SPECIAL ISSUE : SEX CHROMOSOMES AND SPECIATION

WILEY MOLECULAR ECOLOGY

Model-based detection and analysis of introgressed Neanderthal ancestry in modern humans

Matthias Steinrücken[1,2,3] | Jeffrey P. Spence[4] | John A. Kamm[5] | Emilia Wieczorek[6] | Yun S. Song[3,5,7]

[1]Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA

[2]Department of Biostatistics and Epidemiology, University of Massachusetts, Amherst, MA, USA

[3]Department of EECS, University of California, Berkeley, CA, USA

[4]Computational Biology Graduate Group, University of California, Berkeley, CA, USA

[5]Department of Statistics, University of California, Berkeley, CA, USA

[6]Department of Mathematics, University of California, Berkeley, CA, USA

[7]Chan Zuckerberg Biohub, San Francisco, CA, USA

Correspondence
Yun S. Song, Department of Statistics, University of California, Berkeley, CA, USA.
Email: yss@berkeley.edu

Funding information
David and Lucile Packard Foundation, Packard Fellowship for Science and Engineering; National Institute of General Medical Sciences, Grant/Award Number: R01-GM094402

Abstract

Genetic evidence has revealed that the ancestors of modern human populations outside Africa and their hominin sister groups, notably Neanderthals, exchanged genetic material in the past. The distribution of these introgressed sequence tracts along modern-day human genomes provides insight into the selective forces acting on them and the role of introgression in the evolutionary history of hominins. Studying introgression patterns on the X-chromosome is of particular interest, as sex chromosomes are thought to play a special role in speciation. Recent studies have developed methods to localize introgressed ancestries, reporting long regions that are depleted of Neanderthal introgression and enriched in genes, suggesting negative selection against the Neanderthal variants. On the other hand, enriched Neanderthal ancestry in hair- and skin-related genes suggests that some introgressed variants facilitated adaptation to new environments. Here, we present a model-based introgression detection method called DICAL-ADMIX. We demonstrate its efficiency and accuracy through extensive simulations and apply it to detect tracts of Neanderthal introgression in modern human individuals from the 1000 Genomes Project. Our findings are largely concordant with previous studies, consistent with weak selection against Neanderthal ancestry. We find evidence that selection against Neanderthal ancestry was due to higher genetic load in Neanderthals resulting from small effective population size, rather than widespread Dobzhansky–Müller incompatibilities (DMIs) that could contribute to reproductive isolation. Moreover, we confirm the previously reported low level of introgression on the X-chromosome, but find little evidence that DMIs contributed to this pattern.

KEYWORDS
archaic hominin introgression, molecular evolution, population genetics, speciation

1 | INTRODUCTION

In recent years, researchers have gathered an increasing amount of high-quality genomic sequencing data from human individuals that lived thousands of years ago (Mathieson et al., 2015) and individuals of extinct hominin sister groups (Meyer et al., 2012; Prüfer et al., 2014, 2017). These ancient samples provide unprecedented opportunities to elucidate the evolution of modern human populations and their relation to other hominins. Previous genetic evidence revealed that the ancestors of non-African humans exchanged genetic material with Neanderthal individuals after emerging out of Africa. Traces of this introgression can still be found in the genomes of modern-day humans. The emerging high-quality genomic sequence data for ancient hominins not only confirm these findings,

but also allow detection of the exact location of these introgressed sequence fragments in the genomes of modern human individuals and a better understanding of their functional relevance.

Recent studies reported long regions in modern non-African individuals that are depleted of Neanderthal ancestry and enriched in genes, suggesting general negative selection against Neanderthal variants in genes (Sankararaman et al., 2014; Vernot & Akey, 2014). Harris and Nielsen (2016) and Juric, Aeschbacher, and Coop (2016) provided further evidence that natural selection has acted to remove these introgressed segments, but there is some debate about the precise cause of this selective pressure. Dobzhansky–Müller incompatibilities (DMIs) are a classic explanation for selection acting against introgressed alleles and play an important role in the evolution of reproductive isolation during speciation. DMIs involve alleles that have arisen separately in each population and are neutral or adaptive in isolation, but are deleterious when brought together in individuals of hybrid ancestry (Dobzhansky, 1936; Orr, 1995). DMIs have been observed in the hybrids of other species—for example *Drosophila simulans* and *D. melanogaster* (Brideau et al., 2006); *Mimulus guttatus* and *M. nasutus* (Fishman & Willis, 2001); and *Ambystoma californiense* and *A. tigrinum mavortium* (Fitzpatrick, 2008)—and were hypothesized to be the cause (particularly in male hybrids) of selection against Neanderthal ancestry in modern humans (Sankararaman et al., 2014). This hypothesis was motivated by finding significant enrichment of testes-expressed genes in regions of low Neanderthal ancestry and a substantial reduction of Neanderthal ancestry on the X-chromosome. Indeed, sex chromosomes are thought to play a special role in speciation, and it has been observed, for example in *Drosophila*, that loci contributing to reduced male fertility in hybrids are concentrated on the X-chromosome (Presgraves, 2008).

Sankararaman et al. (2014) and Vernot and Akey (2014) reported some allelic variants associated with genetic diseases in genomewide association studies (GWAS) that might have originated from the Neanderthal population. The studies also reported enriched Neanderthal ancestry in hair- and skin-related genes (keratin pathways), which suggests that these introgressed variants could have helped modern non-African populations to adapt to their local environments. All in all, the availability of Neanderthal introgression maps in modern humans has led to numerous follow-up studies investigating the various functional, evolutionary and medical implications of archaic hominin introgression into modern humans (Dannemann, Prüfer, & Kelso, 2017; Gittelman et al., 2016; Rogers, 2015; Schumer et al., 2017; Simonti et al., 2016).

Methods to detect Neanderthal introgression tracts include a machine learning-based approach (Sankararaman et al., 2014) that operates on suitably chosen "features" of the genetic data, and an approach based on sequence identity and divergence (Vernot & Akey, 2014). These methods have been extended to jointly detect Neanderthal and Denisovan introgression (Sankararaman, Mallick, Patterson, & Reich, 2016; Vernot et al., 2016); the latter is found to be more prevalent in Oceania and Southeast Asia.

In this article, we present a modification of the method DICAL 2.0, previously developed by Steinrücken, Kamm, and Song (2015) for the inference of complex demographic histories. This modification, which we call DICAL-ADMIX, can be used to efficiently detect tracts of introgressed Neanderthal DNA. It is based on a hidden Markov model (HMM) approach that explicitly accounts for the underlying demographic history relating modern human and Neanderthal populations, including the introgression event. We first present our model-based method for detecting Neanderthal introgression and demonstrate through extensive simulations that our method is able to efficiently and accurately detect introgression in simulated data. We then apply our method to sequence data of modern humans from the 1000 Genomes Project and a high-coverage genome from a Neanderthal individual from the Altai mountains (Prüfer et al., 2014). Our results are in general agreement with previously obtained results, and we discuss similarities, differences and their functional implications. In particular, we do not find evidence to support the hypothesis of Sankararaman et al. that DMIs played a role in shaping the pattern of introgression either genomewide or specifically on the X-chromosome.

2 | MATERIALS AND METHODS

2.1 | Overview of our method

The method to detect Neanderthal introgression that we present here accounts explicitly for the underlying demographic history relating modern humans and Neanderthals. Therefore, we briefly present some of the key features of this demographic model. Researchers have studied various aspects of the ancestral relations between modern African and non-African individuals using different methodologies. Furthermore, many studies have investigated the divergence of Neanderthals and modern humans. These studies resulted in several different, albeit largely consistent, estimates for the relevant demographic parameters, and here, we closely follow Sankararaman et al. (2014) (specifically Figure S2.1 of that paper) for consistency. However, note that Sankararaman et al. used a mutation rate of 2.5×10^{-8} per-site per-generation, whereas we use 1.25×10^{-8}, thus some of the numbers need to be adjusted.

The demographic model we use is depicted in Figure 1a. The size of the most ancestral population and the size of the population ancestral to modern humans before the expansion are set to $N = 11{,}000$. The size of the population ancestral to modern humans after the expansion and the size of the African population are set to $N = 23{,}000$. These sizes are consistent with the estimates provided by Jouganous, Long, Ragsdale, and Gravel (2017). The size of the non-African population after the split is set to $N = 2{,}000$. The size of the Neanderthal population is set to $N = 2{,}000$ as well, since it has been shown by Prüfer et al. (2014) that the Neanderthal population size declined rapidly as the population neared extinction. The more recent small size will have a stronger impact on genetic variation than the larger ancestral size. Several studies (Jouganous et al., 2017) report a strong population bottleneck in European and Asian populations after the out-of-Africa event, followed by rapid exponential population growth, and Sankararaman et al. (2014)

incorporate this into their demographic model. As we will detail later, our method considers each non-African haplotype one at a time. The genetic processes along a single ancestral lineage are not affected by the exact population size history, and thus, we do not explicitly include these details in the demographic model used here.

Following Prüfer et al. (2014), we set the time of divergence between modern humans and Neanderthals, T_{nean}, to 26,000 generations ago, which corresponds to 650 kya, assuming a generation time of 25 years, and the time that the population ancestral to modern humans expands is set to 12,000 generations ago (Jouganous et al., 2017). The split between African and Non-African T_{div} is set to 4,000 generations ago, or 100 kya (Scally & Durbin, 2012), and the time of the introgression or admixture event T_{admix} is set to 2,000 generations ago (Sankararaman, Patterson, Li, Pääbo, & Reich, 2012), which corresponds to 50 kya (Prüfer et al., 2014). Finally, the introgression coefficient is set as 3%; that is, a non-African individual at the time of introgression had a 3% chance that its parent was a Neanderthal individual. This is consistent with previous estimates of this quantity obtained by Green et al. (2010) and Juric et al. (2016). In Section 2, we will demonstrate the robustness of our method to misspecification of key parameters. It has been debated whether all non-African populations received genetic material from Neanderthals in a single pulse of introgression or several pulses. Here, we assume that there has been only one pulse and will defer disentangling such models to future work.

Before we describe our method to detect tracts of Neanderthal introgression, we provide a brief overview of the methods developed by Sankararaman et al. (2014) and Vernot and Akey (2014) for comparison. In Sankararaman et al. (2014), the authors employ a machine-learning approached based on conditional random fields, discriminative analogs of HMMs. To apply this framework, the authors represent the genotype data of the reference Neanderthal, the reference African population and the focal non-African population in terms of "features" that are informative to distinguish between introgressed and nonintrogressed sequence tracts. The authors chose three different classes of features: the distribution of alleles at informative SNPs; a measure of sequence divergence between the focal individual and both reference populations; and a feature to match the length distribution of the observed tracts to be consistent with the expectation from an introgression event 37–86 kya. The authors train the model on data simulated under a demographic scenario similar to Figure 1a and apply the trained model to detect introgression tracts in individuals from the 1000 Genomes data set (The 1000 Genomes Project Consortium, 2012). A modified version of this methodology was also applied in Sankararaman et al. (2016) to detect Denisovan ancestry in Southeast Asians.

