Selection against archaic DNA in human regulatory regions

Natalie Telis[1], Robin Aguilar[2], and Kelley Harris[1,2]
[1]Department of Genetics, Stanford University
[2]Department of Genome Sciences, University of Washington

Correspondence: natalie.telis@gmail.com, harriske@uw.edu

July 18, 2019

Abstract

Although traces of archaic hominin DNA persist in all human populations outside Africa, these traces have been systematically depleted from the most functionally important regions of the human genome. This depletion suggests that many Neanderthal and Denisovan alleles had harmful effects on the fitness of hybrid individuals, but the nature of these harmful effects is poorly understood. Here, we show that Neanderthal and Denisovan alleles likely dysregulated gene expression in specific human tissues, causing systematic depletion of archaic introgression from enhancer regions annotated by the ENCODE RoadMap project. Highly pleiotropic enhancers that show activity in many tissues are more depleted of introgression than tissue-specific enhancers, and Neanderthal depletion is highly correlated with Denisovan depletion across sets of enhancers active in particular tissues. Fetal brain and fetal muscle are the tissues most depleted of archaic SNPs in their regulatory regions, and by analyzing the site frequency spectra of enhancers compared to control regions, we deduce that brain and muscle enhancers are likely depleted of introgression for different reasons. Brain enhancers appear to accumulate deleterious mutations unusually quickly, which may have caused inbred archaic populations to accumulate disproportionately large amounts of genetic load in these regions. In contrast, fetal muscle enhancers show no evidence of high deleterious mutation rates, and we hypothesize that their depletion of archaic DNA resulted from divergent selection for higher muscle mass in Neanderthals and Denisovans compared to humans.

Introduction

Neanderthals came into contact with humans multiple times during their existence, leaving traces of archaic DNA in the genomes of present day non-Africans [1, 2]. These traces imply that human-Neanderthal hybrids were once viable, fertile and numerous. At the same time, Neanderthal alleles were not always well tolerated on human genetic backgrounds. Large amounts of Neanderthal DNA appear to have been purged from regions of the human genome that show evidence of evolutionary constraint [3, 4], suggesting that deleterious archaic alleles were removed by natural selection after they harmed the fitness of hybrid individuals. A similar pattern has been observed in the genomes of Oceanians whose ancestors admixed with the Denisovan hominid lineage: many genetic "deserts" that are devoid of Neanderthal ancestry in Europeans and East Asians are similarly devoid of Denisovan admixture in Oceanians [5, 6].

The depletion of archaic introgression from coding regions of the genome has been recognized for years [3, 4], but Petr, et al. recently measured an even starker depletion of Neanderthal DNA from conserved noncoding regions of the genome [7]. This is exciting in light of the compounding evidence that gene regulatory DNA evolves more nimbly than the exome does, making gene regulation likely responsible for much of the phenotypic divergence between closely related lineages of hominins [8, 9, 10]. Even the Neanderthal regulatory DNA that remains in the human gene pool is not necessarily benign, but appears to perturb gene expression in deleterious ways [11, 12]. This suggests that extinct Neanderthal alleles likely dysregulated human gene expression even more than the Neanderthal DNA that remains within extant populations.

It can be challenging to measure the fitness effects of gene regulatory mutations because the grammar relating sequence to function is not completely understood [13]. Previous studies showed that the Neanderthal exome is under weaker constraint than the human exome by comparing rates of amino acid change to rates of substitution at synonymous sites [14], but it is less straightforward to decipher the fitness effects of Neanderthal variants that occur in regulatory elements such as enhancers. Allele frequency spectra and patterns of sequence divergence can sometimes provide information about the mode and intensity of selection on noncoding regions [15, 16, 17, 18, 19], but introgressed variants have a highly unusual distribution of allele ages and frequencies that can confound the efficacy of standard methods that assume simple population histories [20]. Reporter assays can shed light on this question

by measuring direct impacts of genetic variants on gene expression *in vitro* [21, 22], and these assays have been instrumental for measuring the functional effects of specific introgressed variants [23]. However, functional assays cannot translate changes in transcription factor binding or gene expression into the subtle effects on survival and reproduction that likely determined which archaic variants were purged from human populations.

Although we cannot often distinguish constrained noncoding sites from neutrally evolving noncoding sites at single base pair resolution, the ENCODE project recently identified enhancer and promoter regions that regulate gene expression in cell lines cultured from a variety of human tissues [24, 25]. These annotations make it possible to test whether archaic variation is systematically depleted or enriched in sets of regulatory regions that act together in functional concert. Silvert, et al. recently used this dataset to deduce that immune cell gene regulation evolved in a unique way following introgression, favoring the rise of some Neanderthal variants to higher frequencies than expected by chance or observed in the regulomes of non-immune cells [26]. This pattern of positive selection underscores the importance of partitioning regulatory SNPs by cell type to understand the range of their fitness effects.

