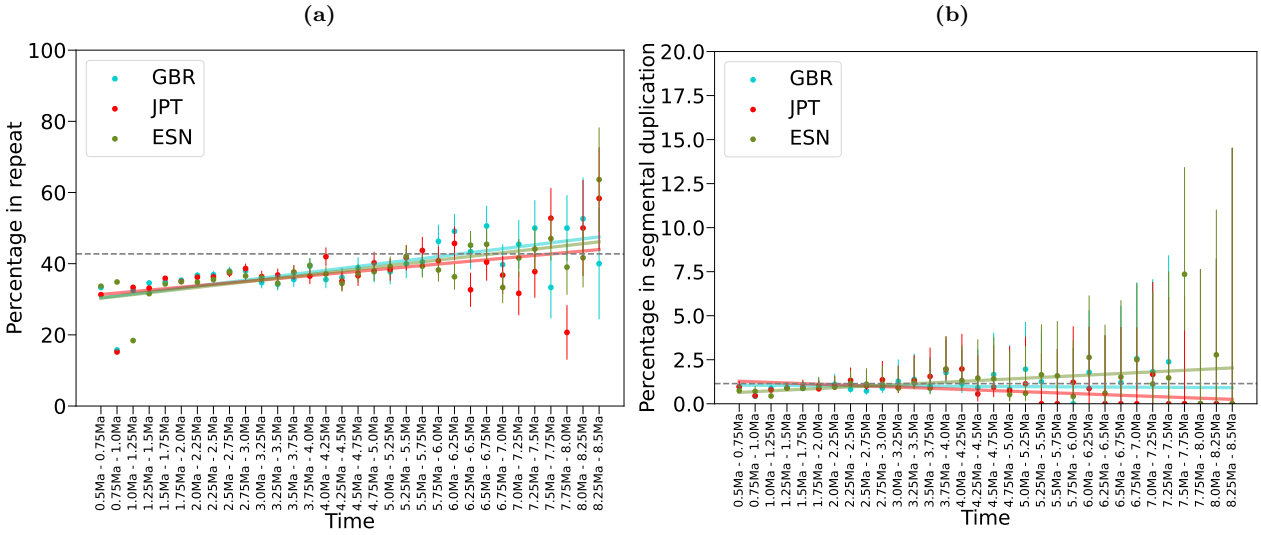

Figure A5: The relationship between the posterior mean TMRCA and the repeat content (a) or segmental duplications (b), on chromosome 1 for one individual each from GBR, JPT, and ESN in the 1000GP. (a) The reported Pearson r^2 for each individual is 0.73, 0.48, 0.69, with p-value<0.005 for all. The grey, dashed line indicates the chromosome wide average repeat content. (b) The reported Pearson r^2 for GBR and ESN is not significant, though for JPT it is -0.5 with p-value 0.003.

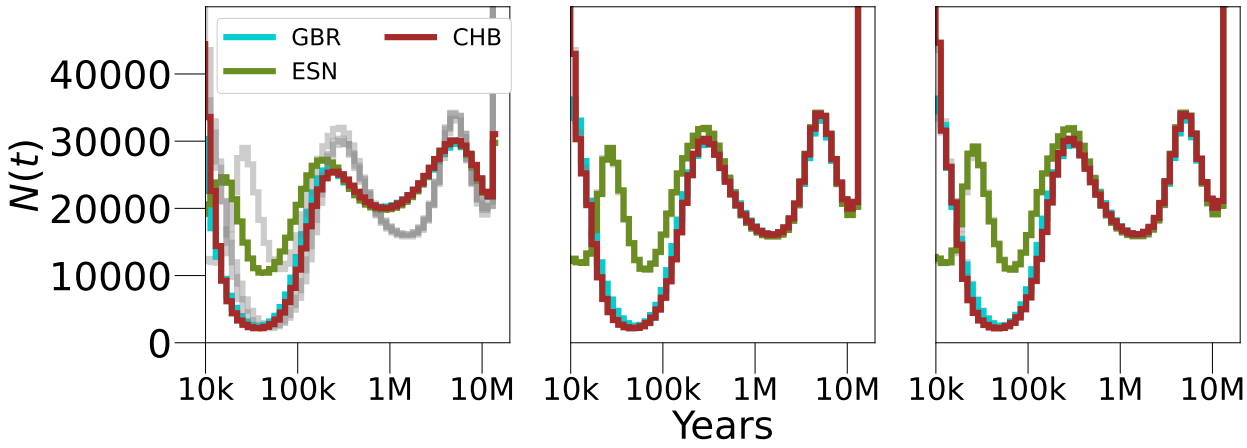

Figure A6: Running PSMC inference on human samples after masking out repeated regions (a), segmental duplications (b), and CpG islands (c). The original, unmasked results are shown in grey.

regions with faster mutation rates will have more mutations than otherwise, which PSMC will interpret as coming from a more ancient TMRCA, and vice versa. Consequently, overestimation and underestimation of TMRCAs can occur, which can confound efforts to infer accurate evolutionary histories.