Vernot and Akey (2014) developed a two-stage procedure to detect introgression tracts. In the first stage, the authors computed S* statistics (Plagnol & Wall, 2006) in sliding windows along the genome. This statistic is sensitive to increased levels of diversity in high linkage disequilibrium, indicating a more ancient T_{MRCA} of a given region and also considers the tract length to detect archaic introgression. Notably, this stage does not require a reference sequence for an archaic individual. In the second stage, the authors compare the identified segments to a reference Neanderthal genome to reliably identify Neanderthal introgression. The same two-stage approach was applied by Vernot et al. (2016) to identify Denisovan introgression.

Here, we apply a modified version of the method DICAL 2.0, developed by Steinrücken et al. (2015) for inference of ancient
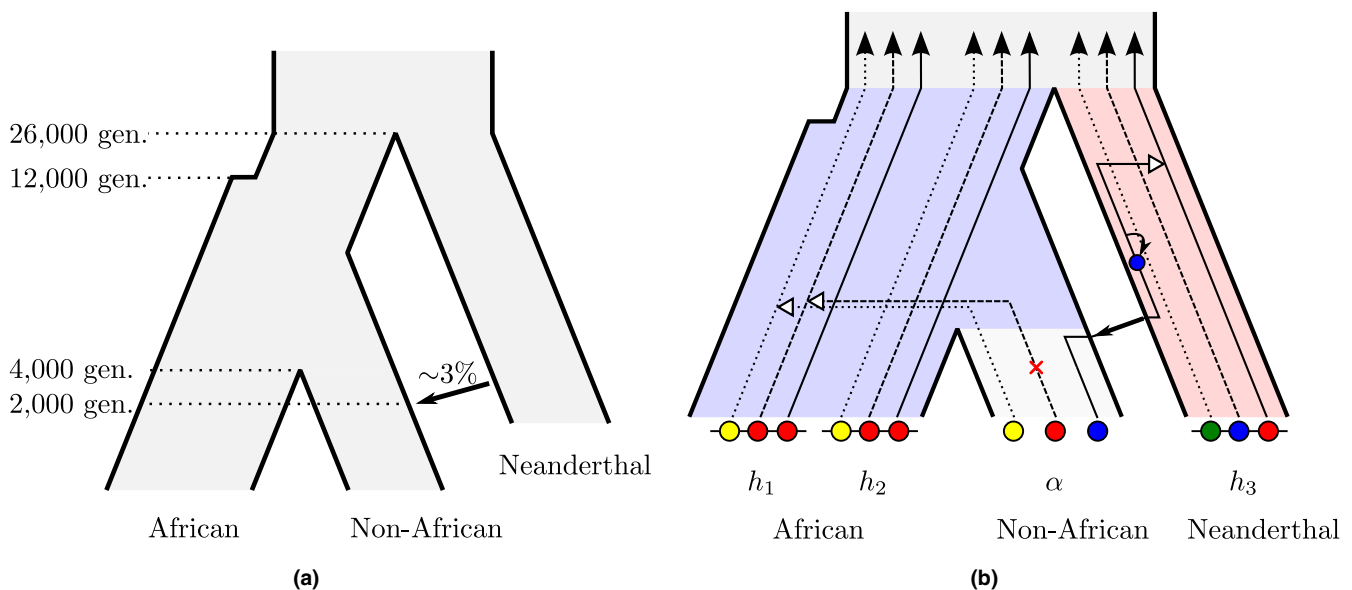

FIGURE 1 Illustration of the demographic model and the coalescent hidden Markov model. (a) A sketch of the demographic history relating African, non-African and the Neanderthal populations that contain the key features. (b) The conditional sampling distribution applied to the detection of Neanderthal introgression. At each locus, the ancestral lineage of the focal Non-African (α) haplotype coalesces either with an African lineage (h_1 or h_2) or a lineage in the Neanderthal population (h_3) through the introgression event [Colour figure can be viewed at wileyonlinelibrary.com]

demographies, to detect sequence tracts of Neanderthal DNA introgressed into modern humans. The method is based on the conditional sampling distribution (CSD; Paul & Song, 2010; Paul, Steinrücken, & Song, 2011; Steinrücken, Paul, & Song, 2013), which is similar to the copying model of Li and Stephens (2003), and is depicted in Figure 1b. The CSD describes the distribution of sampling an additional focal genome or haplotype, conditional on having already observed a certain set of haplotypes. Steinrücken et al. (2015) introduce a version of the CSD that can be applied to haplotypes sampled from several subpopulations, accounting explicitly for the underlying demographic history. Under this model, the unknown genealogy relating the already observed haplotypes is approximated by a trunk genealogy of unchanging ancestral lineages extending infinitely into the past. At each locus, the ancestral lineage of the additional haplotype is absorbed into a lineage of the trunk. The dynamics of absorption depends on the underlying demographic history. In brief, lineages in different subpopulations cannot coalesce, unless continuous or point migration is possible at given rates, and the likelihood of coalescence is larger in small populations, but decreases in large populations. If an ancestral recombination event separates two loci, the haplotype of absorption and the time of absorption can change; thus, different genomic segments can be copied from different haplotypes in the trunk, and the additional haplotype is realized as a mosaic of the observed haplotypes. The CSD can be implemented as an HMM along the genome of the additional haplotype, where the hidden state is the trunk haplotype that the genetic material is currently copied from, and a time of absorption. This absorption time is proportional to the likelihood that mutations can alter the genetic type at a given locus. Steinrücken et al. derived the emission and transition probabilities for the underlying HMM under general demographic models.

This CSD can be applied to detect introgressed tracts of Neanderthal ancestry in modern humans as follows. First, fix the underlying demography given in Figure 1a that relates modern African populations, modern non-African populations and the Neanderthal population. The introgression event is modelled as a point migration. As depicted in Figure 1b, the haplotypes sampled in the African population and a Neanderthal haplotype are used as the trunk haplotypes in their respective subpopulation. Then, each non-African sample is, in turn, used as the additional haplotype in the non-African subpopulation. Computing the forward and backward algorithm under the HMM for this CSD yields a marginal posterior distribution over the hidden states at each locus. Recall that these hidden states consist of both an absorption time and an absorbing haplotype. Marginalizing over the absorption time results in a posterior distribution over the trunk haplotypes, and, grouping haplotypes by subpopulation, this gives a probability at each locus that the locus in the non-African haplotype is obtained from either an African (modern human) ancestor, or from an Neanderthal ancestor through introgression. Note that using an explicit time for the introgression event in the demographic model implicitly specifies a prior distribution for the length of the introgressed tracts. The software implementation of this method, DICAL-ADMIX, is available at http://dical-admix.sourceforge.net/.

2.2 | Simulation study

To demonstrate that DICAL-ADMIX can be used to accurately and efficiently identify tracts of Neanderthal introgression, we performed an extensive simulation study. To this end, we used the coalescent simulator MSPRIME (Kelleher, Etheridge, & McVean, 2016) to simulate sequence data under the demographic model given in Figure 1a. Specifically, we simulated 176 haplotypes in the African population, one haplotype in the Neanderthal population and 20 in the focal non-African population. For each such data set, we simulated 20 Mbp of sequence data, using a per-generation population-scaled mutation and recombination rate of 0.0005 per base (corresponding to a per-base per-generation rate of 1.25×10^{-8}). We simulated 50 replicates in each scenario and estimated the introgression tracts on each focal haplotype.

To investigate how robust our method is to misspecification of the demographic model used for inference, our simulation study was twofold. First, we simulated data under the demographic model given in Figure 1a and varied the demographic model used for the analysis. We kept all parameters fixed and only varied one focal parameter at a time. We varied the divergence time between modern humans and Neanderthal, using 0.5 and 2 times T_{nean} = 26,000 generations, and the divergence time between Africans and non-Africans, using 0.8 and 2 times T_{div} = 4,000 generations. Moreover, we varied the fraction of introgressed Neanderthal individuals, using 0.5 and 2 times admix = 3%, and the time of the introgression event, using 0.8 and 1.25 times T_{admix} = 2,000 generations. We also considered the effect of simulating under spatially heterogeneous recombination rates, taking 20 Mbp from the recombination maps inferred in Myers, Bottolo, Freeman, McVean, and Donnelly (2005); we call this model the "HapMap" model. Finally, we considered four more complicated models. These models are the same as that given in Figure 1a except in the ways that we highlight below. While some of these models may not be indicative of any hypothesized demographic events, they show the robustness of DICAL-ADMIX and highlight its potential utility in other scenarios. In the "CEU ghost" model, a non-African "ghost" population diverges from a basal CEU population 2,300 generations ago, and 500 generations ago the two populations admix with 30% of their ancestry deriving from the ghost population and the remaining 70% coming from the basal CEU population. In the "YRI ghost" model, an ancient hominid ghost population diverges from the Neanderthal lineage 17,000 generations ago, and then there is a 3% pulse of admixture from this ghost population into YRI 1,700 generations ago. In the "Nean into YRI" model, there is a 3% pulse of admixture into both YRI and CEU 2,000 generations ago. Lastly, in the "Mig" model, YRI and CEU exchange migrants at a rate of 0.001% of the population per generation.

For simulated data, the true introgression status at each locus is known. Running DICAL-ADMIX to detect introgressed tracts in the simulated non-African individuals yields a posterior probability of introgression at each locus. Using different thresholds on this posterior probabilities for calling a locus introgressed, we can assess the true-

and false-positive rate and the precision to generate receiver operating characteristic (ROC) and precision-recall curves for each analysis. Figure 2a shows these curves when the data simulated under the "true" model (Figure 1a) are analysed using the different demographic models. We observe an overall good performance and robustness against misspecification of the parameters for analysing the data. From these plots, we determined that a threshold of 0.42 yields a good balance of the different performance metrics. We indicated this threshold along the curves by a red cross. Note that it is possible to increase the true-positive rate by lowering the threshold; however, the decrease in precision would be more severe. Thus, we used this threshold in the remainder to call introgression tracts in the 1000 Genomes data set. The ROC and precision-recall curves

for the four more complicated scenarios and the *HapMap* model are provided in Section S2.1 of the Supporting Information.