In contrast to Silvert, et al., who zeroed in on high frequency Neanderthal alleles that appear to have been positively selected for their beneficial effects on human gene regulation, we aim to identify the tissues whose archaic gene regulatory sequences were likely purged away due to deleterious effects. Our second aim is to quantify the symmetry, or lack thereof, between selection against archaic introgression and selection against ordinary point mutations. Previous studies have found that sequence conservation between species is fairly predictive of introgression depletion [3, 4, 5, 6], implying that conserved, mutation-intolerant regions also tend to be intolerant to the shuffling of variation between groups as divergent as humans and Neanderthals. However, archaic species were not simply reservoirs of variation lying outside the human norm, but were shaped by their own suites of selective pressures that sometimes converged with the human fitness landscape and sometimes likely diverged from it. We aim to identify classes of regulatory DNA where selection against archaic introgression appears to have been qualitatively different from selection against new mutations, thereby shedding light on how the pace of regulatory evolution has varied between human tissues and cell types.

Results

Enhancers are depleted of Neanderthal introgression compared to control regions affected by similar levels of background selection

If Neanderthal enhancers tended to dysregulate gene expression in hybrid individuals, we should expect enhancers to have lower Neanderthal admixture levels than DNA with no annotated regulatory function. To test for such a pattern, we intersected the ENCODE Roadmap enhancer calls with the high confidence Neanderthal and Denisovan SNP calls generated by Sankararaman, et al. from the Simons Genome Diversity Panel [5, 27]. A complicating factor, however, is that enhancers may have been depleted of archaic DNA due to background selection rather than their own functional attributes. Neanderthal gene regulatory variants may have entered the human gene pool in linkage disequilibrium with Neanderthal coding variants in highly conserved genes and been subsequently purged due to selection on protein sequence.

Following Sankararaman, et al., we used McVicker and Green's B-statistic to estimate the amount of background selection that enhancer loci experience [28, 3, 5]. To test whether archaic variants occur within enhancers less often than predicted by their B statistic distribution, we randomly paired each Neanderthal variant with a "control" variant matched for both B-statistic decile and allele frequency (Figure 1A). Chromosome X was excluded given its systematic depletion of Neanderthal and Denisovan variants.

In every population, we found that these matched control variants appear within enhancers significantly more often than introgressed variants do (Figure 1B), with depletion odds ratios ranging from 84% to 91%. As expected, this method also detects negative selection against introgression in exons and indicates that exons and enhancers were less tolerant to the influx of Neanderthal variation than similar regions that do not bind TFs or regulate gene expression. Enard, et al. recently used a related approach to quantify Neanderthal introgression in proteins that interact with viruses [29].

Highly pleiotropic enhancers contain less introgression than tissue-specific enhancers

The enhancers annotated by the RoadMap project exhibit wide variation in tissue specificity. Some are active in only one or two tissues, while others show evidence of activity in 20 tissues or more [25]. When we stratified enhancers by pleiotropy number, meaning

