bioRxiv preprint doi: https://doi.org/10.1101/343087; this version posted June 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Spectrum of Neandertal introgression across modern-day humans indicates multiple episodes of human-Neandertal interbreeding

Fernando A. Villanea and Joshua G. Schraiber

Compiled on June 8, 2018.

1 Abstract

Neandertals and anatomically modern humans overlapped geographically for a period of over 30,000 years following human migration out of Africa. During this period, Neandertals and humans interbred, as evidenced by Neandertal portions of the genome carried by non-African individuals today. A key observation is that the proportion of Neandertal ancestry is ∼12% higher in East Asian individuals relative to European individuals. Here, we explore various demographic models that could explain this observation. These include distinguishing between a single admixture event and multiple Neandertal contributions to either population, and the hypothesis that reduced Neandertal ancestry in modern Europeans resulted from more recent admixture with a ghost population that lacked a Neandertal ancestry component (the "dilution" hypothesis). In order to summarize the asymmetric pattern of Neandertal allele frequencies, we compile the joint fragment frequency spectrum (FFS) of European and East Asian Neandertal fragments and compare it to both analytical theory and data simulated under various models of admixture. Using maximum likelihood and machine learning, we found that a simple model of a single admixture does not fit the empirical data, and instead favor a model of multiple episodes of gene flow into both European and East Asian populations. These findings indicate more long-term, complex interaction between humans and Neandertals than previously appreciated.

2 Introduction

When anatomically modern humans dispersed out of Africa they encountered and hybridized with Neandertals [Green et al., 2010]. The Neandertal component of the modern human genome is ubiquitous in non-African populations, and yet is quantitatively small, representing on average only ∼2% of those genomes [Green et al., 2010, Prüfer et al., 2017]. This pattern of Neandertal ancestry in modern human genomes was initially interpreted as

bioRxiv preprint doi: https://doi.org/10.1101/343087; this version posted June 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

evidence of a single period of admixture, occurring shortly after the out-of-Africa bottleneck [Green et al., 2010, Sankararaman et al., 2012]. However, subsequent research showed that Neandertal ancestry is higher by ∼12% in modern East Asian individuals relative to modern European individuals [Prüfer et al., 2017, Meyer et al., 2012, Wall et al., 2013].

A key point of contention regarding elevated Neandertal ancestry in East Asians centers on whether there were one or multiple episodes of gene flow between modern humans and Neandertals. Sankararaman et al. [2014] proposed that differences in the level of Neandertal ancestry in East Asian individuals could be explained by their lower ancestral effective population size relative to Europeans, which would reduce the efficacy of purifying selection against deleterious Neandertal alleles [Harris and Nielsen, 2016]. However, Kim and Lohmueller [2015] found that differences in the strength of purifying selection and population size are unlikely to explain the enrichment of Neandertal ancestry in East Asian individuals. This conclusion was further strengthened by Juric et al. [2016].

Another hypothesis consistent with a single episode of gene flow is that Neandertal ancestry in modern Europeans was diluted by one of the populations that mixed to create modern Europeans [Lazaridis et al., 2014, 2016]. This population, dubbed 'Basal Eurasian', possibly migrated out of Africa separately than the population receiving the pulse of Neandertal gene flow, and thus had little to no Neandertal ancestry.

One the other hand, admixture may have occurred multiple times; the first pulse of Neandertal gene flow into the population ancestral to East Asians and Europeans was supplemented by additional episodes after both populations had diverged [Vernot and Akey, 2015, Vernot et al., 2016]. The finding of an individual from Peştera cu Oase, Romania with a recent Neandertal ancestor provides direct evidence of additional episodes of interbreeding, although this individual is unlikely to have contributed to modern-day diversity [Fu et al., 2015]. Nonetheless, declining levels of Neandertal ancestry in samples going back tens of thousands of years has raised doubts about substantial additional episodes of gene flow from Neandertals into modern humans [Fu et al., 2016].