To test this effect, we performed a series of simulations of a spatially varying mutation rate that changes according to a gamma distribution, in line with the models proposed in [100]. We generated a mutation rate map as follows. First, we divided the genome into disjoint intervals, whose length is exponentially distributed according to a rate parameter, CR. Then, for each interval, we drew a mutation rate from a

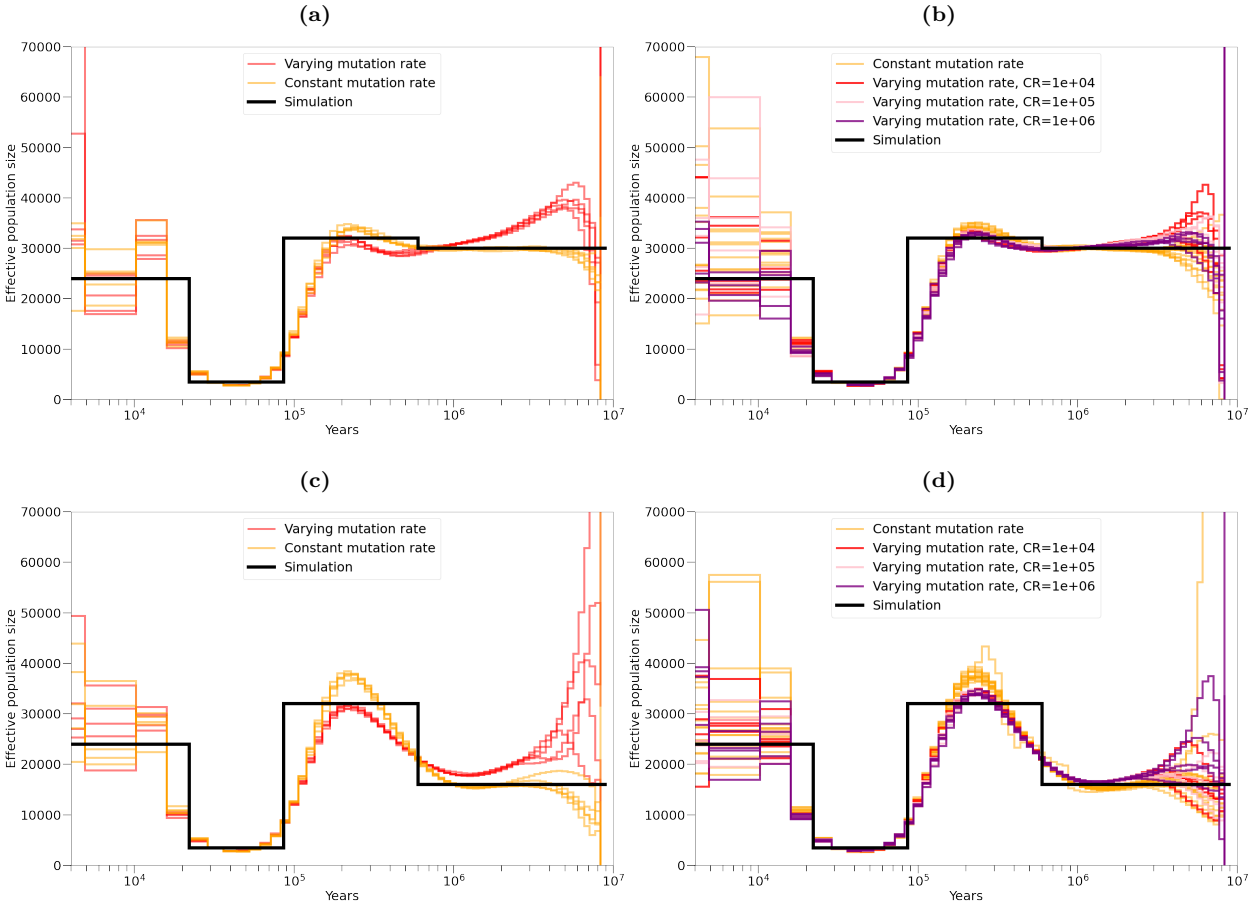

Figure A7: PSMC inference on simulations with a mutation rate that varies across the genome. The mutation rate changes at disjoint intervals, whose length is exponentially distributed with rate CR. In each interval, the mutation rate is drawn from gamma distribution with mean 1.25e-8 and a coefficient of variation v. The red, pink and purple lines show inference from PSMC on a simulation where the mutation rate varies, and the gold lines show inference from PSMC on a simulation where the mutation rate is constant. **(a)** and **(c)** $v = 0.25$ and CR =10kb. **(b)** and **(d)** $v = 0.15$ and CR=10kb (red), 100kb (pink), and 1Mb (purple).

gamma distribution whose mean is fixed to be 1.25e-8 [90, 91, 92, 93, 94, 95] with a coefficient of variation, v. We simulated over $CR \in \{10\text{kb}, 100\text{kb}, 1\text{Mb}\}$ and $v \in \{0.15, 0.25\}$, covering the range of values of v in reference [100]. We simulated diploid genomes according to a coarse piecewise population size trajectory using msprime [97], then generated pointwise mutations according to the mutation rate map. We generated 5 replicates of this simulation and then inferred the population size history using PSMC. Finally, as a control, we generated an additional 5 replicates using the same procedure, but with a constant mutation rate of 1.25e-8, and infer $N(t)$ using PSMC as well.