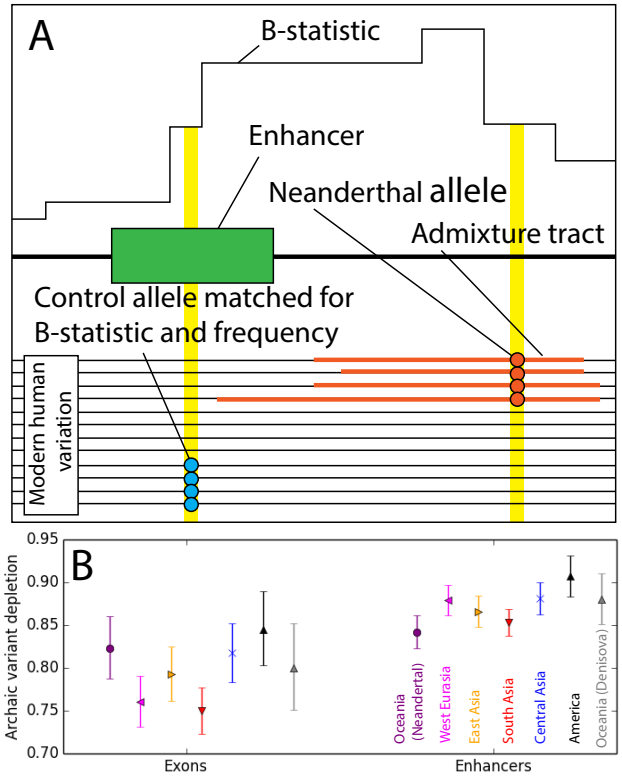

Figure 1: A. This schematic illustrates the process of matching archaic variants to "control" variants with identical allele frequencies and B statistic values. B. Introgressed-to-control variant odds ratios show that Neanderthal DNA has been depleted from both exons and enhancers in every population sequenced by the Simons Genomic Diversity Project. In the case of Oceanians, a similar pattern holds for DNA introgressed from Denisovans. Error bars span 95% confidence intervals.

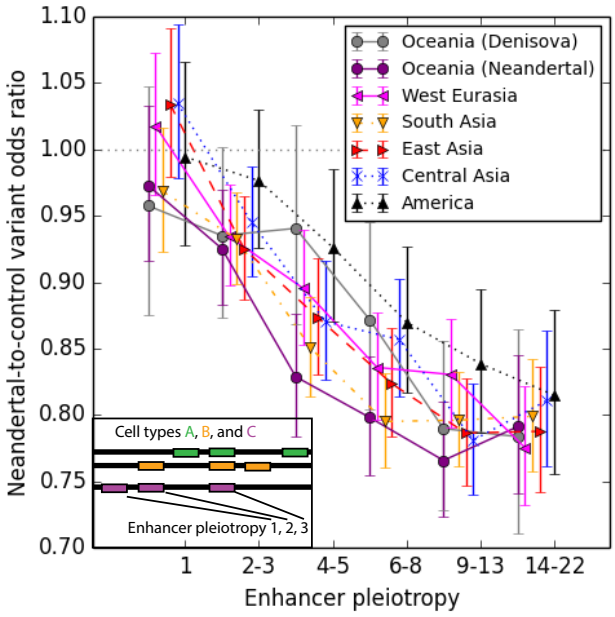

Figure 2: Enhancers active in only a single cell type are not depleted of introgressed variation, whereas enhancers that are active in multiple cell types contain up to 20% fewer introgressed variants than expected.

the number of tissues where the enhancer is active (Figure 2), we found pleiotropy to be positively correlated with the magnitude of archaic ancestry depletion (Figure 2).

We found that some tissues' enhancers were much more depleted of archaic DNA than others' were (Figure 3A; Figure 3–Figure Supplement 1), with significant correlation between depletion of Neanderthal DNA and depletion of Denisovan DNA ($r^2 = 0.537$, $p < 4e{-}5$). Enhancers active in fetal muscle, fetal brain, and neurosphere cells are the most highly depleted of introgression; conversely, enhancers active in fetal blood cells and T-cells , as well as mesenchymal cells, were consistently less depleted of introgression than enhancers active in other cell types. Mesenchymal cells, T cells, and other blood cells are among the cell types where some adaptively introgressed regulatory variants are thought to be active [30, 31, 32, 26], but our results suggest that selection decreased introgression levels overall even within the regulatory networks of these cells. The excess introgression depletion in brain and fetal muscle is a pattern that holds robustly across populations (Figure 3–Figure Supplement 2).

Although brain and fetal muscle enhancers are significantly more depleted of introgression than other enhancers, this pattern cannot explain much of the correlation between introgression levels and pleiotropy (Figure 3B). When Neanderthal and Denisovan variants from