Here, we study the the asymmetry in the pattern of Neandertal introgression in modern human genomes between individuals of East Asian and European ancestry. We summarize the asymmetric distribution of Neandertal ancestry tracts in the East Asian and European individuals in the 1000 Genomes Project panel in a joint fragment frequency spectrum (FFS) matrix. We fit analytical models using maximum likelihood to explain the distribution of fragments in European and Asian individuals marginally. Then, we compare the joint FFS to the output of genomic data simulated under specific models of admixture between Neandertals and AMH. Our results support a complex model of admixture, with early admixture occurring before the diversification of European and East Asian lineages, and secondary episodes of gene flow into both populations independently.

bioRxiv preprint doi: https://doi.org/10.1101/343087; this version posted June 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

3 Results

We constructed the joint fragment frequency spectrum by analyzing published datasets of Neandertal fragment calls in 1000 genomes individuals [Sankararaman et al., 2014, Steinrücken et al., 2018]. To avoid complications due to partially overlapping fragments and difficulties calling the edges of fragments, we computed fragment frequencies by sampling a single site every 100kb and asking how many haplotypes were introgressed with confidence above a certain cutoff (Figure 1). Our main results make use fragments called with posterior probability of 0.45 in the Steinrücken dataset, although we verified robustness across a range of cutoffs and between datasets (Supplementary Material).

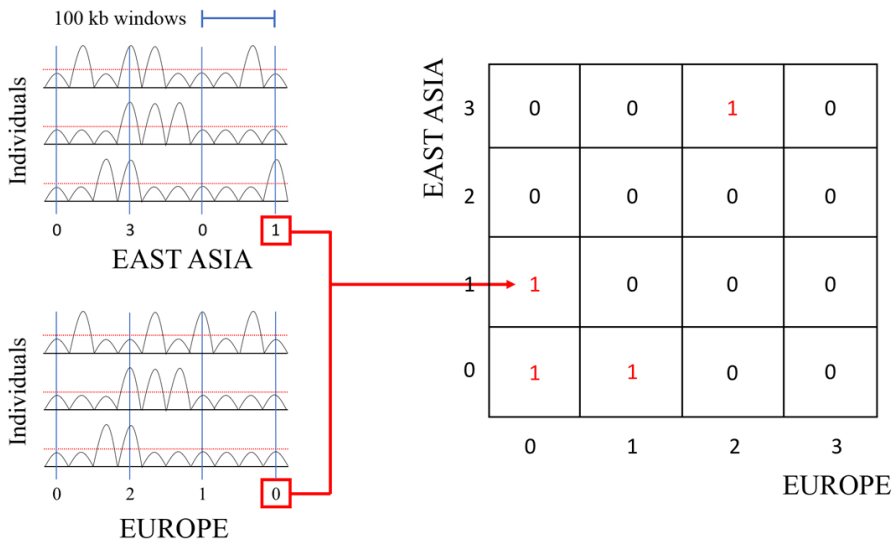


Figure 1: Graphic representation of the process to make an FFS based on the posterior probability calls from Steinrücken et al. [2018]. At a single position every 100 kb, introgression is assigned for individuals presenting a posterior probability above the global cut-off. The count of individuals between East Asian and European samples determines the cell in the FFS where that SNP is counted

We first developed analytic theory to understand what the FFS would look like in each population separately under different demographic models. To our surprise, we found that when looking only at the marginal distribution of introgressed fragment frequencies, the one pulse model and the dilution model are not statistically identifiable (Supplementary Material). However, the two pulse model is identifiable. With this in mind, we used a maximum likelihood procedure to fit a one pulse and a two pulse model to the European and East Asian marginal spectra (Methods), and found strong support for the two pulse model in both cases ($\Lambda = 193.91$ in East Asians, $\Lambda = 212.64$ in Europeans; Figure 2a

bioRxiv preprint doi: https://doi.org/10.1101/343087; this version posted June 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and b). A subsequent goodness-of-fit test strongly rejected the fit of the one pulse model ($p = 2 \times 10^{-26}$ in East Asians, $p = 0.0$ in Europeans; χ^2 goodness of fit test) but could not reject the fit of the two pulse model in either populations ($p = 1$ in East Asians, $p = 0.95$ in Europe; χ^2 goodness of fit test). Thus, we concluded from analyzing each population in isolation that that the history of admixture was complex, and involved multiple matings with Neandertals.