This simulation study shows that if the variation in mutation rate is strong enough (v=0.15), and the rate of change along the genome is fast enough (CR=10kb), PSMC detects a spurious peak in ancient time (Figure A7a). However, we note that qualitatively, the false peaks across different replicates vary more in location and height than in the human inference, for which the ancient peaks are well aligned (Figure 1a). Additionally, this simulation used v=0.25, which is larger than mutation rate model fitted by [100] for the

datasets in [101, 102] (0.18 and 0.15, respectively) but smaller than in [103] (0.27). In a simulation with v=0.15, we do not consistently observe a second peak (Figure A7b) for any value of L. In a simulation with lower ancient $N(t)$, v=0.25 does generate an artificial inflation of ancient $N(t)$, but not a discernible peak (Figure A7c). A simulation with $v = 0.15$ and L=1Mb again can generate a false peak, though it is not consistent across replicates (purple line, Figure A7d); CR=10kb or CR=100kb with v=0.15 does not seem to have much of an effect as the inference is similar to the constant mutation rate inference.

If the ancient peak was an artefact caused by variable mutation rates, we would expect to see a correlation between the mutation rate and the inferred TMRCA. To test this, we used de-novo mutation abundance counts obtained from trios [104] to generate a pointwise mutation rate map that depends on the local trinucleotide context (see Methods). The generated mutation map is positively correlated with SNP density as reported in 1000GP's dbSNP (r^2=0.09, p-value<1e-100, averaged across non-overlapping windows of 100bp), indicating positions with higher inferred rates have an elevated probability of mutations. We compared the variable mutation rates with the maximum inferred TMRCA across the 26 individuals in 1000GP, evaluated every 1kb. We observed a statistically insignificant Pearson's correlation (r=0.002, p-value=0.29) and a significant negative Spearman's correlation (rho=-0.04, p-value=3.7e-93) between TMRCA and mutation rate map. This suggests variable mutation rates do not play a significant role in causing the ancient peak.

Finally, we explored the role of CpG islands in affecting $N(t)$ inference. CpG islands are regions of DNA characterised by a high frequency of cytosine-guanine dinucleotides. They are often found near the transcription start site of genes and are associated with gene regulation. Due to their regulatory role, CpG islands have fewer mutations than expected by their high GC content, and therefore may contribute to mutation rate variability along the genome. The fraction of the genome in a CpG island is 0.007, and we stratify this by inferred TMRCA in Figure A8. The reported Pearson association between TMRCA and the fraction of regions in a CpG island is negative, though often this is not significant. Moreover in ancient time the CpG fraction seems to increase; to test the effect on inference we masked out CpG islands and inferred $N(t)$. We still observed an ancient peak (Figure A6, right), and thus concluded that CpG islands are not its cause.

6.2.3 Variable mutation rate through time

The PSMC model also assumes a constant mutation rate across past generations. While this is a sufficient approximation to allow inference in the last ∼50 kya [31], it may not hold in more ancient times. Indeed, it has been suggested that the mutation rate has changed over the course of human evolution, also referred to as the "hominoid rate slowdown" [47, 48, 49, 50]. This is based on the observation that the yearly mutation rate estimated from human pedigree studies is almost half the rate inferred by considering observed differences between human and chimpanzee genomes. The effect of a possible slowdown on PSMC estimates is unclear.

To this end, we simulated 10 replicates of a diploid genome, arising from a constant population size, and a mutation rate that is fixed at 2.5e-8 at 10 Mya then decreases to 1.25e-8 at present time (Figure A9a). To simulate a temporally changing mutation rate, we discretised time into D=64 segments with 65 time interval boundaries according to equation 1, and set the changes in mutation rate as piecewise constant in these intervals, $\mu = [\mu_1, \ldots, \mu_{64}]$. We simulate the coalescent process using msprime [97] with Hudson's