For the second part of the simulation study, we simulated data under the different demographic models. We then analysed each data set using the same parameters as used for the simulation on the one hand, and using the parameters of the "true" model (Figure 1a) on the other hand. The ROC and the Precision-recall curves for varying the introgression percentage are depicted in Figure 2b,c, and the curves for the remaining scenarios are given in Section S2.1 of the Supporting Information. Again, we observe that misspecifying the parameters of the analysis does not affect the performance substantially. The ROC curves demonstrate a good performance in terms of true- and false-positive rate in most scenarios; however,
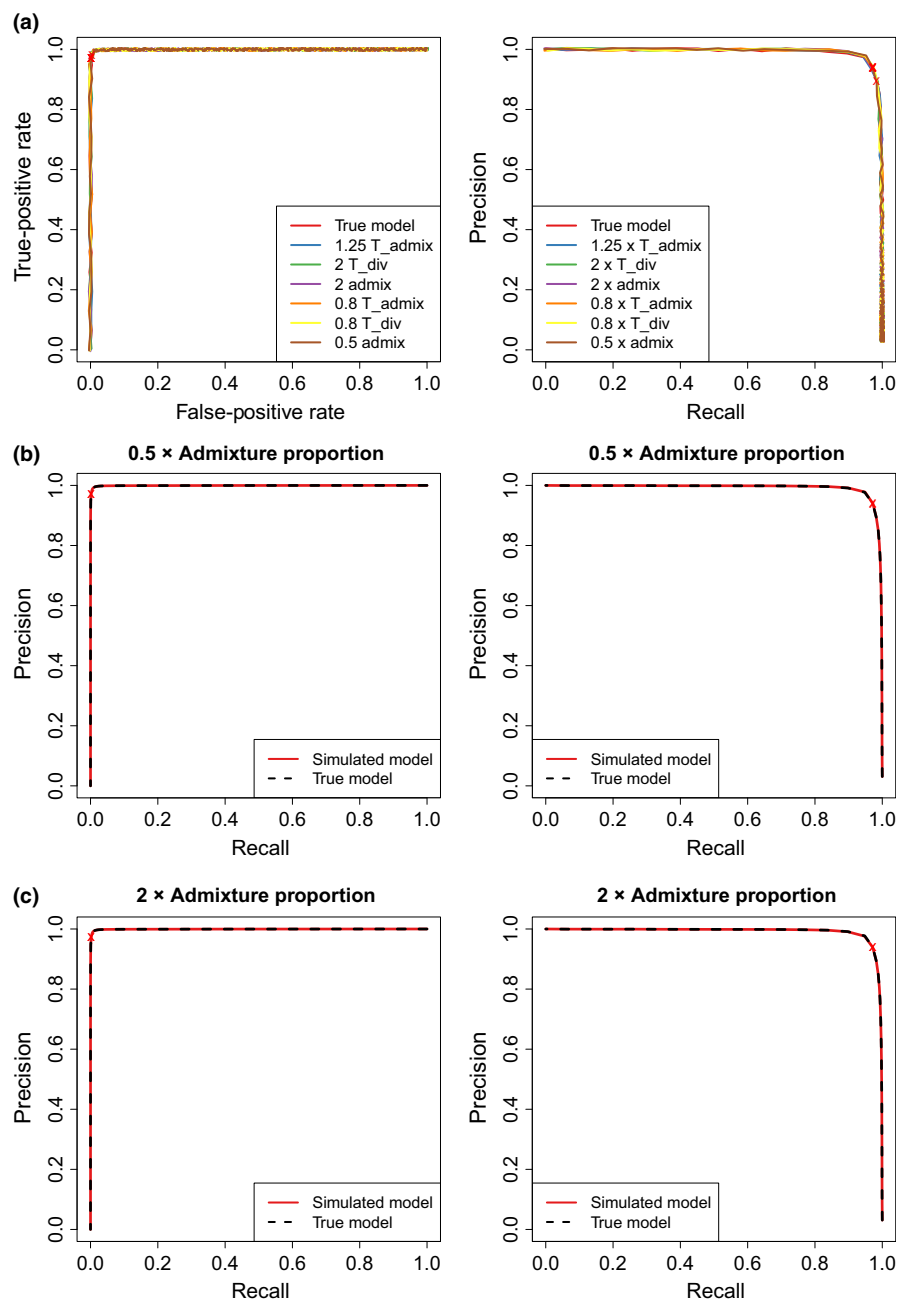

FIGURE 2 Receiver operating characteristic and Precision-recall curves from the simulation study. (a) For data simulated under the "true" model, and the parameters for the analysis varied. We added a small amount of random "jitter" to the curves as they are hard to distinguish otherwise. (b) For data simulated under a model where the admixture proportion is 1.5% and analysed using the same model and the "true" model. (c) For data simulated under a model where the admixture proportion is 6% and analysed using the same model and the "true" model [Colour figure can be viewed at wileyonlinelibrary.com]

the precision-recall curves exhibit a poorer performance in some scenarios. This is to be expected, as in some scenarios, the Neanderthal and African population are genetically closer, for example, when the divergence time between Africans and non-Africans is increased, or the divergence time between Neanderthal and modern humans is decreased. If the populations are more closely related, it becomes more difficult to distinguish introgressed variation from variation shared between modern human populations. Lastly, we examine the detection power of our method with varying tract length in Section S2.2 of the Supporting Information and observe good performance across all sizes.

3 | RESULTS

3.1 | Neanderthal introgression in the 1000 Genomes data

We applied DICAL-ADMIX to detect tracts of Neanderthal introgression in non-African individuals from the Phase I data set of the 1000 Genomes Project (The 1000 Genomes Project Consortium, 2012), focusing on Europeans (CEU) and East Asians (CHB and CHS) in particular. These data were collected using low-coverage short-read sequencing, potentially affecting the quality of the genotype calls. However, we believe that this should not introduce substantial bias into our results. Moreover, we applied the strict mappability mask (The 1000 Genomes Project Consortium, 2012) to exclude from our analysis the genomic regions where no confident genotype calls were made in the 1000 Genomes data set. We used the 88 YRI individuals (176 haplotypes) from this data set as reference African haplotypes, assumed to have no introgressed genetic material from Neanderthals, that serve essentially as a modern human reference panel. We applied DICAL-ADMIX to compute the marginal posterior introgression probability along the genomic sequences of each of the 85 CEU individuals (170 haplotypes), 97 CHB individuals (194 haplotypes) and 100 CHS individuals (200 haplotypes) in turn. We used a high-coverage genomic sequence from an Altai Neanderthal individual (Prüfer et al., 2014) as a Neanderthal reference. Prüfer et al. (2014) presented different genome alignability filters (Prüfer et al., 2014, SI 5b), and we used the `map35_50%`-filter, as this filter was suggested by the authors to be most appropriate for population genomic analyses.

DICAL-ADMIX requires that the genomic data be phased into haplotype sequences. The 1000 Genomes data set is computationally phased, so we could use this data as provided; however, the diploid sequence of the Neanderthal individual cannot be phased using standard statistical methods. We instead used an additional preprocessing step to obtain a pseudohaplotype sequence. As noted by Prüfer et al. (2014), the Altai Neanderthal individual exhibits only a sixth of the heterozygosity of modern non-African individuals. Thus, the number of ambiguous sites that require phasing is small. We tested three different methods to obtain a haplotype allele for these remaining sites: choosing an allele uniformly at random, using the ancestral allele only and using the derived allele only, where the

ancestral states at each locus were determined using a six-primate consensus (Paten, Herrero, Beal, Fitzgerald, & Birney, 2008). We observed little difference in our results between the different approaches, and thus, we only present the results using the derived-allele approach. We used a mutation rate of 1.25×10^{-8} per-site per-generation (Scally & Durbin, 2012) and chromosome-specific recombination rates obtained by averaging the fine-scale rates provided by Kong et al. (2010). Note that we made the simplifying assumption that the recombination rate is constant within each chromosome. Due to computational considerations, we did not compute the posterior at every genomic site, but rather grouped sites together into 500 bp windows; the details of this procedure are provided in Steinrücken et al. (2015). Furthermore, we applied a moving average filter of length 15 kbp to the raw posterior in a postprocessing step, to smooth sudden changes. We empirically observed that this filtering step improves detection.

Moreover, we obtained from Sankararaman et al. (2014) the likelihoods of Neanderthal introgression that they computed for the same individuals. To compare these calls obtained at the SNPs to the ones obtained using DICAL-ADMIX, we interpolated the Sankararaman et al. likelihoods at the position in the middle of the 500-bp windows employed by DICAL-ADMIX. As advised by Sankararaman et al., we used a threshold of 0.89 to call Neanderthal introgression tracts. Vernot and Akey (2014) also identified tracts of Neanderthal ancestry in the individuals from the 1000 Genomes data set, excluding the X-chromosome. We downloaded the population summaries from http://akeylab.princeton.edu/downloads.html and compared them to the results obtained with DICAL-ADMIX, when possible.

We computed the average of the marginal posterior introgression probability obtained at each locus using DICAL-ADMIX across each chromosome and across all CEU individuals, and separately across all CHB + CHS individuals. We performed the same averaging for the posterior probabilities obtained by Sankararaman et al. (2014). Figure 3a shows the results for each chromosome in the CEU population, while Figure 3b shows the results for the CHB+CHS population. We find an average introgression probability of 1.48% in the CEU autosomes, and 1.80% in the CHB + CHS autosomes, whereas Sankararaman et al. report 1.13% and 1.35%, respectively. However, the average amount of introgression detected using DICAL-ADMIX varies, from as low as 0.57% on chromosome 17 in the CEU population, to as high as 3.29% on chromosome 9 in the CHB + CHS population. Compared to the autosomes, Sankararaman et al. previously reported a lower amount of introgression on the X-chromosome in CEU (0.18%) as well as in CHB + CHS (0.24%). Similarly, we observed a roughly fourfold decrease on the X-chromosome when compared to the autosomes in CEU (0.38%) and CHB + CHS (0.54%). The general patterns of introgression stratified by chromosome are in good agreement with Sankararaman et al. However, DICAL-ADMIX detects on average 30% more introgression, which might be attributable to the fact that it detects more short tracts (see also Figure 7). Note that most Neanderthal ancestry proportions are lower than the 3% that was assumed in our model, which indicates that

Neanderthal ancestry has been preferentially removed. In Section 3.3, we discuss in more detail possible mechanisms for this purging of Neanderthal ancestry.