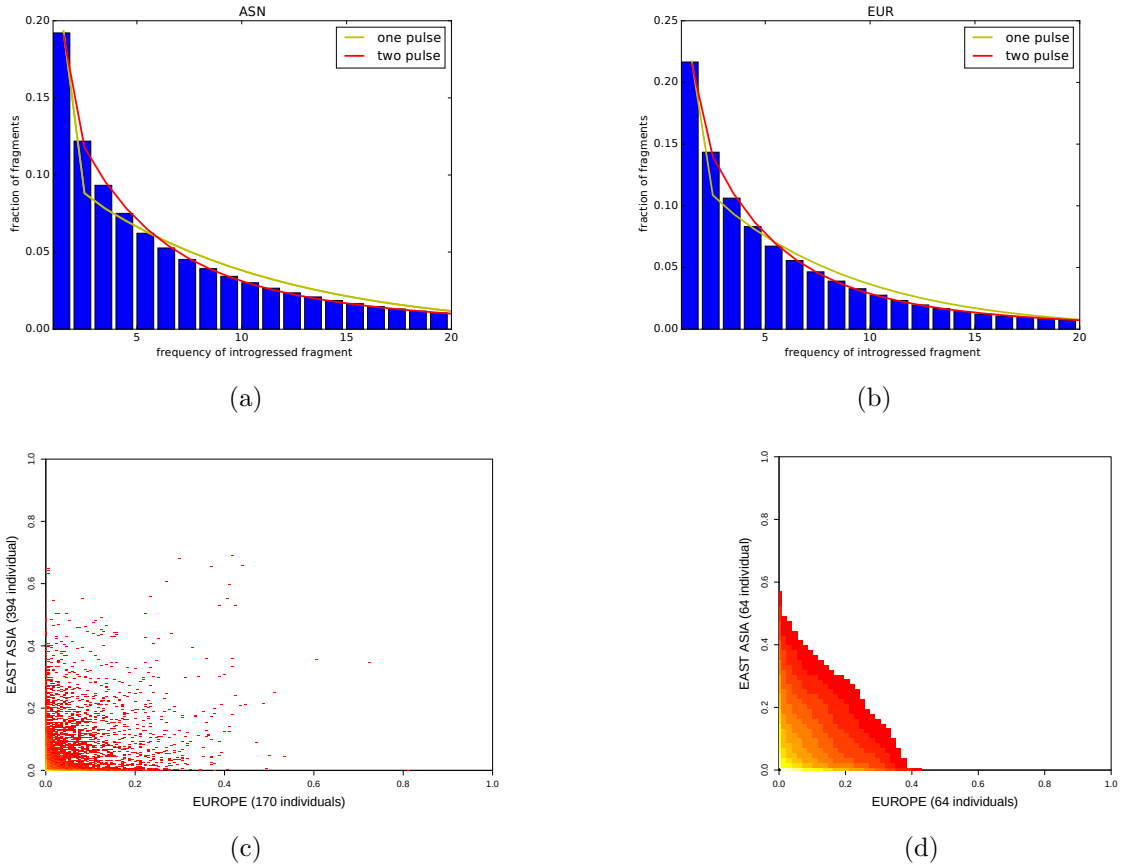


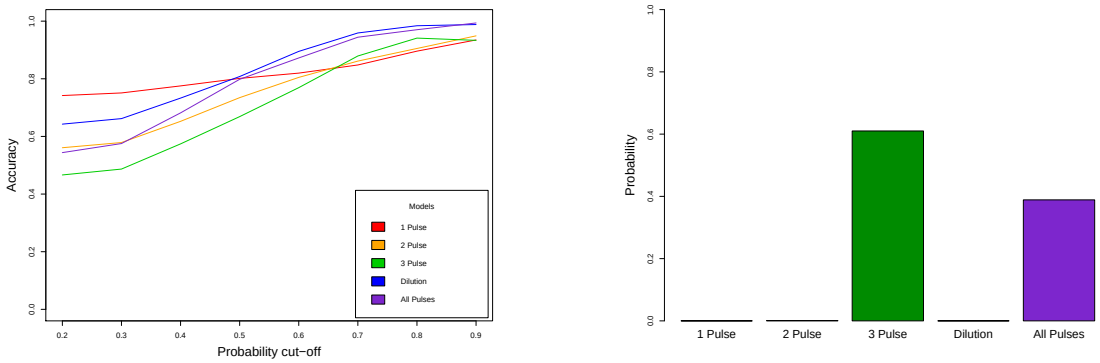
Figure 2: a,b) Marginal SFS for the East Asian and European populations (first 20 bins, excluding the 0 bin). The lines represent the best fitted one pulse and two pulse model for each population. c) FFS of the Steinrücken et al. [2018] introgression data. d)FFS of the Steinrücken et al. [2018] introgression data projected down to 64x64 bins, as used to train the FCNN.

Nonetheless, looking at each population individually, we did not have power to estimate the relative contribution of dilution and multiple admixtures in shaping the patterns of Neandertal fragments seen between Europe and Asia. To gain a more global picture of the

bioRxiv preprint doi: https://doi.org/10.1101/343087; this version posted June 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

history of human-Neandertal interbreeding, we developed a supervised machine learning approach. We simulated Neandertal admixture by specifying five demographic models with different numbers of admixture events (Supplementary Figure 1), and produce FFS under a range of parameters. We used the simulated FFS to train a fully-connected neural network (FCNN). The trained network classified models successfully ∼58% of the time, well above the 20% expected by chance, and was not overfit to the training data (Supplementary Figure 2). We then examined how the accuracy of the prediction changed when we required different levels of support for the chosen model (Figure 3a). Crucially, we see that when the classifier has high confidence in a prediction, it is very often correct, and that multiple pulse models are not often confused the dilution model (Supplementary Figure 3).

Finally, we applied the he trained FCNN to our empirical joint FFS (Figure 3b). Strikingly, we found that the model supported our two most complicated models, favoring a model with 3 pulses of admixture (Posterior probability ~ 0.61), and with a lower probability, a model with 3 pulses of admixture and dilution (Posterior probability ~ 0.39). These results are consistent across a range of cutoffs for calling introgressed fragments (Supplementary Figure 4), and dovetail with our maximum likelihood results showing that the best fit model must include multiple episodes of human-Neandertal interbreeding.



(a) (b)

Figure 3: (a) Accuracy of the predictions for all models under different levels of support for the chosen model. The x axis shows the probability cutoff that we used to classify models, and the y axis shows the accuracy. Each line corresponds to a different model. Simulated datasets in which no model surpassed the cutoff were deemed unclassified. (b) Posterior probability of the empirical introgression data matching each of the five demographic models, determined by the FCNN classifier.

bioRxiv preprint doi: https://doi.org/10.1101/343087; this version posted June 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