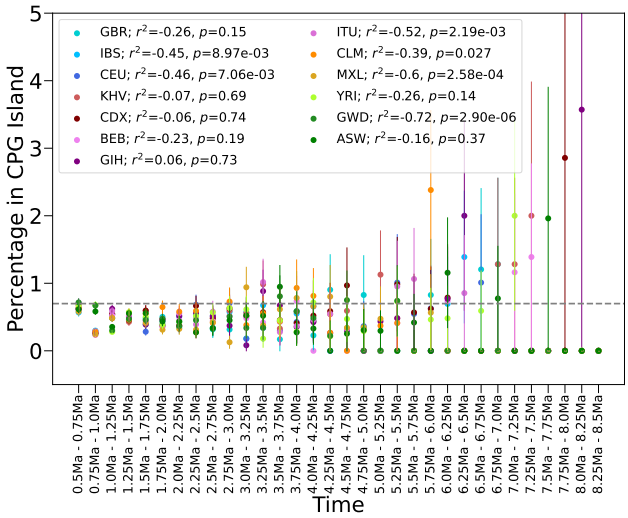

Figure A8: The correlation between the posterior mean TMRCA and the fraction of regions in a CpG island, on chromosome 1 for one individual from GBR, JPT, and ESN in the 1000GP. The Pearson r^2 and p-value are shown in the legend, although the relationship appears to be non-linear, and dominated by noise in ancient TMRCA bins. The dashed line indicates the genome-wide fraction of regions in a CpG island.

model of recombination [105], then utilise the memoryless property of the exponential distribution to add mutations at each position. With a coalescence time of t where $\tau_i \leq t < \tau_{i+1}$, measuring time in generations the probability of at least one mutation arising on either lineage is:

$$P(\text{At least one mutation arises in [0,t]}) = 1 - \exp\left(-4N\left(\sum_{j=1}^{i-1}\mu_j\Delta_j + \mu_i(t-\tau_i)\right)\right) \tag{2}$$

where $\Delta_j = \tau_j - \tau_{j-1}$.

We used PSMC to infer an $N(t)$ curve, and compared the results to $N(t)$ curves inferred from simulations with a fixed mutation rate of 1.25e-8. We observed that PSMC infers an increasingly inflated $N(t)$ in ancient times (Figure A9b). With the same mutation model, we also simulated changes in $N(t)$ similar to those inferred in the ESN (Figure 1). Again, we observe the inference of $N(t)$ from PSMC is increasingly inflated in ancient times, though the general shape of the trajectory is recovered (Figure A9c). We note that if the changes in mutation rate are known, we can simply scale the inference appropriately to correct the error (Figure A9d and e). Even though changes in the mutation rate over time can affect the estimation of $N(t)$, it is unlikely that the observed ancient peak in humans is due to this, as producing a peak as an artefact of changing mutation rates would require several rapid and severe mutation rate fluctuations, which seems unlikely. We conclude that mutation rate variation through time is not a likely cause of an ancient peak, though we note that this may contribute to the differences in ancient $N(t)$ magnitudes in humans and chimpanzees (Figures 1a and b).

6.2.4 Variable recombination rate across the genome

The PSMC model also assumes a constant recombination rate across the genome, and uniformly through all past generations. However, recombination rate varies along the genome, to a much greater extent than muta-