The posterior distributions along the chromosomes allow for a more detailed view of Neanderthal introgression into modern humans as it varies along the genome. We determined whether a given locus is admixed on a particular haplotype by thresholding the posterior generated using DICAL-ADMIX at 0.42 and thresholding the posterior from Sankararaman et al. (2014) at 0.89. We then averaged these calls across 1 Mbp windows and across the individuals in the respective populations and plotted the result as piecewise constant functions. The skyline plots in Figure 4 show the percentage of Neanderthal introgression along chromosome 4 in CEU and the X-chromosome in the CEU population. (In Section S3 of the Supporting Information, we provide skyline plots for all chromosomes in the CEU and the CHB + CHS populations.) In addition, we indicated
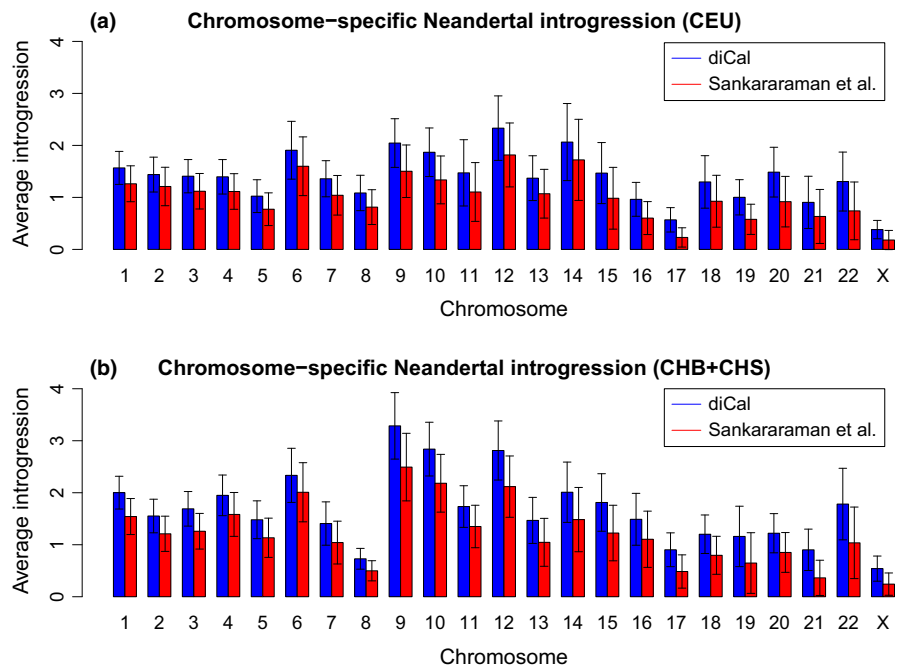

FIGURE 3 The amount of Neanderthal introgression in modern-day individuals, stratified by chromosome. The height of each bar gives the average introgression, and the whiskers indicate the standard deviation across the sample. (a) CEU. (b) CHB + CHS [Colour figure can be viewed at wileyonlinelibrary.com]
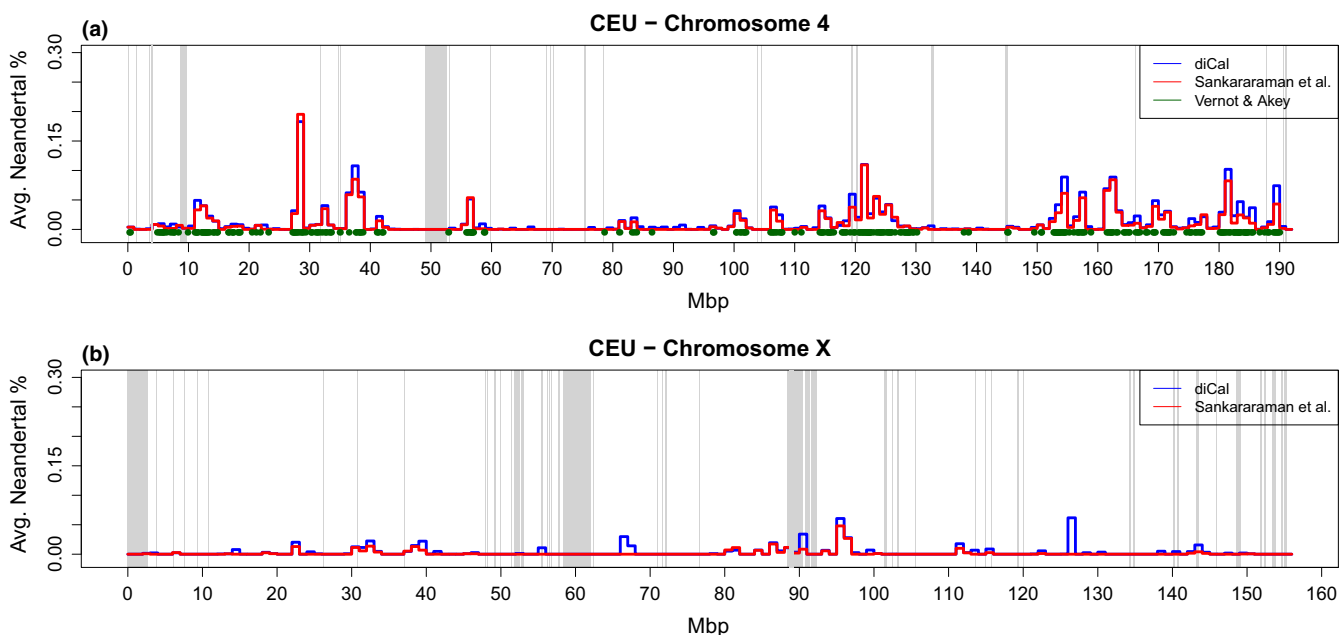

FIGURE 4 Skyline plot of the amount of Neanderthal introgression in the CEU population, averaged over all individuals in 1 Mbp windows. The results from DICAL-ADMIX are indicated in blue, and the results from Sankararaman et al. (2014) indicated in red. The regions reported as introgressed by Vernot and Akey (2014) are indicated in green. The grey bars denote the regions were no calls were made in the 1000 Genomes data set, which include the centromeres. (a) Chromosome 4. (b) Chromosome X [Colour figure can be viewed at wileyonlinelibrary.com]

the regions on the autosomes that were identified in Vernot and Akey (2014) to be introgressed. As mentioned earlier, the X-chromosome was excluded in their study. We see good agreement between the calls made using DICAL-ADMIX and the calls from Sankararaman et al. (2014). Furthermore, the regions of introgression detected by Vernot and Akey (2014) cluster in regions were the skyline plots indicate introgressed genetic material.

To investigate the shared features and differences between the introgression call sets, we generated Venn diagrams. For the DICAL-ADMIX posterior and the posterior from Sankararaman et al. (2014), we used the aforementioned thresholds to call introgression tracts in the CEU and CHB + CHS individuals. For each individual, we assessed at each locus whether either method, both methods, or no method detected Neanderthal introgression, and averaged these indicators to get population-wide percentages for the whole genome. Figure 5a,b shows the Venn diagrams for the different call sets in the CEU population, for the autosomes and the X-chromosome, respectively. Figure 5c,d depicts the results for the CHB + CHS population, on the autosomes and the X-chromosome, respectively. We observe a large overlap between the calls based on DICAL-ADMIX and Sankararaman et al. (2014) on the autosomes, but less agreement on the X-chromosome. We also generated population-wide introgression maps in each population, called a *tiling path* by Sankararaman et al. (2014). To this end, we identified those regions on the chromosomes where introgression was called for at least one individual in the respective population. We then compared these population-wide introgression maps with the population-level introgression maps published by Vernot and Akey (2014). We generated three-way Venn diagrams for the autosomes in the CEU population and the CHB + CHS population, shown in Figure 6, both in units of percentage of the whole autosome. Again, we observe a large overlap between DICAL-ADMIX and Sankararaman et al. (2014), but less so with Vernot and Akey (2014). This discordance might be explained to some degree by the fact that in the
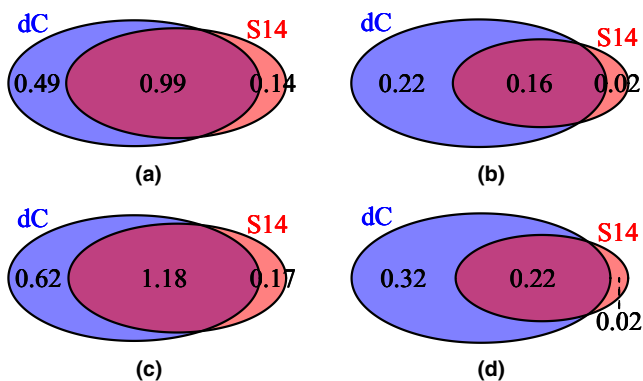

FIGURE 6 Venn diagrams of the average percentage of regions on the autosomes where Neanderthal introgression was called for at least one individual in the population, by DICAL-ADMIX (dC), Sankararaman et al. (2014) (S14), and Vernot and Akey (2014) (V14). (a) CEU. (b) CHB + CHS [Colour figure can be viewed at wileyonlinelibrary.com]

two-stage procedure of Vernot and Akey (2014), the first step does not use sequence information from the Neanderthals as in the other methods. Thus, regions of high sequence identity between modern non-Africans and Neanderthals might be missed in this first step.

We also investigated the distribution of fragment lengths that were detected by the different methods. For all individual from a given population, we counted the number of times an introgression tract of a specific length was detected. Figure 7a,c depicts the distributions of the absolute frequencies in the autosomes of the individuals in the CEU population and the CHB + CHS population, respectively. Figure 7b,d shows the same distributions for the X-chromosomes. In addition to the empirical tract length distribution obtained from the 1000 Genomes individuals, we plotted the neutral expectation of the absolute frequencies. This neutral expectation of the tract length distribution is computed under the following simple model. Approximating the chromosome as continuous, and considering the introgression tract in an individual at present, the distance between recombination break points is exponentially distributed with parameter $g \times r$, where r is the per-generation per base pair recombination probability and g is the number of generations since the introgression event, because in each generation, there is a chance that recombination breaks down the introgression tract. Here, we used $g = 2,000$ and $r = 1.19 \times 10^{-8}$ per-base per-generation for the autosomes and $r = \frac{2}{3} \times 1.18 \times 10^{-8}$ per-base per-generation for the X-chromosome. The exponential rate $g \times r$ can also be used to obtain the expected number of sequence tracts in a genome of a certain size, 3% of which are introgressed from an ancestral Neanderthal individual, which yields the expected absolute frequency.

This simple model for the neutral expectation is certainly oversimplified, but it serves as a first approximation. It is not discernible in these plots whether deviation from the neutral expectation is due to incorrect detection of the tracts, or the true underlying tracts actually being subject to non-neutral evolution. The fact that both methods deviate from the neutral expectation in qualitatively similar ways suggests both factors may be playing a role. However, it is