4 Discussion

Despite initial indications of a simple history of admixture between humans and Neandertals, more detailed analyses suggested that there might be additional, population specific episodes of admixture. By analyzing the joint fragment frequency spectrum of introgressed Neandertal haplotypes in modern Europeans and Asians, we found strong support for a model of multiple admixture events. Specifically, our results support a model in which the original pulse of introgression into the ancestral Eurasian population is supplemented with additional pulses to both European and East Asian populations after those populations diverge, resulting in elevated Neandertal ancestry in East Asians relative to Europeans. This is similar to a model recently proposed by Vernot et al. [2016] for explanation differential levels of Neandertal ancestry across Europe, Asia, and Melanesia. Importantly, our results exclude a demographic model where the difference in Neandertal ancestry between Europeans and East Asians is driven primarily through dilution of Neandertal ancestry in Europe due to recent admixture with Basal Eurasians, a population lacking Neandertal ancestry. Nonetheless, we cannot exclude dilution as playing a role in the differences in Neandertal ancestry between Europe and East Asia; a model which includes multiple pulses of Neandertal introgression and dilution through Basal Eurasians was the second likeliest model in the five model comparison. Given the strong evidence that Basal Eurasians contributed to the modern European gene pool, we suspect that dilution does play a role in shaping the pattern of Neandertal ancestry across Eurasia. However, a large amount of dilution would be necessary if it were the only factor explaining the $\sim$12% difference in Neandertal ancestry between Europe and East Asia, in contrast with recent work that inferred only a small ($\sim$9.4%) contribution of Basal Eurasians to modern European individuals [Kamm et al., 2018].

Several confounding factors could impact our inference. Although it is unlikely that differential purifying selection is responsible for the discrepancy between European and East Asian Neandertal ancestry [Kim and Lohmueller, 2015, Vernot and Akey, 2015, Juric et al., 2016], some Neandertal ancestry was likely deleterious [Harris and Nielsen, 2016, Juric et al., 2016] and our models assume neutrality. However, the strength of selection against introgressed fragments is likely to be small compared to the demographic forces at work; moreover, there is relatively little evidence of strong differences in the strength of selection between different non-African populations [Simons et al., 2014, Do et al., 2015]. Of additional concern is differential power to detect fragments in each population. To address this, we implemented an error model to analyze the one-dimensional data (Supplementary Material). We find that false positive and false negative rates are likely to be very low in both Europe and East Asia, and do not cause substantial shifts in the FFS (Supplementary Figure 5). Moreover, when we examined the weights assigned to different frequency categories, we found that moderate frequency haplotypes are most important to distinguish between models (Supplementary Figure 6), whereas errors in calling fragments are most likely to impact low frequency haplotypes. This, combined with the fact that our results

bioRxiv preprint doi: https://doi.org/10.1101/343087; this version posted June 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

are robust across two different datasets and a range of cutoffs for determining archaic ancestry, convince us that our results are robust to errors in fragment calling. Nonetheless, it is possible that some of the Neandertal ancestry in East Asia has been misclassified, and in fact originated from Denisovan introgression, mimicking the signal of additional pulses of Neandertal introgression. These problems are likely to be resolved as our ability to make accurate introgression calls for the various ancient human species improves in the future.

Our work provides additional evidence for the ubiquity of archaic admixture in recent human history, consistent with recent work shows that humans mixed with Denisovans multiple times [Browning et al., 2018]. Though we find that additional pulses of admixture in both East Asians and Europeans are necessary to explain the distribution of Neandertal ancestry in Eurasia, we are unable to settle *why* East Asians have elevated Neandertal ancestry. Interestingly, in contrast to Denisovans, there does not seem to be evidence of Neandertal population structure within introgressed fragments [Browning et al., 2018]. Combined with our results, this indicates that the Neandertal population or populations that admixed with Eurasians must have been relatively closely related. A likely explanation is that gene flow between humans and Neandertals was intermittent and ongoing, but in a somewhat geographically restricted region. Thus, differential levels of admixture between different Eurasian groups may primarily reflect how long those populations coexisted with Neandertals in that region.