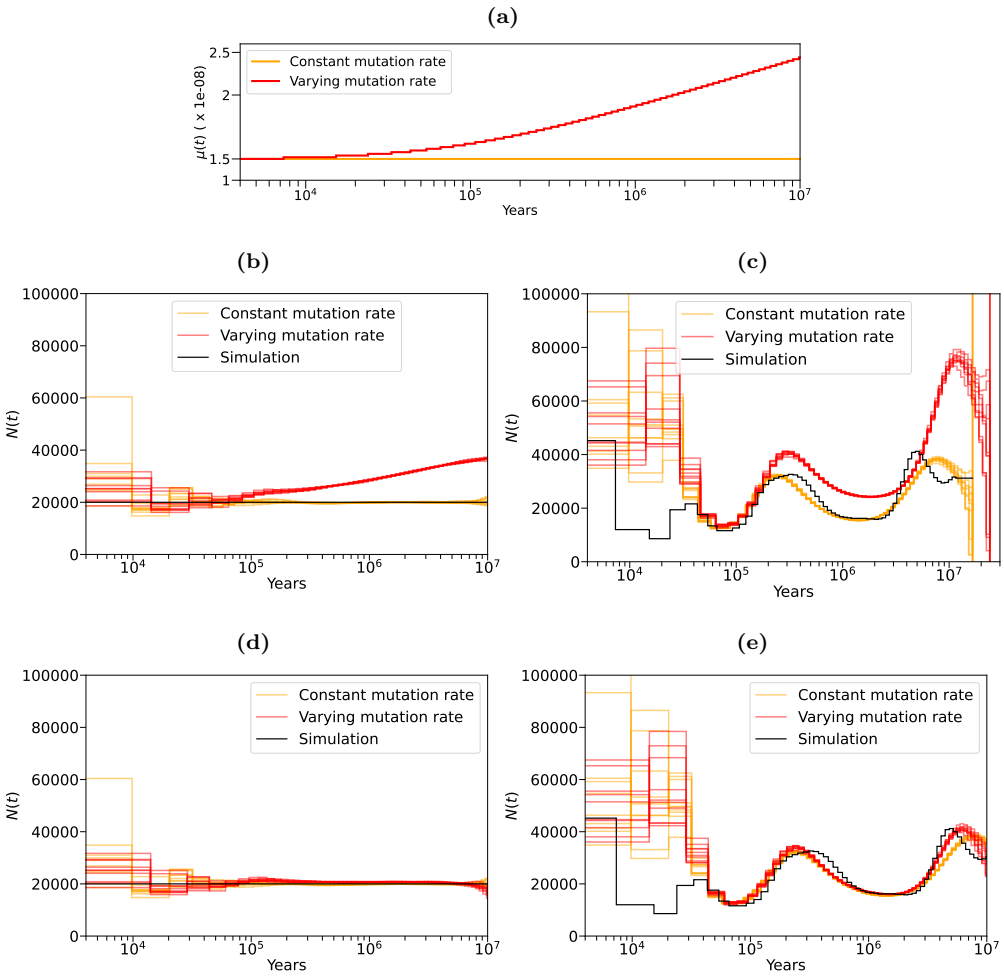

Figure A9: PSMC inference on simulations with a mutation rate that varies through time. **(a)** The mutation rate model, $\mu(t)$: at 10 Mya the mutation rate is 2.5e-8, then it slows down to 1.25e-8 at present. **(b)** Inference from PSMC on a simulation with constant population size, where the mutation rate changes through time according to the model in (a) in red, and a constant mutation rate (1.25e-8) in gold. **(c)** The same simulation as in (b) but with a population undergoing size changes. **(d)** and **(e)** are the same as (b) and (c), respectively, although the inferred $N(t)$ on the simulation with a varying mutation rate (red) has been divided through by $\mu(t)$ (red line in (a)).

tiona rate, with high rates at recombination hotspots and lower rates in different regions [106, 107, 53, 54, 55]. Looking back in time, the landscape of recombination is known to transform on a time scale of a few hundred thousand years [108, 109, 110, 111]. It is thus unclear how recombination rate variation affects $N(t)$ inference.

Previous work [6] has demonstrated that $N(t)$ inference in PSMC is robust even under simulations including recombination hotspots. As additional confirmation, we tested the effect of mis-specifying the recombination rate. Denote the simulated scaled recombination rate as ρ; we fix PSMC's recombination rate $\hat{\rho}$ as $0.2\rho, 0.5\rho, 1\rho, 2\rho$, or 5ρ, and infer $N(t)$ (Figure A10a). We observed that PSMC is able to recover $N(t)$ relatively accurately for all values of the recombination rate except $\hat{\rho}/\rho$ equal to 5. Finally, we saw a significant correlation between the local recombination probability (genetic map taken from HapMap [112, 113]; downloaded from `https://alkesgroup.broadinstitute.org/Eagle/`) and inferred TMRCAs