FIGURE 5 Venn diagrams of the average percentage of introgressed genetic material detected by DICAL-ADMIX (dC) and by Sankararaman et al. (2014) (S14). (a) For autosomes in the CEU individuals. (b) For the X-chromosome in the CEU individuals. (c) For autosomes in the CHB and CHS individuals. (d) For the X-chromosome in the CHB and CHS individuals [Colour figure can be viewed at wileyonlinelibrary.com]
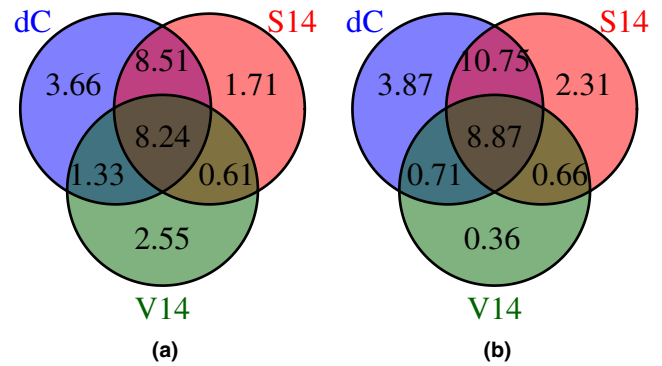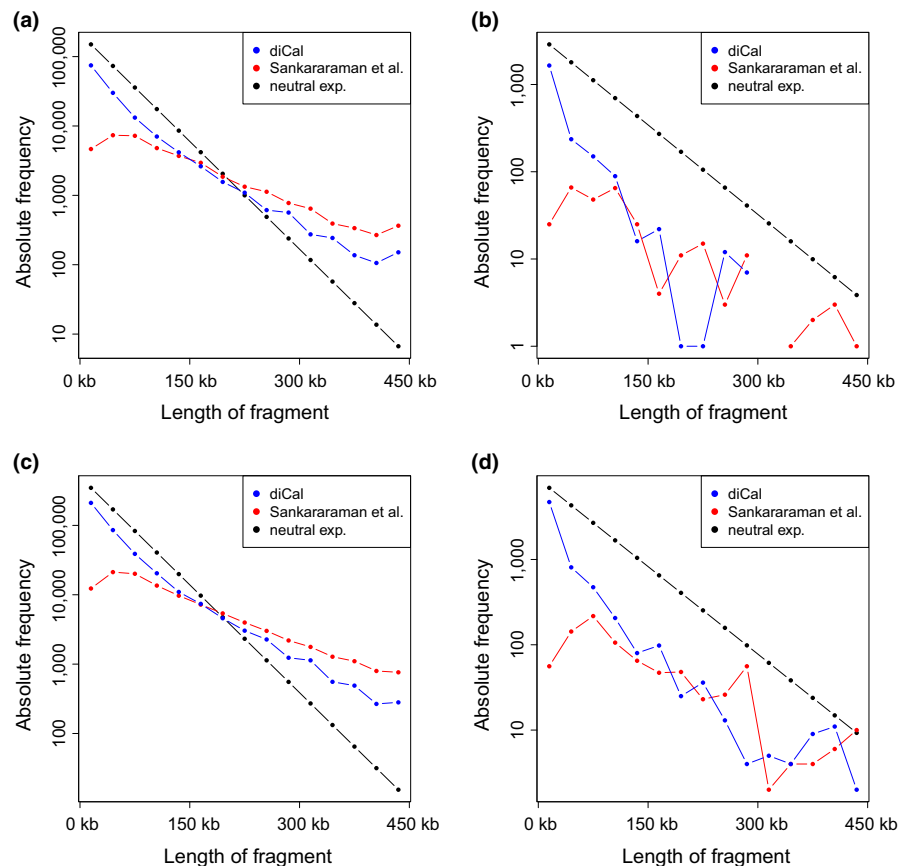
FIGURE 7 The empirical distribution of the lengths of the introgression tracts, accumulated across individuals. The different length is binned into classes of size 30 kbp. Missing values indicate unobserved classes. (a) Absolute frequency of tracts of a given length on the autosomes in the CEU population. (b) Absolute frequency of tracts of a given length on the X-chromosome in the CEU population. (c) Absolute frequency of tracts of a given length on the autosomes in the CHB + CHS population. (d) Absolute frequency of tracts of a given length on the X-chromosome in the CHB + CHS population [Colour figure can be viewed at wileyonlinelibrary.com]

surprising that, for the autosomes, both methods detect more long fragments than expected under the simple neutral model. This may be due to recombination rate variation along the genome or due to our slightly lower power to detect shorter fragments (see Section S2.2 of the Supporting Information). It could also suggest that either there was an additional introgression event that happened more recently than 2,000 generations ago, or that some form of selection is acting that favours longer fragments.

In general, DICAL-ADMIX detects more short fragments and fewer long fragments than reported by Sankararaman et al. (2014). Moreover, the empirical distribution of DICAL-ADMIX is closer to the neutral model. This and the other statistics of the empirical distribution of the introgressed Neanderthal tracts presented in this section suggest that there is merit in applying different methodologies for the detection of introgression. While all methods perform reasonably well on simulated data, they seem to be sensitive to slightly different features of the introgression tracts. Thus, we suggest using the consensus of the three methods for highly confident introgression calls and using regions unique to only some of the methods for more exploratory research.

3.2 | Functional implications of Neanderthal introgression

To explore the functional implications of Neanderthal introgression, we performed a gene ontology (GO) analysis using GORILLA (Eden,

Lipson, Yogev, & Yakhini, 2007; Eden, Navon, Steinfeld, Lipson, & Yakhini, 2009), which looks for overrepresentation of GO terms at the top of a ranked list of genes. For each population, we ranked genes by their mean posterior probability of introgression (as determined by the DICAL-ADMIX posterior decoding) and looked for GO terms associated with either a lack of Neanderthal introgression or an enrichment of introgression. We restricted our analyses to the 500 bp resolution introgression calls where no more than half of the bases are masked by the 1000 Genomes strict mappability mask (The 1000 Genomes Project Consortium, 2012). Furthermore, we only included genes where <10% of sites were masked. The results (shown in Section S1 of the Supporting Information) are broadly concordant between populations. Like Sankararaman et al. (2014), we find that genes associated with keratin are more likely to be introgressed than other genes, which hint at the possibility of adaptive introgression for these genes. Intriguingly, sensory perception, particularly olfaction, both had genes more likely to be introgressed as well as genes less likely to be introgressed. It is possible that adaptive introgression has played a role for some of these genes, for example by helping to adapt to local environments and that selection has removed introgression at other olfaction-related genes. It is also possible, however, that this is an artefact of either the introgression calls (e.g., due to high amount of polymorphism in olfactory genes Malnic, Godfrey, and Buck (2004)) or high variance in mean introgression rate (e.g., due to being smaller than other genes, or being spatially clustered in the genome).

We also investigated whether SNPs associated with particular phenotypes are more or less likely to be introgressed on average. To this end, we downloaded the results of 2,419 GWASs that were performed on data from the UK Biobank (Global Biobank Engine, 2017; Sudlow et al., 2015) and extracted all of the SNPs that were significant at a genomewide significance level of 5.0×10^{-8} for each GWAS. We then tested whether the mean posterior probability of introgression at these SNPs was significantly higher or lower than expected using our bootstrap-like test (described below). Perhaps due to the large number of tests performed, we found only one statistically significant result: loci associated with being treated with desloratadine, a drug used to treat allergies, were significantly more likely to be introgressed (Bonferroni-corrected $p < .025$ in CEU and CHB + CHS, two-sided bootstrap-like test). The possibility that Neanderthal introgression may play a role in immune disorders such as allergic diseases is a tantalizing direction for future research.

Meanwhile, a number of tests were nominally significant at the $p = .05$ level in both populations and may be of interest for future research. In particular, we found that loci associated with sodium in urine may be less likely to be introgressed in both populations (nominal $p = .013$ in CEU, nominal $p = .0035$ in CHB + CHS), as are loci associated with the mean time to correctly identify matches, a test in which subjects attempted to quickly determine whether abstract symbols matched (Miller et al., 2016; nominal $p = .0061$ in CEU, nominal $p = .0229$ in CHB + CHS). Some sets of loci may be more likely to be introgressed in both populations: loci associated with vaginal/uterine prolapse (nominal $p = .0049$ in CEU, nominal $p = .174$ in CHB + CHS; loci associated with being treated with glyceryl trinitrate, a drug used to treat heart disease (nominal $p = .0224$ in CEU, nominal $p = .0017$ in CHB + CHS); loci associated with being treated with budesonide, a steroid used to treat asthma, COPD, allergies and Crohn's disease (nominal $p = .0118$ in CEU, nominal $p = .0161$ in CHB + CHS); loci associated with being treated with cardioplen, another drug used to treat heart disease (nominal $p = .0205$ in CEU, nominal $p = .0082$ in CHB + CHS); loci associated with being treated with diclomax, a drug used to treat inflammation, for example in rheumatoid arthritis (nominal $p = .0186$ in CEU, nominal $p = .0106$ in CHB + CHS); and loci associated with living in small towns (nominal $p = .011$ in CEU, nominal $p = .0233$ in CHB + CHS). We again urge caution in interpreting these results, due to the multiple testing burden making the above results not statistically significant.

3.3 | Causes of selection against Neanderthal introgression

As mentioned in the Introduction, Sankararaman et al. (2014) hypothesized DMIs (particularly in male hybrids) to be a cause of selection against Neanderthal ancestry in modern humans. This hypothesis was motivated by noting significant depletion of introgression in testes-expressed genes, as well as a substantial reduction of Neanderthal ancestry on the X-chromosome, which had been associated with DMIs in *Drosophila* (Presgraves, 2008). Our results

obtained using DICAL-ADMIX also show a reduction in Neanderthal ancestry on the X-chromosome, but this reduction can be explained without appealing to DMIs, as we will discuss in more detail in Section 3.4. Here, we focus on global features of the genome and how they relate to potential DMIs.

Juric et al. (2016) and Harris and Nielsen (2016) recently independently proposed that selection acts against Neanderthal ancestry due to a higher mutational load in Neanderthals rather than DMIs. In their study, Juric et al. (2016) developed a likelihood method to explicitly infer the strength of selection against Neanderthal introgression based on the introgression maps obtained by Sankararaman et al. (2014). The authors estimated a selection coefficient for deleterious exonic Neanderthal alleles around -3×10^{-4} for the autosomes. As the estimated coefficient is on the order of the inverse of the effective population size in humans, they hypothesized that the deleterious alleles could have accumulated as a result of the small long-term effective population size in Neanderthals (Prüfer et al., 2014), which reduced the efficacy of selection in this population. When these deleterious alleles entered the larger human population through introgression, they were subjected to more efficient selection, and this led to the observed widespread selection against Neanderthal alleles. The authors use simulations to confirm that the population size history of Neanderthals could have indeed allowed for the accumulation of deleterious alleles on the observed order of magnitude. Harris and Nielsen (2016) arrived at similar conclusions while studying the strength of selection against Neanderthal introgression using forward simulations of autosomal genetic material.

In an attempt to disentangle the DMI and mutational load hypotheses, we performed a number of statistical tests. As detailed below, we found evidence that there was selection against Neanderthal ancestry, but could not find statistically significant evidence that the selection was due to DMIs. We interpret this as evidence in favour of the mutational load hypothesis, although it is possible that more data or more powerful statistical tests may show evidence in favour of the DMI hypothesis. In particular, our tests are underpowered to detect individual DMIs and would not be able to detect if there were a small number of strong DMIs. On the other hand, if there are relatively few DMIs, then they would not be able to explain the reduction in Neanderthal ancestry across the whole genome.

We repeatedly used a test similar to bootstrapping, which we describe presently and refer to subsequently as the *bootstrap-like test*. When performing hypothesis tests, we must account for the spatial correlation of both our introgression calls and many genomic features of interest (e.g., gene locations or local recombination rates). We also would like to account for uncertainty in the introgression calls themselves. To this end, we left the genomic features of interest in place and then sampled new introgression calls for each chromosome by drawing, with replacement, nonoverlapping 5 Mb segments of our original introgression calls from the same chromosome. We then recalculated our test statistic using this resampled set of introgression calls. Repeating this resampling procedure many times provided an approximate empirical distribution of our test

statistic under the null hypothesis of no association between our introgression calls and the genomic feature of interest. We then used this distribution to compute approximate p-values. For all of the tests presented below, we again restricted our analyses to the 500 bp windows where no more than half of the bases were masked by the 1000 Genomes strict mappability mask (The 1000 Genomes Project Consortium, 2012).