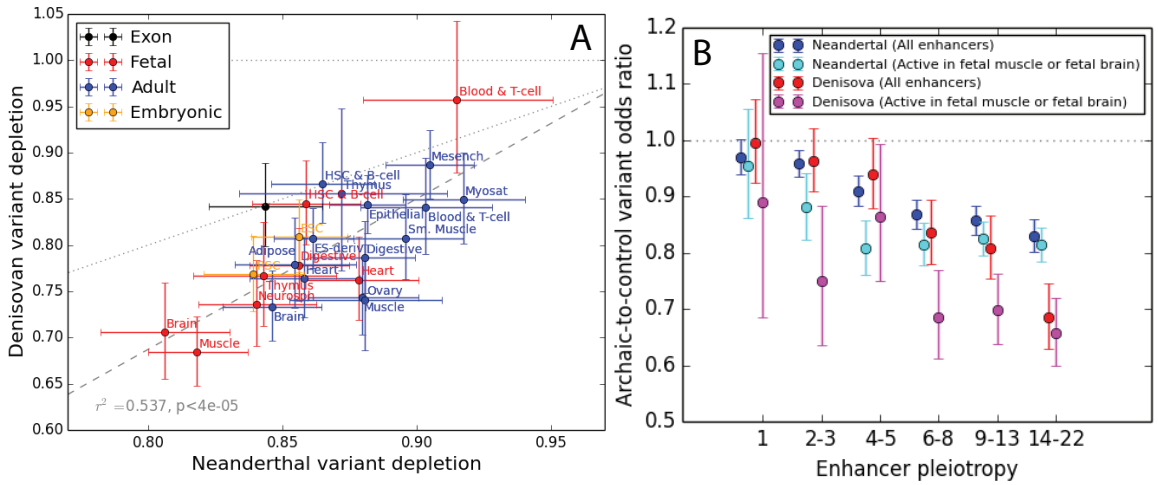

Figure 3: A. Neanderthal and Denisovan introgression levels vary between enhancers active in different tissues. Data points that lie below the dotted line correspond to tissues whose enhancers are more depleted of Denisovan introgression compared to Neanderthal introgression. B. Even after restricting to enhancers active in fetal muscle or fetal brain, the two tissue types most depleted of archaic introgression, pleiotropy remains negatively correlated with introgression depletion. The difference in introgression depletion between these two tissues and other tissues is driven mainly by enhancers of intermediate pleiotropy.

different populations are pooled together, Denisovan variants appear to be more strongly depleted from regulatory DNA than Neanderthal variants are. Although exons are slightly more depleted of archaic introgression than enhancers as a whole (Figure 1B), they are actually less depleted than brain or fetal muscle enhancers.

Old variation shared by Neanderthals and Denisovans was likely less deleterious to humans than variation that arose in these species more recently

The ascertainment of introgression tracts is still a challenging technical feat that is being actively refined by computational methods developers [33, 34, 35]. Motivated by this, we wanted to assess whether differences in the parameters used to infer archaic ancestry tracts would affect our conclusions about how selection has shaped the introgression landscape. To this end, we compared two different introgression call sets that Sankararaman, generated from the SGDP data [5] using conditional random field inference.

The conditional random field approach developed by Sankararaman, et al. calls a Eurasian DNA segment as introgressed if it is sufficiently more similar to a reference archaic sequence than to an outgroup panel. One set of archaic calls, denoted "Set 1," was generated using an outgroup panel composed entirely of Africans. A potential problem with this call set is

that Neanderthals and Denisovans were more closely related to each other than to humans, meaning that the CRF might misidentify Neanderthal tracts as Denisovan and vice versa. To avoid this issue, Sankararaman, et al. generated a second Neanderthal call set, "Set 2," using an outgroup panel that contained the Altai Denisovan as well as African humans. Similarly, Denisovan call set 2 was generated using a panel that included the Altai Neanderthal. The analyses presented in Figures 1 through 3 utilize the Set 2 Neanderthal calls as well as Set 2 Denisovan calls that were generated using the Altai Neanderthal as well as Africans as outgroup sequences.

Compared to Set 2, we hypothesized that the more inclusive Set 1 calls should contain more old variation that arose in the common ancestral population of Neanderthals and Denisovans, and that this older variation might be better tolerated when it introgressed into humans because it rose to high frequency in an ancestral population that had had less time to diverge from humans and accumulate deleterious mutations compared to the mature Neanderthal and Denisovan populations that had experienced long periods of isolation and inbreeding. To test this hypothesis, we compiled sets of "distal" Neanderthal and Denisovan variation by including their respective Set 1 introgression calls and excluding all Set 2 introgression calls. In contrast to the "proximal" Set 2 introgression calls, these distal calls are not measurably depleted from enhancers compared to control variants matched for allele frequency and B statistic (Figure 4). This suggests that the introgression landscape was shaped mainly by selection against Neanderthal and Denisovan genetic variants that arose relatively close to the time that gene flow occurred, not variation that arose soon after their isolation from humans. Many populations actually show a slight enrichment of distal archaic variants in enhancers compared to controls, possibly because these old, shared variants are enriched for beneficial alleles that swept to high frequency in the common ancestor of Neanderthals and Denisovans.