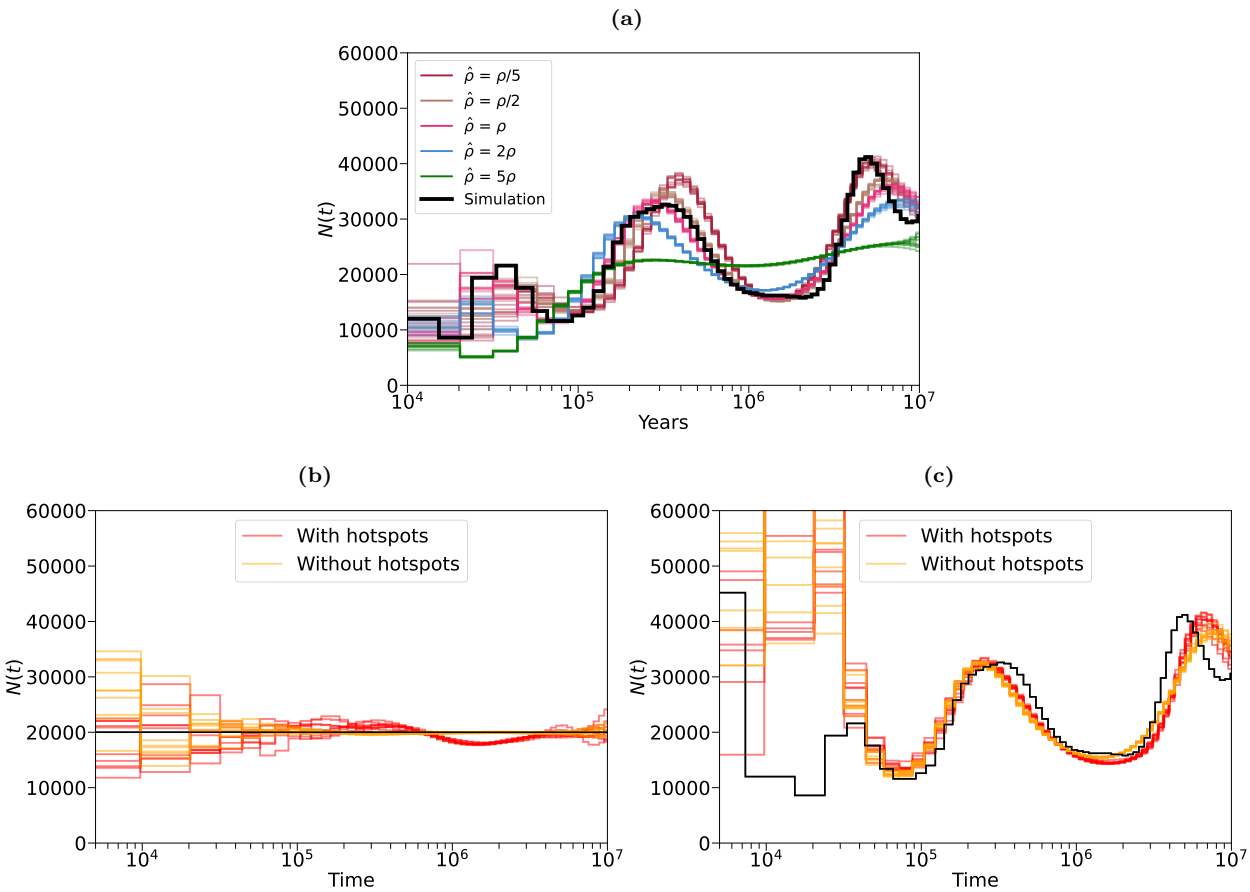

Figure A10: Inference from PSMC when its model of recombination is violated. **(a)** In a simulation with constant recombination rate ρ, we deliberately fixed the recombination rate used by the PSMC algorithm $\hat{\rho}$ at an incorrect value. The black line indicated the simulated $N(t)$ and the coloured lines indicate the various values of $\hat{\rho}$, which are expressed relative to the simulated value. **(b)** and **(c)** Inference from PSMC on simulations with changes in the recombination rate according to HapMap (red lines) and a constant recombination rate (gold lines).

(Figure A11), though the relationship appears to be non-linear. We investigated this by simulating from this genetic map, and inferring an $N(t)$ curve assuming a constant recombination rate. We observed that this type of model mis-specification does not significantly alter inference (Figure A10b and A10c). Given these observations, we believe it is unlikely that a spatially varying recombination rate could generate a fake ancient peak.

6.2.5 Recurrent mutations

PSMC assumes an infinite sites model, in which a site can only experience one mutation. In reality, some ancient sites will have experienced more than one mutation. Under a Jukes-Cantor model of mutation [114], which assumes that each base pair mutates with uniform probability to another, with probability 1/3 a second mutation at the same site will revert to its ancestral state. These will be observed as a homozygote and therefore will tend to lower the estimate of TMRCA by PSMC. With probability 2/3 a second mutation at the same site induces a distinct biallelic SNP, which is still observed as a heterozygote and so will not affect TMRCA inference. Overall, repeated mutations at the same site will tend to uniformly and gradually

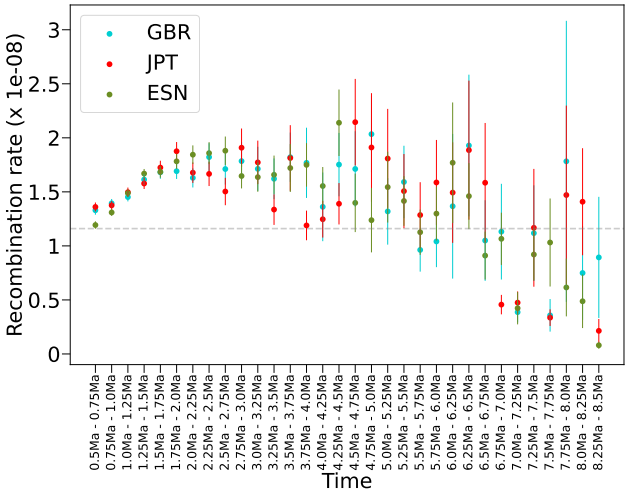

Figure A11: The relationship between the posterior mean TMRCA and the HapMap recombination rate map, on chromosome 1 for one individual from GBR, JPT, and ESN in the 1000GP. The reported Pearson r^2 for each population is 0.45, 0.41, and 0.50, with p-value<0.02 for each. However, clearly the relationship is non-linear; it appears that recombination rate increases as does TMRCA up until ∼2 Mya , after which they become negatively correlated. This may come from model violations, in that PSMC assumes a uniform rate of recombination, or that the recombination landscape changes over time [108]. The grey, dashed line indicates the chromosome wide average recombination rate.

underestimate TMRCA with increasing time, but the effect will be very small given hominin heterozygosity, of the order of 1% out to 10 million years.