To begin, we looked into whether selection against Neanderthal ancestry has occurred. First, note that the admixture proportion has previously been estimated as 3% (Green et al., 2010; Juric et al., 2016). If there was no subsequent "dilution" of Neanderthal ancestry (Sankararaman et al. (2016), but see Vernot and Akey (2015), and see Slatkin and Racimo (2016) for a comprehensive review), then under neutrality, we would expect about 3% ancestry on average in present-day populations, and we would expect about half of chromosomes to have more than 3% Neanderthal introgression on average and about half of chromosomes to have <3% introgression. Yet, no chromosome in either CEU or CHB + CHS has, on average, more than 3.29% introgression, and most chromosomes have <1.5% average Neanderthal ancestry. Thus, under this simple null model, we can reject neutrality ($p = 2.4 \times 10^{-7}$ in CEU, $p = 5.7 \times 10^{-6}$ in CHB + CHS, two-sided sign test, $n = 23$) in both CEU and CHB + CHS. While the above test assumed that the admixture proportion was 3%, we would be able to reject the null hypothesis of neutrality for any admixture proportion greater than 1.48% in CEU or 1.96% in CHB + CHS, both of which are much lower than the findings of the previous studies discussed above. Indeed, Juric et al. (2016) report a confidence a 95% confidence interval of [3.22%, 3.52%] for the admixture proportion in CEU and [3.45%, 3.86%] for the admixture proportion in CHB + CHS (Juric et al., 2016). Using the estimates of Juric et al. (2016) to infer that selection has occurred is a somewhat circular argument because Juric et al. (2016) used the assumption of selection against Neanderthal ancestry to infer their admixture proportions. Yet, an ancient European individual has been found with >6% Neanderthal ancestry, and a number of other ancient Europeans have been found with >3% Neanderthal ancestry without the assumption of selection against Neanderthal ancestry (Fu et al., 2015). These ancient genomes together with our findings suggest that there has been a significant reduction in Neanderthal ancestry.

To explore if this reduction in Neanderthal ancestry is more pronounced in genic regions, we compared the mean (across individuals and loci) frequency of introgression in regions marked as exons in the RefSeq annotation (O'Leary et al., 2016) to the chromosome-wide mean. The results were largely concordant when we considered transcripts or coding sequences instead of exons, so we present below only results based on exons. For CEU, we found that in 13 of the 23 chromosomes, there is less introgression in genic regions than in the rest of the chromosome, which is not statistically significant ($p = .678$, two-sided sign test, $n = 23$). Likewise, for CHB + CHS, we again found that 13 chromosomes have less introgression in genic regions, which are also not statistically significant ($p = .678$, two-sided sign test, $n = 23$). Furthermore, only nine chromosomes

showing a reduction in introgression at exons were shared between the two populations, which are not statistically significant ($p = .09$, two-sided permutation test). We also performed our bootstrap-like test to see if there was any significant decrease in genic regions on any chromosome, but we did not find any significant results (unadjusted $p > .05$) on any chromosome in either population, except for the X-chromosome in CHB + CHS (unadjusted $p = .02$) which can be explained simply by the multiple testing burden. We interpret these results as indicating that either selection against Neanderthal introgression is fairly weak if it is acting on most or all genes, or has only acted on a subset of genes. It is also possible that selection against Neanderthal introgression has acted on genomic elements other than exons, such as regulatory elements.

Meanwhile, we found evidence that a measure of conservation, phastCONS (Siepel et al., 2005), was significantly negatively correlated with mean introgression at a given locus (Spearman's $\rho = -0.029$, $p = .003$ in CEU and $\rho = -.024$, $p = .043$ in CHB + CHS, bootstrap-like test), which indicates that selection was more likely to remove Neanderthal ancestry at highly conserved loci. We also tested if proportion of Neanderthal ancestry was positively correlated with local population-scaled recombination rate (as inferred by Myers et al. (2005)), which would be suggestive of selection against Neanderthal ancestry because regions of high recombination would be more likely to separate neutral regions of Neanderthal ancestry from linked deleterious regions, an idea recently explored elegantly and in more detail by Schumer et al. (2017). Similar to Schumer et al. (2017), we found a positive association between local population-scaled recombination rate and frequency of introgression (Spearman's $\rho = 0.055$, $p < .001$ in CEU and $\rho = .054$, $p < .001$ in CHB + CHS, bootstrap-like test) lending further credence to the hypothesis that selection is acting against certain regions of Neanderthal ancestry.

If DMIs were the cause of selection against introgression, then we would expect that genes that code for proteins with more binding partners would be less likely to be introgressed; each protein–protein interaction (PPI) can be thought of as a possible DMI. To test this hypothesis, we used the PICKLE2.0 PPI network (Gioutlakis, Klapa, & Moschonas, 2017; Klapa, Tsafou, Theodoridis, Tsakalidis, & Moschonas, 2013) and associated a number of binding partners to each gene by counting the number of PPIs in which the protein coded by that gene participates. We found an insignificant correlation between number of binding partners and mean frequency of introgression in both populations (Spearman's $\rho = -.016$, $p = .314$ in CEU, $\rho = -0.003$, $p = .858$ in CHB + CHS, two-sided bootstrap-like test), which provide weak indirect evidence against the DMI hypothesis.

As a more direct test of the DMI hypothesis, we tested whether proteins that interact (according to the PICKLE2.0 PPI network) are more likely to be co-introgressed. In particular, for each gene in the PPI network, we say that that gene is introgressed if any part of any of its exons is in a called introgression tract. For each individual, we then assign a weight to each edge in the PPI network as follows. Let gene A and gene B be the genes that code for the proteins involved

in the interaction corresponding to the edge of interest. If each copy of gene A and gene B (i.e., on autosomes, we assume there are two copies of each gene and on the X-chromosome males have only one copy) has the same ancestry, the edge is assigned a weight of one—in this individual, this interaction is always between proteins from genes of the same ancestry. Meanwhile, if all of the copies of gene A are of one type of ancestry and all of the copies of gene B are of the other type of ancestry, then the edge is assigned a weight of zero—this interaction is never between proteins of the same ancestry. Finally, if either gene has mixed ancestry (i.e., one copy from one ancestry and the other copy from the other ancestry) the edge is assigned a weight of 0.5—in this case, it can be shown that if one randomly selects a copy of gene A and a copy of gene B, then the proteins produced by those copies will have the same ancestry 50% of the time. Thus, these edge weights are the probabilities that compatible proteins interact, assuming that both ancestry types at each locus produce the same amount of protein and the probability that a given protein is involved in a particular interaction does not depend on its ancestry. We then averaged these weights across individuals and across edges in the PPI network to obtain a test statistic. Using this test, we failed to obtain a significant result in either population ($p = .117$, CEU, $p = .647$, CHBS, one-sided permutation test), which again provides some evidence against the DMI hypothesis. If DMIs are not widespread, then this test would likely not have power to detect them, so we are unable to rule out the possibility that there are a small number of DMIs. Yet, a small number of DMIs would be unable to explain the overall reduction in Neanderthal ancestry across the genome.

Taken as a whole, we find that while there has been selection against Neanderthal introgression, for example at highly conserved loci, it seems that the negative selection is likely not due to widespread DMIs, which lends more credence to the mutational load hypothesis. We also note that in contrast to the broad findings presented here, a small number of specific loci have been found to have experienced positive selection for archaic introgression (Racimo, Marnetto, & Huerta-Sanchez, 2017; Racimo, Sankararaman, Nielsen, & Huerta-Sanchez, 2015; Sams et al., 2016).

3.4 | Patterns of introgression on the X-chromosome

To further explore whether selection against introgressed Neanderthal variation differed between autosomes and the X-chromosome, we performed forward simulations in a Wright–Fisher model, focusing on the dynamics of Neanderthal alleles in the modern human population after the introgression event 2,000 generations ago. For the autosomes, we modelled each diploid individual to be comprised of two chromosomes. Each chromosome consisted of 5,000 loci, and recombination could act between these loci. The recombination rate was calibrated such that this corresponds to a 150 Mbp chromosome with a recombination rate of 1.25×10^{-8} per-generation per base pair. Similar to Harris and Nielsen (2016), the population size was set to $N = 1,860$ for the first 900 generations, followed by an instantaneous decrease to $N = 1,032$ with subsequent exponential growth at 0.38% per generation. For computational reasons, we limited the population size to $N = 10,000$, which is reached roughly 400 generations before present. In the generation immediately after the introgression event, the chromosomes of 97% of the individuals in the population carry modern human alleles at all 5,000 loci on both chromosomes, and 3% carry Neanderthal alleles, representing the introgressed individuals. In the subsequent generations, the fitness of an individual is $(1-s)^D$, with selection coefficient s, and D denoting the number of Neanderthal alleles that a diploid individual carries. Figure 8a depicts the amount of Neanderthal introgression measured in the autosomes in the CEU and the CHB + CHS populations, as well as the amount of Neanderthal introgression in a sample of individuals at present from populations simulated with different values for s, repeated 16 times. These simulations suggest that a selection coefficient on the order of $s = -2 \times 10^{-5}$ is more than sufficient to explain the observed reduction in Neanderthal ancestry from the initial proportion of 3% on the autosomes. These results are largely consistent with the estimates of selection against introgression provided in Juric et al. (2016), table 1, where -3×10^{-8} is estimated for the effective selection strength per exonic site. In our simulations, each locus
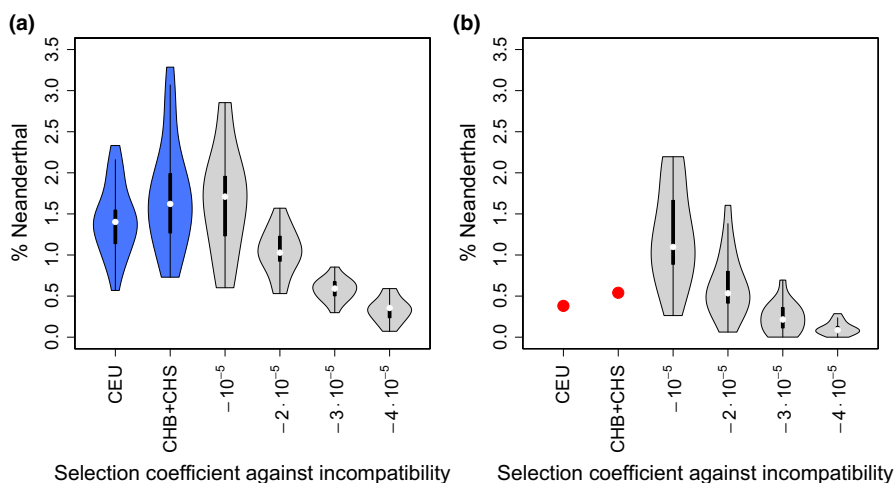
FIGURE 8 Distribution of amount of Neanderthal introgression on the different chromosomes. Given empirically in CEU and CHB + CHS, and sampled from the present-day generation under the Wright–Fisher simulations for different selection coefficients, in 16 replicates. (a) The distributions on the autosomes. (b) The distributions on the X-chromosomes [Colour figure can be viewed at wileyonlinelibrary.com]

corresponds to 30,000 sites. According to the annotation from the UCSC genome browser (https://genome.ucsc.edu/), the genomewide density of exonic sites is 2.8%, and thus, a simulated locus contains roughly 840 exonic sites. The simulated selection strength of $s = -2 \times 10^{-5}$ then corresponds to a selection strength of -2.4×10^{-8} per exonic site.