Introgression depletion and the site frequency spectrum reveal different landscapes of evolutionary constraint

Since introgression levels are broadly correlated with genomic conservation, we sought to measure whether the enhancer categories that are most depleted of archaic introgression are simply the categories with the highest intolerance to new mutations. A useful indicator of mutation intolerance is the shape of the site frequency spectrum: if purifying selection keeps a substantial fraction of new mutations from rising to high frequency, we expect a higher

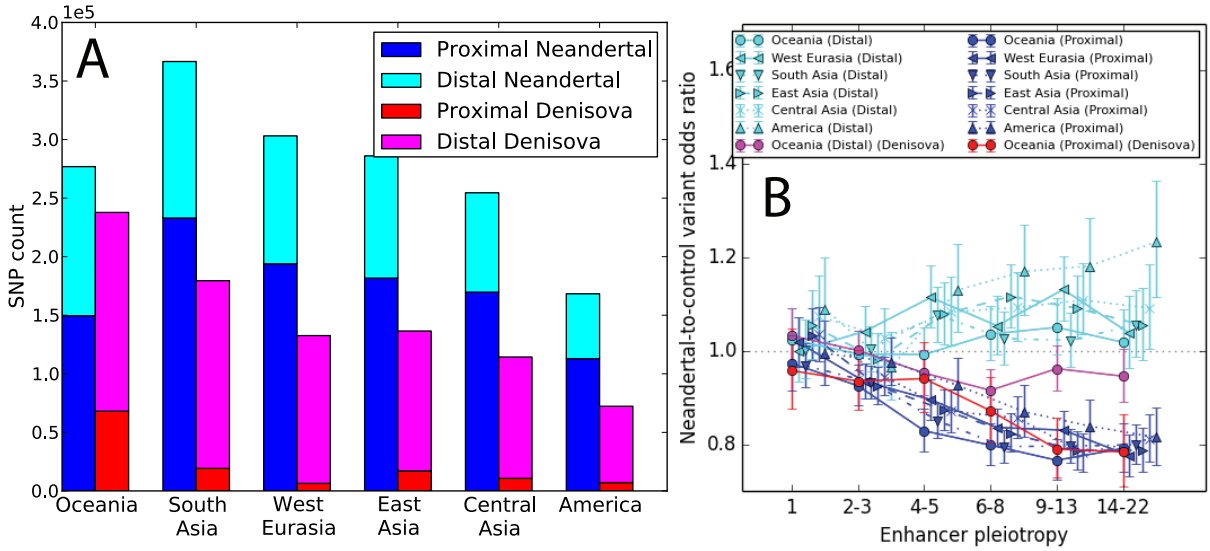

Figure 4: A. Sankaraman, et al. generated two sets of Neanderthal ancestry calls in the SGDP data: a more inclusive Set 1 and a more stringent Set 2, which included only segments that are sufficiently distinct from the Denisovan reference genome. We refer to Set 2 as "proximal" variation since it is closer to the Altai Neanderthal reference. About 30% of Set 1 is excluded from Set 2, and we refer to this as "distal" Neanderthal variation. Proximal and distal Denisovan variation is defined in the same way, but a much smaller fraction of Denisovan introgression falls into the proximal category, probably because the Denisovan reference individual was relatively distantly related to the Denisovans who contributed to the human gene pool [34]. B. In contrast to the proximal archaic variation considered elsewhere in this paper, distal archaic variation is not measurably depleted from enhancers, even enhancers active in numerous tissues.

proportion of segregating variants to be rare compared to a scenario where new mutations segregate neutrally or have beneficial effects [36] (Figure 5A).

Motivated by this, we analyzed the site frequency spectrum of African enhancer variation from the 1000 Genomes Project. Neanderthals contributed little or no genetic material to sub-Saharan African populations [1], meaning that Neanderthal alleles should have no direct effect on the African site frequency spectrum. One caveat is that this strategy will not identify regions that are prone to strongly deleterious mutations–mutations that are lethal or nearly lethal in heterozygotes will not segregate long enough to affect the frequency spectrum's shape. However, strongly deleterious mutations are not expected to contribute to mutation load differences between populations, making it appropriate to focus on a metric that will identify regions experiencing high rates of weakly deleterious mutations.

To estimate the strength of purifying selection against weakly deleterious alleles that has been acting on a particular set of enhancers, we computed the proportion of segregating variants that are singletons and compared this to the proportion of variants that are singletons in the enhancer-sized regions immediately upstream of the annotated regulatory elements (Figure 5B). By comparing enhancers to immediately adjacent regions, we control for the potentially confounding effects of recombination rate, background selection, and sequencing read depth. There is, admittedly, a base composition difference between enhancers and adjacent control regions, with enhancers having higher GC content. In the absence of functional differences between enhancers and adjacent regions, this should bias the enhancer site frequency spectrum toward a lower singleton frequency, seeing that biased gene conversion tends to increase the frequencies of mutations from AT to GC and depress the frequencies of mutations from GC to AT. However, we observe the opposite pattern in empirical site frequency data: enhancers contain a significantly higher proportion of rare variants than adjacent regions do. The same pattern holds if we count only SNPs with AT ancestral alleles, and likewise if we count only SNPs with GC ancestral alleles (Figure 5–Figure Supplement 1). Across tissue categories, there is no apparent correlation between GC content and the enrichment of rare variants in enhancers relative to control regions (Figure 5–Figure Supplement 2). This suggests that purifying selection is driving the difference between the site frequency spectra of enhancers and control regions, not base composition or biased gene conversion.