Indeed, in almost all the simulations previously presented a Jukes-Cantor model [114] was used to generate mutations, in which recurrent mutations are allowed to occur. Generally, in simulations with uniform μ in time and space, we observe that inference of ancient $N(t)$ is reasonably accurate in ancient time (Figures A7, A9, and A10).

6.2.6 Background selection

Background selection (BGS) is a form of linked selection, where the removal of deleterious mutations reduces genetic diversity in the surrounding regions due to linkage [115, 116]. It has been demonstrated that BGS is pervasive throughout the human genome, and that this explains roughly 60% of the variance in diversity at the megabase scale [56, 57]. Most demographic inference methods, however, including PSMC, assume that the genome evolves neutrally. This is problematic, as it has been shown that wrongly assuming the absence of selection means that inference methods are not able to accurately reconstruct the demographic history [58, 60, 59].

We correlated the inferred TMRCAs against a high resolution map that describes the strength of BGS across the human genome [57]. This map infers a local B value, which is a measure of the reduction in local diversity due to BGS and so inversely related to the strength of BGS. The B value is calculated by integrating functional and genetic map information. We saw a significant positive correlation (Pearson $r^2 \approx 0.7$, p-value<1e-5; Figure A12) between B value and TMRCA, which is consistent with expectations that reduced diversity will be correlated with reduced coalescent time [116]. To explore the effect of BGS on

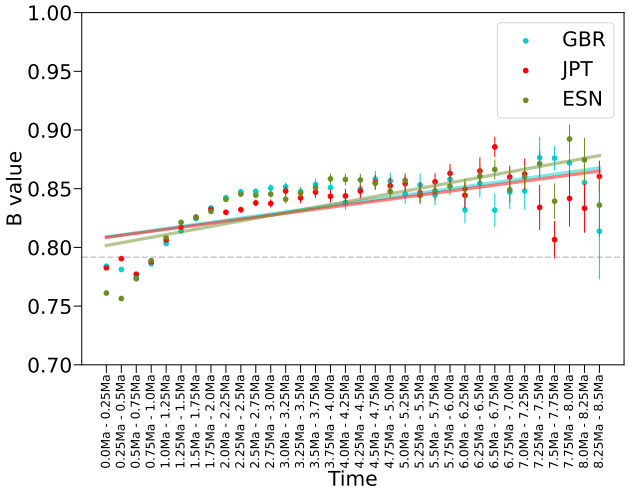

Figure A12: The relationship between the posterior mean TMRCA and B, which is a high resolution map of the effects of background selection [57], on chromosome 1 for one individual from GBR, JPT, and ESN in the 1000GP. B values indicate the local strength of BGS, with 1 being no reduction in genetic diversity due to selection and 0 being full removal of genetic diversity. The reported Pearson r^2 for each population is 0.67, 0.67, and 0.75, with p-value<1.3e-5 for each. In general this suggests that B value increases with TMRCA (a linear line of best fit has been added to show this), though the relationship appears to be non-linear. The grey, dashed line indicates the chromosome wide average B value.

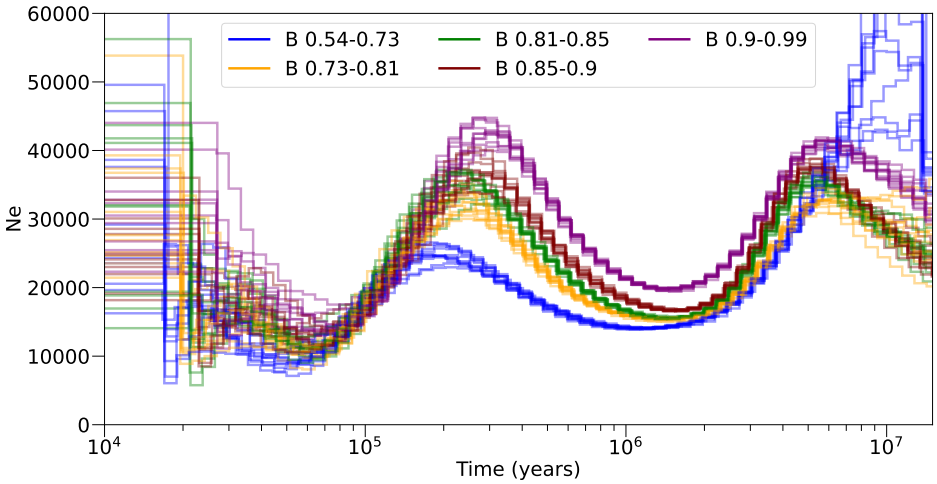

Figure A13: PSMC inference of $N(t)$ in 10 YRI individuals from the 1000GP project, stratified into quintiles according to the strength of BGS (as inferred by [57]). This figure is reproduced with permission from [62].

inference of $N(t)$ in real data, the authors in [62] binned the genome of 10 YRI individuals into quintiles according the strength of BGS and ran PSMC in each (reproduced in Figure A13). In all quintiles except the one with strongest BGS, a clear second peak is seen in ancient time. The peak is not seen in the strongest BGS quintile likely because all of the input sequence has already coalesced. In none of the simulations with realistic parameters of BGS, as shown in [59, 117, 62, 118], does PSMC infer a false ancient peak. We thus find it implausible that widespread BGS could generate an ancient peak as observed in humans.