We also performed simulations for the X-chromosome to see if the strength of selection is similar on the autosomes and X-chromosome. In our simulation, the fitness of males was determined solely by their single X-chromosome, and their Y-chromosome was modelled as selectively neutral. In females, to model X-inactivation, we chose one chromosome randomly to determine the fitness. There are several methods to calibrate the selection coefficient, s. Here, we chose to calibrate s such that, in both females and males, carrying only Neanderthal variants at a certain locus has the same affect on fitness for the X-chromosome as for an autosome. Consequently, to determine the exponent for the fitness in females, the number of Neanderthal alleles on the active chromosome is multiplied by two. This calibration effectively corresponds to an additive selection model in females. In males, the number of Neanderthal alleles on the single X-chromosome is also multiplied by two to obtain the exponent. Note that Juric et al. (2016) used a different calibration, where selection in males is half as strong. Figure 8b shows the Neanderthal proportions on the X-chromosome in both populations, and the results of the simulations for different values of s. We observe that using this calibration for the selection coefficient, for the same strength of selection, the amount of Neanderthal ancestry in the population at present is reduced on the X-chromosome compared with the autosomes. While this might seem at odds with the reduced effective population size for the X-chromosome, and hence a lower efficacy of selection, it can be explained by the fact that genetic variants in males have a stronger impact than they would in a diploid autosomal population of reduced effective size. Moreover, we observe in Figure 8b that a selection coefficient not much stronger than $s = -2 \times 10^{-5}$ can result in the reduction of Neanderthal introgression observed on the X-chromosomes in the 1000 Genomes data.

To explore the possibility of hybrid male infertility, we modified the simulations for the X-chromosome as follows. In addition to the global selection with coefficient s against Neanderthal alleles at all loci, we designated 0.5% of the 5000 loci to be *incompatibility* loci. These *incompatibility* loci only affect fitness in male individuals. The fitness is multiplied by $(1 - s_I)^{\frac{4C}{M}(M-C)}$, where s_I is the selection coefficient against incompatibility, M is the total number of *incompatibility* loci and C is the number of Neanderthal alleles a male individual carries at these loci. The exponent is proportional to the number of incompatible pairs. Thus, if an individual carries only modern human or only Neanderthal alleles at these loci, the exponent is zero, and the fitness is not affected. The exponent equals its maximal value of $\frac{M}{2}$ when $C = \frac{M}{2}$, that is, half of the *incompatibility* loci carry the Neanderthal allele, and the other half carries the human allele. Figure 9 depicts the results of the simulations for $s = -10^{-5}$, and different *incompatibility* selection coefficients s_I. Note that the order of
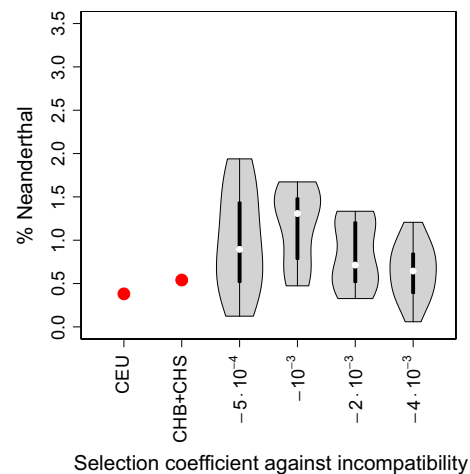

FIGURE 9 Distribution of amount of Neanderthal introgression on the X-chromosome. Given empirically in CEU and CHB + CHS, and sampled from the present day generation under the Wright–Fisher simulations for different models of male hybrid infertility, in 16 replicates [Colour figure can be viewed at wileyonlinelibrary.com]

magnitude of s_I is higher than s, because it is acting on fewer loci. The simulations show that a mechanism like this type of hybrid incompatibility could indeed decrease the introgression further than global weak selection against Neanderthal variants by itself, although such an explanation is not necessary to fit the observed levels of introgression.

4 | DISCUSSION

In this study, we introduced a modification of the method DICAL 2.0, which was developed by Steinrücken et al. (2015) to infer complex demographic histories from full-genomic sequence data. We applied this modification (DICAL-ADMIX) to detect tracts of genetic material in modern non-African individuals that introgressed into the population when non-Africans and Neanderthals exchanged genetic material about 2,000 generations ago. We demonstrated in an extensive simulation study that DICAL-ADMIX can accurately and efficiently detect tracts of Neanderthal introgression. Furthermore, we applied DICAL-ADMIX to detect introgression in the individuals sampled from the CEU, CHB and CHS populations as part of the 1000 Genomes Project (The 1000 Genomes Project Consortium, 2012). We exhibited some of the methodological and empirical differences between DICAL-ADMIX and previous results reported by Sankararaman et al. (2014) and by Vernot and Akey (2014). While they are generally in good agreement, we observed some differences. This highlights the importance of the development of different methodologies to generate a consensus. We also reported some of the functional implications of introgression, which confirms previously reported findings of widespread selection against introgression, enrichment of Neanderthal introgression in certain classes of genes, but a general signal of depleted Neanderthal introgression in conserved regions of the genome.

However, the role of the X-chromosome remains intriguing. As in previous studies, we observe a substantially lower amount of Neanderthal introgression on the X-chromosome compared with the autosomes. Sankararaman et al. (2014) hypothesized that this reduction is due to DMIs reducing male fertility, further supported by significantly reduced Neanderthal introgression in genes expressed in testes. However, we do not find significant evidence for DMIs, and the GO-term enrichment analyses did not reveal any patterns of enrichment or depletion of Neanderthal ancestry in genes related to spermatogenesis, the testes or infertility more broadly.

Additionally, compared to modern humans and Neanderthals (separated by tens of thousands of generations), the species in which DMIs have been observed are substantially more diverged—for example about 20 million generations between *D. simulans* and *D. melanogaster* (Li, Satta, & Takahata, 1999); 200 to 500 thousand generations between *M. guttatis* and *M. nasutus* (Brandvain, Kenney, Flagel, Coop, & Sweigart, 2014); and 750 thousand to 5 million generations between *A. californiense* and *A. tigrinum mavortium* (Fitzpatrick et al., 2010; Shaffer & McKnight, 1996). The divergence time separating modern humans and Neanderthals is only about a factor of two older than the divergence time between the most diverged human populations (e.g., as inferred by Veeramah et al. (2012) and Gronau, Hubisz, Gulko, Danko, and Siepel (2011) using an updated mutation rate as discussed in Scally and Durbin (2012)), and no DMIs are known to occur in admixtures of modern human populations, raising the question of how such incompatibilities between modern humans and Neanderthals could have arisen so quickly.

Other studies (Harris & Nielsen, 2016; Juric et al., 2016) and our simulations suggest that only a moderate strength of selection is required to explain the observed reduction on the X-chromosome. However, the evidence that has been collected to date does not seem to be sufficient to fully characterize the importance of the X-chromosome. Resolving these questions will require a more comprehensive analysis of larger samples of contemporary genetic data like the Simons Genome Diversity Project (Mallick et al., 2016) and the individuals from Phase III of the 1000 Genomes Project (The 1000 Genomes Project Consortium, 2012). Moreover, additional high-quality data for hominin sister groups (Meyer et al., 2012; Prüfer et al., 2017) will improve the detection of introgression. Detecting introgression in genetic samples of ancient modern humans (Mathieson et al., 2015) will also allow resolution of the evolutionary trajectory of introgressed genetic material over time. Incorporating the distribution of tracts on an individual level and different models of diploid selection into the inference frameworks will improve the inference of the strength of selection and allow reliable testing of different models. Additionally, a better understanding of the evolution of incompatibilities and more careful investigation of the gene content on the X-chromosome will help shed light on the role of the X-chromosome in the Neanderthal introgression landscape. In general, maps of introgressed Neanderthal and Denisovan ancestry will facilitate the interpretation of patterns of human genomic variation, and further the understanding of how archaic introgression influenced the trajectory of human evolution.

DATA ACCESSIBILITY

The marginal posterior probabilities of Neanderthal introgression for the CEU, CHB and CHS haplotypes computed using DICAL-ADMIX are available at http://dical-admix.sourceforge.net/

REFERENCES

Brandvain, Y., Kenney, A. M., Flagel, L., Coop, G., & Sweigart, A. L. (2014). Speciation and introgression between *Mimulus nasutus* and *Mimulus guttatus*. *PLOS Genetics*, *10*(6), 1–15.

Brideau, N. J., Flores, H. A., Wang, J., Maheshwari, S., Wang, X., & Barbash, D. A. (2006). Two dobzhansky-muller genes interact to cause hybrid lethality in drosophila. *Science*, *314*(5803), 1292–1295. https://doi.org/10.1126/science.1133953

Dannemann, M., Prüfer, K., & Kelso, J. (2017). Functional implications of Neandertal introgression in modern humans. *Genome Biology*, *18*(1), 61. https://doi.org/10.1186/s13059-017-1181-7

Dobzhansky, T. (1936). Studies on hybrid sterility. II. localization of sterility factors in drosophila pseudoobscura hybrids. *Genetics*, *21*(2), 113–135.

Eden, E., Lipson, D., Yogev, S., & Yakhini, Z. (2007). Discovering motifs in ranked lists of DNA sequences. *PLOS Computational Biology*, *3*(3), 1–15.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., & Yakhini, Z. (2009). GORILLA: A tool for discovery and visualization of enriched go terms in ranked gene lists. *BMC Bioinformatics*, *10*(1), 48. https://doi.org/10.1186/1471-2105-10-48

Fishman, L., & Willis, J. H. (2001). Evidence for dobzhansky-muller incompatibilites contributing to the sterility of hybrids between *Mimulus guttatus* and *M. nasutus*. *Evolution*, *55*(10), 1932–1942. https://doi.org/10.1111/j.0014-3820.2001.tb01311.x

Fitzpatrick, B. M. (2008). Dobzhansky-muller model of hybrid dysfunction supported by poor burst-speed performance in hybrid tiger salamanders. *Journal of Evolutionary Biology*, *21*(1), 342–351. https://doi.org/10.1111/j.1420-9101.2007.01448.x

Fitzpatrick, B. M., Johnson, J. R., Kump, D. K., Smith, J. J., Voss, S. R., & Shaffer, H. B. (2010). Rapid spread of invasive genes into a threatened native species. *Proceedings of the National Academy of Sciences*, *107*(8), 3606–3610. https://doi.org/10.1073/pnas.0911802107

Fu, Q., Hajdinjak, M., Moldovan, O. T., Constantin, S., Mallick, S., Skoglund, P., . . . Pääbo, S. (2015). An early modern human from Romania with a recent Neanderthal ancestor. *Nature*, *524*, 216–219. https://doi.org/10.1038/nature14558

Gioutlakis, A., Klapa, M. I., & Moschonas, N. K. (2017). PICKLE 2.0: A human protein-protein interaction meta-database employing data integration via genetic information ontology. *PLoS One*, *12*(10), 1–17.