5 Methods

5.1 Data

We obtained the joint fragment frequency spectrum by first downloading publicly available Neandertal introgression calls from two sources [Sankararaman et al., 2014, Steinrücken et al., 2018]. The Sankararaman data consists of the location of introgressed fragments along the genome in each phased haplotype for the 1000 Genomes Project populations. The Steinrücken data consists of the probability of Neandertal origin in 500 bp windows of the genome across each phased haplotype for the Central European (CEU) and Han Chinese (CHB) and Southern Han Chinese (CHS) from the 1000 Genomes Project. We computed fragment frequencies by sampling a single site every 100kb, from both sources of data. To compute the Joint Fragment Frequency Spectrum (FFS), we counted how many haplotypes were introgressed at each position. For the Steinrücken data, we called a site introgressed if it has a posterior probability of being introgressed above 45%. We then applied the 1000 Genomes accessibility mask (downloaded from `ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502/supporting/accessible_genome_masks/20141020.pilot_mask.whole_genome.bed`). In the supplement, we show that our results are robust to the cutoff and consistent between both datasets. Finally, we masked the (0,0) position of the FFS matrix, in order to reduce the impact of false negative fragment calls.

bioRxiv preprint doi: https://doi.org/10.1101/343087; this version posted June 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

5.2 Analytical Model

We model introgression of intensity f as injection of alleles at frequency f into the population at the time of introgression. In the one pulse model, this results in an exact expression for the expected fragment frequency spectrum under the Wright-Fisher diffusion model (Supplementary Material). Multiple pulse models could be solved analytically using a dynamic programming algorithm as in Kamm et al. [2018], but we instead approximate the expected frequency spectrum by making the approximation that the probability of sampling k introgressed haplotypes in a sample of size $n+1$ is the same as sampling k haplotypes in a sample of size n for large n (c.f. Jouganous et al. [2017]). This results in closed-form expressions for the expected frequency spectrum under both the two-pulse and the dilution model (Supplementary Material). With an expected frequency spectrum given parameters θ and model $\mathcal{M}$, $p_{n,k}(\theta, \mathcal{M}) = \mathbb{P}(\text{k out of n haplotypes are introgressed}|\theta, \mathcal{M})$, we compute the likelihood

$$L(\theta, \mathcal{L}) = \sum_{k=0}^{n} x_k \log(p_{n,k}(\theta, \mathcal{L}))$$

where x_k is the number of fragments found in k out of n individuals. We optimized the likelihood using `scipy`, and compared models using the likelihood ratio statistic

$$\Lambda = 2(L(\theta_2, \mathcal{M}_2) - L(\theta_1, \mathcal{M}_1))$$

where θ_i and $\mathcal{M}_i$ correspond to the i-pulse model. Under the null, Λ should be χ^2 distributed with 2 degrees of freedom (since there are 2 additional parameters in the two pulse model). Nonetheless, because we are examining a composite likelihood, we choose not to report explicit p-values, becuase they are likely to be miscalibrated.

5.3 Simulations

We used `msprime` [Kelleher et al., 2016], to simulate Neandertal introgression into two modern populations with multiple potential admixture episodes and dilution from Basal Eurasians (Supplementary Figure 1). For each replicate, we simulated the complete genomes for 170 European individuals, and 394 East Asian individuals, matching the sampling available from the 1000 Genomes Project panel. We used the human recombination map (downloaded from `http://www.well.ox.ac.uk/~anjali/AAmap/`; Hinch et al. [2011]). In each simulation we mimicked out sampling scheme on the real data by sampling 1 site every 100kb and calling a Neandertal fragment by asking which individuals coalesced with the Neandertal sample more recently than the human-Neandertal population split time.

For each simulation, we drew demographic parameters, including effective population sizes and divergence times, from uniform distributions. For effective population sizes we used 500-5000 individuals for Neandertals, 5000-50000 for Eurasians, 5000-100000 individuals for the European and East Asian populations. For divergence times, we used

bioRxiv preprint doi: https://doi.org/10.1101/343087; this version posted June 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

26000-12000 generations for Neandertals and humans, and 1300-2000 generations for the Eurasian split. The divergence between Basal Eurasians and Eurasians was fixed at 3000 generations. Lastly, we drew introgression times between 1500-3000 generations for gene flow into Eurasians, 800-2000 generations for gene flow into the European and East Asian populations. The time for the introgression event between Basal Eurasians and Europeans (dilution) was drawn from an uniform distribution of 200-2000 generations.