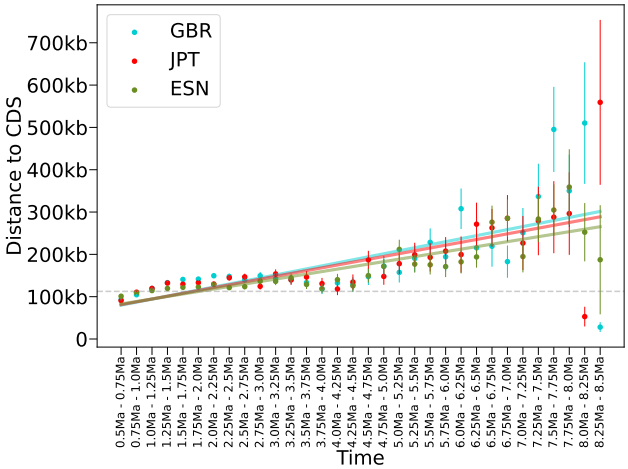

Figure A14: The relationship between the posterior mean TMRCA and mean distance to coding sequence (CDS), on chromosome 1 for one individual from GBR, JPT, and ESN in the 1000GP. The reported Pearson r^2 for each population is 0.63, 0.67, and 0.83, with p-value<0.0001 for each. In general this suggests that dcCDS increases with TMRCA (a linear line of best fit has been added to show this), although the relationship appears reversed from ∼3 Mya to ∼4.5Ma. The grey, dashed line indicates the chromosome wide average distance to CDS.

6.2.7 Balancing selection

Balancing selection (BLS) is a type of natural selection that maintains genetic diversity in a population by favouring the maintenance of multiple alleles. This type of selection can occur through various mechanisms, such as heterozygote advantage, frequency-dependent selection, or spatially varying selection [61]. BLS operating for a long time period will maintain advantageous polymorphism and result in an older TMRCA than expected under neutrality. Therefore, if BLS were sufficiently prevalent in the genome, this would manifest as enrichment of older coalescence times, which would increase inferred $N(t)$ in the past.

The prevalence of BLS in the human genome is unclear. Initially considered a rarity [119, 120] and overlooked, balanced polymorphisms have recently received renewed attention with several lines of evidence showing their relevance in human evolution [121, 122, 123]. In one study, hundreds of loci were implicated in possible trans-species BLS, maintained since earlier than human-chimpanzee speciation [124]. By finding regions of ancient shared ancestry across 54 individuals from Complete Genomics [125], in [126] the authors suggest numerous regions that are under BLS. We calculated the correlation between inferred TMRCAs and distance to coding sequence (CDS), and found that ancient TMRCAs tend to be increasingly far away from CDS (Figure A14; Pearson $r^2 \approx 0.71$, p-value<0.0001). As BLS usually acts on or near functionally important parts of the genome, this makes it unlikely that the ancient TMRCAs underlying our ancient peak are driven by BLS and that this is the cause of the ancient peak.

6.3 Methods

6.3.1 TMRCA Correlations with annotations

We downloaded the the tracks for repeats, segmental duplications, and CpG islands from the UCSC Genome Browser `https://genome.ucsc.edu/cgi-bin/hgTables` [127]. The positions of coding sequence were

obtained from Gencode `ftp.ebi.ac.uk/pub/databases/gencode/Gencode_human/release_45/gencode.v45.chr_patch_hapl_scaff.basic.annotation.gff3.gz`. We realigned all 1000 Genomes Project data to GRCh38 for comparison with various annotations. We used the B-map as inferred by Murphy et al. [57] and lifted over from GRCh37 to GRCh38 [127].

We took the posterior mean TMRCA across chromosome 1 for the 26 1000GP samples. We stratified the TMRCAs into windows of 250 kya and analysed how various functional annotations correspond to these TMRCA windows, and ensured that positions in the analysis passed the strict mappability mask.

To generate a pointwise mutation rate map, we used strand-symmetric de-novo abundance counts obtained from trios [104] to calculate the relative frequencies of each possible mutation in its trinucleotide context. To get a pointwise mutation rate per trinucleotide, we scaled these frequencies appropriately to obtain the genomewide mutation rate of 1.25e-8.