Although enhancers broadly show evidence of purifying selection against both archaic

variation and new mutations, the strength of selection against these two types of perturbation is poorly correlated among tissues (Figure 5C, D). Fetal brain, neurosphere cells, and to a lesser extent, adult brain, are the only tissues whose enhancers show elevated selective constraint, suggesting that mutations perturbing brain development have an outsize probability of deleterious consequences. Brain tissues are also some of the tissues most strongly depleted of archaic DNA, and with all tissues included, singleton enrichment appears correlated with Neanderthal depletion ($r^2 = 0.31$, $p < 0.004$) and Denisovan depletion ($r^2 = 0.27$, $p < 0.009$). However, when these brain tissue outliers are excluded, there remains no correlation between the strength of selection against new mutations and the strength of selection against introgression (Neanderthal $p < 0.42$; Denisovan $p < 0.10$; see Figure 5–Figure Supplement 3). In particular, fetal muscle enhancers show no evidence of unusual selective constraint despite their strong depletion of both Neanderthal and Denisovan ancestry.

Enhancer pleiotropy is positively correlated with singleton enrichment as well as the depletion of archaic ancestry (Figure 5E). One difference, however, is that enhancers active in only a single tissue (pleiotropy number 1) still show significant evidence of selection against new mutations despite their lack of any evidence for selection against archaic introgression. Overall, enhancer pleiotropy appears to have affected the introgression landscape more strongly than it affects the tendency to retain new mutations. This observation may be related to experimental evidence that the most highly pleiotropic enhancers tend to have the most consistently conserved functioning across species [37].

Discussion

The weakness of the correlation between introgression depletion and singleton enrichment provides us within insight into the fitness effect differences between *de novo* human mutations and alleles that entered the human population through introgression. This difference appears starkest when comparing enhancers to exons. Exons are well known to evolve more slowly than enhancers over phylogenetic timescales [38, 39, 40], implying that selection acts more strongly against new coding mutations compared to new regulatory mutations. Enhancers show much stronger evidence of selection against introgression than we might expect given their relatively modest intolerance between new mutations, suggesting that regulatory effects played a significant role in shaping the landscape of Neanderthal and Denisovan introgression.

Two sources of dysfunction are thought to drive selection against archaic introgression:

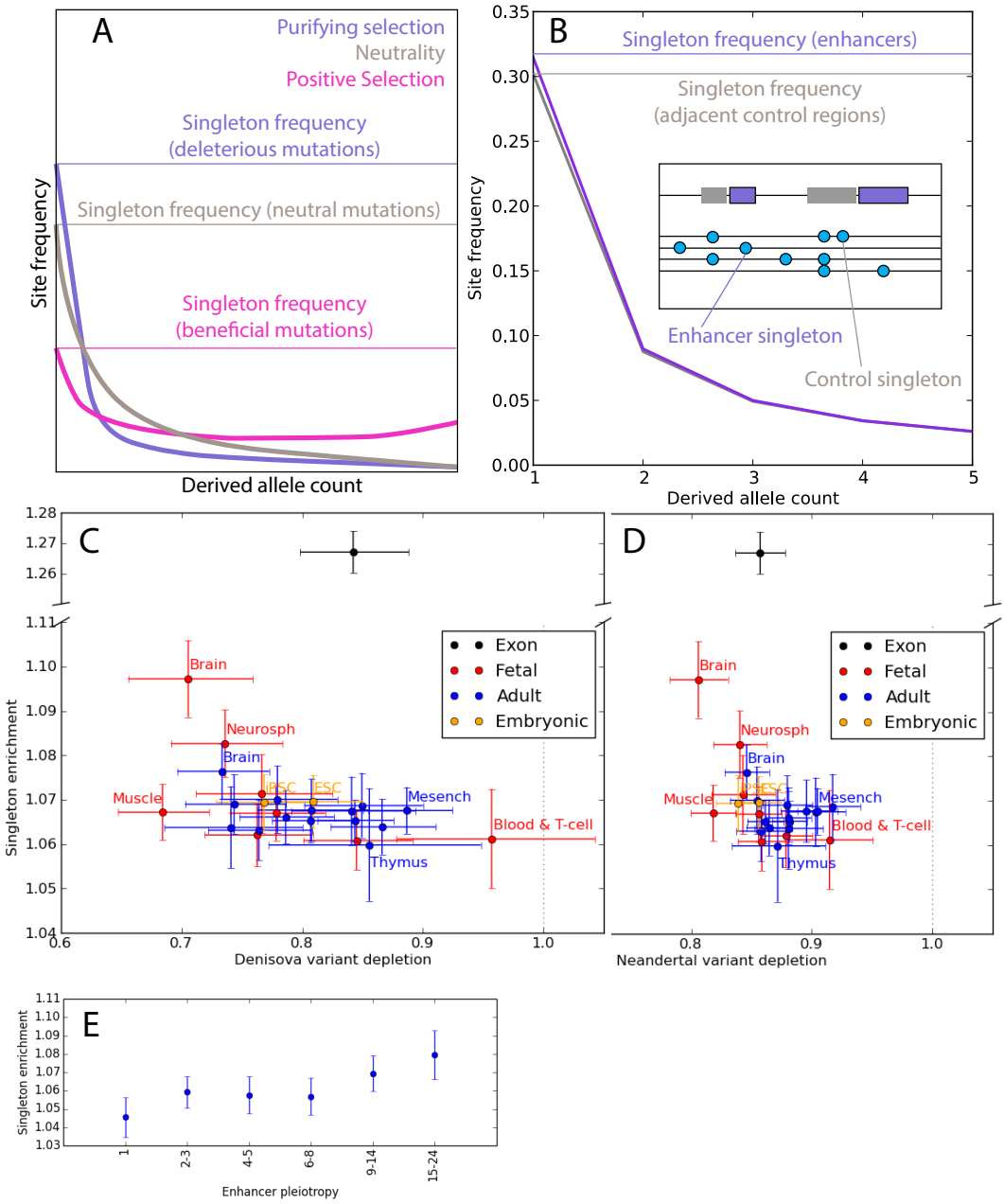

Figure 5: A. Theory predicts that the site frequency spectrum (SFS) becomes skewed toward rare variants by the action of purifying selection. B. In African data from the 1000 Genomes project, the enhancer SFS has a higher proportion of singletons compared to control regions adjacent to enhancers. C. Every tissue type's enhancer complement is enriched for singletons compared to adjacent control regions. This comparison of singleton enrichment odds ratios to Denisovan depletion odds ratios shows that fetal brain, neurosphere cells, and adult brain are outliers under stronger constraint. The y axis has been split to accommodate the magnitude of singleton enrichment in exons. Error bars span 2 binomial test standard errors. D. Comparison of the singleton enrichment landscape to the Neanderthal depletion landscape. E. Enhancer pleiotrophy is negatively correlated with singleton enrichment, though even enhancers of pleiotropy 1 have a singleton enrichment odds ratio significantly greater than 1.

one is the accumulation of deleterious mutation load in inbred Neanderthal and Denisovan populations [41, 42] and the other is the accumulation of epistatic incompatibilities due to divergent selective landscapes [3, 5, 43]. Both forces have the potential to affect enhancers, and our comparison between the introgression landscape and the site frequency spectrum confers some ability to distinguish the two. When a set of regulatory elements is more depleted of introgression than we would expect given the strength of singleton enrichment in these regions, this suggests that archaic variation in these regions is more deleterious than expected, possibly due to selection in Neanderthals and Denisovans for alleles that are damaging on the human genetic background. Fetal muscle enhancers appear to fit this profile, with unremarkable levels of singleton enrichment but strong depletion of Neanderthal and Denisovan introgression. Archaeological evidence indicates that Neanderthals had higher muscle mass, strength, and anatomical robustness compared to humans [44, 45], supporting the idea that the two species had different fetal muscle growth optima. We have no direct knowledge of Denisovan muscle anatomy, but the depletion of Denisovan DNA from muscle enhancers suggests that they may have shared Neanderthals' robust phenotype.

In contrast to the case of muscle enhancers, mutation load is an attractive candidate cause for the depletion of archaic DNA from brain enhancers. Brain enhancers' high singleton enrichment implies that mutations in these regions are more often deleterious than mutations in enhancers not active in brain cells, an idea that is bolstered by the discovery of severe developmental disorders caused by noncoding *de novo* mutations affecting brain gene expression [46, 47, 48]. High mutational fitness effects could cause brain enhancers to accumulate genetic load more quickly than other enhancers during periods of inbreeding.