In order to ensure that our simulations focused on the correct parameter space, we constrained the resulting amount of Neandertal introgression in the modern European and East Asian genomes. The average Neandertal ancestry a was drawn from an uniform distribution between 0.01-0.03, and the difference in ancestry d between the East Asian and European populations was drawn from an uniform distribution between 0-0.01. We then determined the introgression intensity given a and d (Supplementary Material).

Our simulation pipeline is available at `https://github.com/Villanea/Neandertal_admix/blob/master/n_admix_10.py`, which generates the simulated genomes, identifies Neandertal introgression in 100kb windows, applies the 1000 Genomes accessibility mask, and outputs a FFS and accompanying parameters for each replicate.

5.4 Machine Learning (FCNN)

Using the resulting joint FFS, we trained a simple fully-connected neural network (FCNN) to categorize a joint FFS into one of five demographic models. The network was implemented in Keras [Chollet et al., 2015] using a TensorFlow back-end. The network used a simple architecture of three Dense layers (from 1024 nodes, to 512 nodes, to 64 nodes), each followed by a dropout layer (0.20). The code used to implement the network can be found at `https://github.com/Villanea/Neandertal_admix/blob/master/Fully_connected_network.py`.

6 Acknowledgments

We are grateful to Sara Mathieson and Jeff Spence for several useful discussions about neural network architecture and appropriate methods for training neural networks. We also wish to thank Jeff Spence and Matthias Steinrücken for extensive discussions on errors in fragment calling. Iain Mathieson, Sara Mathieson, and Jeff Spence provided invaluable feedback on an early draft of this manuscript that helped improve its clarity. This work was supported by NIH grant R35 GM124745.

References

Sharon R Browning, Brian L Browning, Ying Zhou, Serena Tucci, and Joshua M Akey. Analysis of human sequence data reveals two pulses of archaic denisovan admixture.

bioRxiv preprint doi: https://doi.org/10.1101/343087; this version posted June 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Cell, 173(1):53–61, 2018.

François Chollet et al. Keras. `https://github.com/fchollet/keras`, 2015.

Ron Do, Daniel Balick, Heng Li, Ivan Adzhubei, Shamil Sunyaev, and David Reich. No evidence that selection has been less effective at removing deleterious mutations in europeans than in africans. *Nature genetics*, 47(2):126, 2015.

Qiaomei Fu, Mateja Hajdinjak, Oana Teodora Moldovan, Silviu Constantin, Swapan Mallick, Pontus Skoglund, Nick Patterson, Nadin Rohland, Iosif Lazaridis, Birgit Nickel, et al. An early modern human from romania with a recent neanderthal ancestor. *Nature*, 524(7564):216, 2015.

Qiaomei Fu, Cosimo Posth, Mateja Hajdinjak, Martin Petr, Swapan Mallick, Daniel Fernandes, Anja Furtwängler, Wolfgang Haak, Matthias Meyer, Alissa Mittnik, et al. The genetic history of ice age europe. *Nature*, 534(7606):200, 2016.

Richard E Green, Johannes Krause, Adrian W Briggs, Tomislav Maricic, Udo Stenzel, Martin Kircher, Nick Patterson, Heng Li, Weiwei Zhai, Markus Hsi-Yang Fritz, et al. A draft sequence of the neandertal genome. *science*, 328(5979):710–722, 2010.

Kelley Harris and Rasmus Nielsen. The genetic cost of neanderthal introgression. *Genetics*, 203(2):881–891, 2016.

Anjali G Hinch, Arti Tandon, Nick Patterson, Yunli Song, Nadin Rohland, Cameron D Palmer, Gary K Chen, Kai Wang, Sarah G Buxbaum, Ermeg L Akylbekova, et al. The landscape of recombination in african americans. *Nature*, 476(7359):170, 2011.

Julien Jouganous, Will Long, Aaron P Ragsdale, and Simon Gravel. Inferring the joint demographic history of multiple populations: beyond the diffusion approximation. *Genetics*, pages genetics–117, 2017.