Gittelman, R. M., Schraiber, J. G., Vernot, B., Mikacenic, C., Wurfel, M. M., & Akey, J. M. (2016). Archaic hominin admixture facilitated adaptation to out-of-africa environments. *Current Biology*, *26*(24), 3375–3382. https://doi.org/10.1016/j.cub.2016.10.041

Global Biobank Engine (2017). Retrieved from http://gbe.stanford.edu/.

Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., ... Pääbo, S. (2010). A draft sequence of the Neandertal genome. *Science*, *328*(5979), 710–722. https://doi.org/10.1126/science.1188021

Gronau, I., Hubisz, M. J., Gulko, B., Danko, C. G., & Siepel, A. (2011). Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics*, *43*, 1031. https://doi.org/10.1038/ng.937

Harris, K., & Nielsen, R. (2016). The genetic cost of Neanderthal introgression. *Genetics*, *203*(2), 881–891. https://doi.org/10.1534/genetics.116.186890

Jouganous, J., Long, W., Ragsdale, A. P., & Gravel, S. (2017). Inferring the joint demographic history of multiple populations: Beyond the diffusion approximation. *Genetics*, *206*(3), 1549–1567. https://doi.org/10.1534/genetics.117.200493

Juric, I., Aeschbacher, S., & Coop, G. (2016). The strength of selection against Neanderthal introgression. *PLOS Genetics*, *12*(11), e1006340. https://doi.org/10.1371/journal.pgen.1006340

Kelleher, J., Etheridge, A. M., & McVean, G. (2016). Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLOS Computational Biology*, *12*(5), e1004842. https://doi.org/10.1371/journal.pcbi.1004842

Klapa, M. I., Tsafou, K., Theodoridis, E., Tsakalidis, A., & Moschonas, N. K. (2013). Reconstruction of the experimentally supported human protein interactome: What can we learn? *BMC Systems Biology*, *7*(1), 96. https://doi.org/10.1186/1752-0509-7-96

Kong, A., Thorleifsson, G., Gudbjartsson, D. F., Masson, G., Sigurdsson, A., Jonasdottir, A., ... Stefansson, K. (2010). Fine-scale recombination rate differences between sexes, populations and individuals. *Nature*, *467*(7319), 1099–1103. https://doi.org/10.1038/nature09525

Li, Y.-J., Satta, Y., & Takahata, N. (1999). Paleo-demography of the *Drosophila melanogaster* subgroup: Application of the maximum likelihood method. *Genes & Genetic Systems*, *74*(4), 117–127. https://doi.org/10.1266/ggs.74.117

Li, N., & Stephens, M. (2003). Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics*, *165*(4), 2213–2233.

Mallick, S., Li, H., Lipson, M., Mathieson, I., Gymrek, M., Racimo, F., ... Reich, D. (2016). The simons genome diversity project: 300 genomes from 142 diverse populations. *Nature*, *538*(7624), 201–206. https://doi.org/10.1038/nature18964

Malnic, B., Godfrey, P. A., & Buck, L. B. (2004). The human olfactory receptor gene family. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(8), 2584–2589. https://doi.org/10.1073/pnas.0307882100

Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S. A., ... Reich, D. (2015). Genome-wide patterns of selection in 230 ancient eurasians. *Nature*, *528*(7583), 499–503. https://doi.org/10.1038/nature16152

Meyer, M., Kircher, M., Gansauge, M.-T., Li, H., Racimo, F., Mallick, S., ... Pääbo, S. (2012). A high-coverage genome sequence from an archaic Denisovan individual. *Science*, *338*(6104), 222–226. https://doi.org/10.1126/science.1224344

Miller, K. L., Alfaro-Almagro, F., Bangerter, N. K., Thomas, D. L., Yacoub, E., Xu, J., ... Smith, S. M. (2016). Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nature Neuroscience*, *19*, 1523–1536. https://doi.org/10.1038/nn.4393

Myers, S., Bottolo, L., Freeman, C., McVean, G., & Donnelly, P. (2005). A fine-scale map of recombination rates and hotspots across the human genome. *Science*, *310*(5746), 321–324. https://doi.org/10.1126/science.1117196

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., ... Pruitt, K. D. (2016). Reference sequence (refseq) database at ncbi: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, *44*(D1), D733–D745. https://doi.org/10.1093/nar/gkv1189

Orr, H. A. (1995). The population genetics of speciation: The evolution of hybrid incompatibilities. *Genetics*, *139*(4), 1805–1813.

Paten, B., Herrero, J., Beal, K., Fitzgerald, S., & Birney, E. (2008). Enredo and pecan: Genome-wide mammalian consistency-based multiple alignment with paralogs. *Genome Research*, *18*(11), 1814–1828. https://doi.org/10.1101/gr.076554.108

Paul, J. S., & Song, Y. S. (2010). A principled approach to deriving approximate conditional sampling distributions in population genetics models with recombination. *Genetics*, *186*(1), 321–338. https://doi.org/10.1534/genetics.110.117986

Paul, J. S., Steinrücken, M., & Song, Y. S. (2011). An accurate sequentially markov conditional sampling distribution for the coalescent with recombination. *Genetics*, *187*(4), 1115–1128. https://doi.org/10.1534/genetics.110.125534

Plagnol, V., & Wall, J. D. (2006). Possible ancestral structure in human populations. *PLOS Genetics*, *2*(7), e105. https://doi.org/10.1371/journal.pgen.0020105

Presgraves, D. C. (2008). Sex chromosomes and speciation in drosophila. *Trends Genetics*, *24*(7), 336–343. https://doi.org/10.1016/j.tig.2008.04.007

Prüfer, K., de Filippo, C., Grote, S., Mafessoni, F., Korlevic, P., Hajdinjak, M., ... Pääbo, S. (2017). A high-coverage Neandertal genome from Vindija Cave in Croatia. *Science*, *358*, 655–658.

Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., ... Pääbo, S. (2014). The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*, *505*(7481), 43–49. https://doi.org/10.1038/nature12886

Racimo, F., Marnetto, D., & Huerta-Sanchez, E. (2017). Signatures of archaic adaptive introgression in present-day human populations. *Molecular Biology and Evolution*, *34*(2), 296–317.

Racimo, F., Sankararaman, S., Nielsen, R., & Huerta-Sanchez, E. (2015). Evidence for archaic adaptive introgression in humans. *Nature Reviews Genetics*, *16*, 359. https://doi.org/10.1038/nrg3936

Rogers, R. L. (2015). Chromosomal rearrangements as barriers to genetic homogenization between archaic and modern humans. *Molecular Biology and Evolution*, *32*(12), 3064–3078.

Sams, A. J., Dumaine, A., Nédélec, Y., Yotova, V., Alfieri, C., Tanner, J. E., ... Barreiro, L. B. (2016). Adaptively introgressed Neandertal haplotype at the oas locus functionally impacts innate immune responses in humans. *Genome Biology*, *17*(1), 246. https://doi.org/10.1186/s13059-016-1098-6

Sankararaman, S., Mallick, S., Dannemann, M., Prufer, K., Kelso, J., Paabo, S., ... Reich, D. (2014). The genomic landscape of Neanderthal ancestry in present-day humans. *Nature*, *507*(7492), 354–357. https://doi.org/10.1038/nature12961

Sankararaman, S., Mallick, S., Patterson, N., & Reich, D. (2016). The combined landscape of Denisovan and Neanderthal ancestry in present-day humans. *Current Biology*, *26*(9), 1241–1247. https://doi.org/10.1016/j.cub.2016.03.037

Sankararaman, S., Patterson, N., Li, H., Pääbo, S., & Reich, D. (2012). The date of interbreeding between Neandertal and modern humans. *PLOS Genetics*, *8*(10), e1002947. https://doi.org/10.1371/journal.pgen.1002947

Scally, A., & Durbin, R. (2012). Revising the human mutation rate: Implications for understanding human evolution. *Nature Reviews Genetics*, *13*(10), 745–753. https://doi.org/10.1038/nrg3295

Schumer, M., Xu, C., Powell, D., Durvasula, A., Skov, L., Holland, C., ... Przeworski, M. (2017). Natural selection interacts with the local recombination rate to shape the evolution of hybrid genomes. *bioRxiv*, https://doi.org/10.1101/212407.

Shaffer, H. B., & McKnight, M. L. (1996). The polytypic species revisited: Genetic differentiation and molecular phylogenetics of the tiger

salamander *ambystoma tigrinum* (amphibia: Caudata) complex. *Evolution*, *50*(1), 417–433. https://doi.org/10.1111/j.1558-5646.1996.tb04503.x

Siepel, A., Bejerano, G., Pedersen, J. S., Hinrichs, A. S., Hou, M., Rosenbloom, K., ... Haussler, D. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Research*, *15*(8), 1034–1050. https://doi.org/10.1101/gr.3715005

Simonti, C. N., Vernot, B., Bastarache, L., Bottinger, E., Carrell, D. S., Chisholm, R. L., ... Capra, J. A. (2016). The phenotypic legacy of admixture between modern humans and Neandertal. *Science*, *351*(6274), 737–741. https://doi.org/10.1126/science.aad2149

Slatkin, M., & Racimo, F. (2016). Ancient dna and human history. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(23), 6380–6387. https://doi.org/10.1073/pnas.1524306113

Steinrücken, M., Kamm, J., & Song, Y. (2015). Inference of complex population histories using wholegenome sequences from multiple populations. Retrieved from https://doi.org/10.1101/026591

Steinrücken, M., Paul, J. S., & Song, Y. S. (2013). A sequentially Markov conditional sampling distribution for structured populations with migration and recombination. *Theoretical Population Biology*, *87*, 51–61. https://doi.org/10.1016/j.tpb.2012.08.004

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., ... Collins, R. (2015). Uk biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLOS Medicine*, *12*(3), 1–10.

The 1000 Genomes Project Consortium (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature*, *491*(7422), 56–65.

Veeramah, K. R., Wegmann, D., Woerner, A., Mendez, F. L., Watkins, J. C., Destro-Bisol, G., ... Hammer, M. F. (2012). An early divergence of khoesan ancestors from those of other modern humans is supported by an abc-based analysis of autosomal resequencing data. *Molecular Biology and Evolution*, *29*(2), 617–630. https://doi.org/10.1093/molbev/msr212

Vernot, B., & Akey, J. M. (2014). Resurrecting surviving Neandertal lineages from modern human genomes. *Science*, *343*(6174), 1017–1021. https://doi.org/10.1126/science.1245938

Vernot, B., & Akey, J. M. (2015). Complex history of admixture between modern humans and Neandertal. *The American Journal of Human Genetics*, *96*(3), 448–453. https://doi.org/10.1016/j.ajhg.2015.01.006

Vernot, B., Tucci, S., Kelso, J., Schraiber, J. G., Wolf, A. B., Gittelman, R. M., ... Akey, J. M. (2016). Excavating Neandertal and Denisovan dna from the genomes of melanesian individuals. *Science*, *352*(6282), 235–239.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.