Ivan Juric, Simon Aeschbacher, and Graham Coop. The strength of selection against neanderthal introgression. *PLoS genetics*, 12(11):e1006340, 2016.

John A. Kamm, Jonathan Terhorst, Richard Durbin, and Yun S. Song. Efficiently inferring the demographic history of many populations with allele count data. *bioRxiv*, 2018. doi: 10.1101/287268. URL `https://www.biorxiv.org/content/early/2018/03/23/287268`.

Jerome Kelleher, Alison M Etheridge, and Gilean McVean. Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLoS Comput Biol*, 12(5):1–22, 05 2016.

Bernard Y Kim and Kirk E Lohmueller. Selection and reduced population size cannot explain higher amounts of neandertal ancestry in east asian than in european human populations. *The American Journal of Human Genetics*, 96(3):454–461, 2015.

bioRxiv preprint doi: https://doi.org/10.1101/343087; this version posted June 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Iosif Lazaridis, Nick Patterson, Alissa Mittnik, Gabriel Renaud, Swapan Mallick, Karola Kirsanow, Peter H Sudmant, Joshua G Schraiber, Sergi Castellano, Mark Lipson, et al. Ancient human genomes suggest three ancestral populations for present-day europeans. *Nature*, 513(7518):409, 2014.

Iosif Lazaridis, Dani Nadel, Gary Rollefson, Deborah C Merrett, Nadin Rohland, Swapan Mallick, Daniel Fernandes, Mario Novak, Beatriz Gamarra, Kendra Sirak, et al. Genomic insights into the origin of farming in the ancient near east. *Nature*, 536(7617):419, 2016.

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.

Kay Prüfer, Cesare de Filippo, Steffi Grote, Fabrizio Mafessoni, Petra Korlević, Mateja Hajdinjak, Benjamin Vernot, Laurits Skov, Pinghsun Hsieh, Stéphane Peyrégne, et al. A high-coverage neandertal genome from vindija cave in croatia. *Science*, page eaao1887, 2017.

Sriram Sankararaman, Nick Patterson, Heng Li, Svante Pääbo, and David Reich. The date of interbreeding between neandertals and modern humans. *PLoS genetics*, 8(10): e1002947, 2012.

Sriram Sankararaman, Swapan Mallick, Michael Dannemann, Kay Prüfer, Janet Kelso, Svante Pääbo, Nick Patterson, and David Reich. The genomic landscape of neanderthal ancestry in present-day humans. *Nature*, 507(7492):354, 2014.

Yuval B Simons, Michael C Turchin, Jonathan K Pritchard, and Guy Sella. The deleterious mutation load is insensitive to recent population history. *Nature genetics*, 46(3):220, 2014.

Matthias Steinrücken, Jeffrey P. Spence, John A. Kamm, Emilia Wieczorek, and Yun S. Song. Modelbased detection and analysis of introgressed neanderthal ancestry in modern humans. *Molecular Ecology*, 2018. doi: 10.1111/mec.14565.

Benjamin Vernot and Joshua M Akey. Complex history of admixture between modern humans and neandertals. *The American Journal of Human Genetics*, 96(3):448–453, 2015.

Benjamin Vernot, Serena Tucci, Janet Kelso, Joshua G Schraiber, Aaron B Wolf, Rachel M Gittelman, Michael Dannemann, Steffi Grote, Rajiv C McCoy, Heather Norton, et al. Excavating neandertal and denisovan dna from the genomes of melanesian individuals. *Science*, page aad9416, 2016.

bioRxiv preprint doi: https://doi.org/10.1101/343087; this version posted June 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Jeffrey D Wall, Melinda A Yang, Flora Jay, Sung K Kim, Eric Y Durand, Laurie S Stevison, Christopher Gignoux, August Woerner, Michael F Hammer, and Montgomery Slatkin. Higher levels of neanderthal ancestry in east asians than in europeans. *Genetics*, 194(1): 199–209, 2013.