Both genetic load and hybrid incompatibilities might contribute to the correlation between enhancer pleiotropy and introgression depletion. Steinrücken, et al noted that epistatic incompatibilities are most likely to arise in genes with many interaction partners [49], and though they did not find evidence that such a correlation existed for genes, the pleiotropy correlation that we describe here reveals that introgression is most depleted from enhancers with many interaction partners. However, an enhancer's number of interaction partners may also correlate with its deleterious mutation rate. Genes that are expressed in many tissues evolve more slowly than genes expressed in few tissues because they have greater potential for functional tradeoffs [50, 10], and a mutation that disrupts the balance of a functional tradeoff is likely to have a deleterious effect. This idea is supported by our finding that sin-

gleton enrichment correlates with enhancer pleiotropy, meaning that genetic load is likely at least a partial contributing factor to selection against introgression in enhancers with many interaction partners.

Both genetic load and epistatic incompatibilities are expected to "snowball" over time, making young archaic variation more likely to be deleterious in hybrids compared to older, high frequency archaic variation. We were able to observe this by leveraging the existence of multiple pulses of archaic gene flow from Neanderthal and Denisovan lineages that were isolated from each other at the time of admixture but still sister species with respect to humans. Part of this effect might be due to positive selection on beneficial introgressed alleles that have risen to high frequency in multiple populations. As more methods for inferring admixture tracts are developed, our results underscore the importance of investigating how they might be biased toward young or old archaic variation and using this information to update our understanding of how selection shapes introgression landscapes.

Our work stops short of numerically estimating the selection coefficients of introgressed variants, but suggests that the DFE of variants that enter a population by introgression tends to be fairly similar between coding and noncoding regions, whereas the fitness effects of new coding mutations tend to be greater than the fitness effects of new noncoding mutations. Regulatory mutations appear to have created incompatibilities between many species that are already in the advanced stages of reproductive isolation [51, 52, 53, 54], and our results suggest that they also acted as incipient barriers to gene flow between humans and Neanderthals. As more introgression maps and functional genomic data are generated for hybridizing populations of non-model organisms, it should be possible to measure the prevalence of weak regulatory incompatibility in more systems that exist in the early stages of reproductive isolation and test how many of the patterns observed in this study occur repeatedly outside the hominoid speciation continuum.

Code availability statement

Summary data files and custom python scripts for reproducing the paper's main figures are available at `https://github.com/kelleyharris/hominin-enhancers/`.

Author contribution statement

N.T. and K.H. conceived and designed the project. N.T., R.A., and K.H. performed the analyses. K.H. wrote the paper.

Acknowledgements

We are grateful to Jonathan Pritchard, Sriram Sankararaman, Josh Schraiber, and members of the Harris Lab for helpful discussions, and we thank Rasmus Nielsen for manuscript comments. We acknowledge financial support from the Burroughs Wellcome Fund, the Searle Foundation, the Alfred P. Sloan Foundation, and the Pew Charitable Trusts.

Methods

Data Access

All datasets analyzed here are publicly available at the following websites:

SGDP	https://www.simonsfoundation.org/simons-genome-diversity-project/
RoadMap	https://personal.broadinstitute.org/meuleman/reg2map/HoneyBadger2_release/
1000 Genomes Phase 3	http://www.1000genomes.org/category/phase-3/

Extraction of Neanderthal and Denisovan variant sets

Neanderthal and Denisovan variant call sets were downloaded from https://sriramlab.cass.idre.ucla.edu/public/sankararaman.curbio.2016/summaries.tgz. These files classify a haplotype as archaic if it is classified as archaic with $\geq 50\%$ probability by the conditional random field analyses reported in [5]. Using these summaries, we classify a variant as archaic if 100% of the haplotypes on which it appears are classified as such. Unless otherwise stated, all Neanderthal and Denisovan variants are obtained from the respective summary call set "2," which we refer to in the text as the proximal call sets. To construct the distal Neanderthal call set analyzed in Figure 4, we included all variants from Neanderthal Set 1 except any variants that also appeared in Neanderthal Set 1 or Denisovan Set 1. Similarly, the distal Denisovan call set included all variants present in Denisovan Set 1 except those variants also present in Neanderthal Set 2 or Denisovan Set 2.

Classifying enhancers by tissue type and pleiotropy number

Cell lines were classified into tissue types using the the tissue assignment labels from the file `Consolidated_EpigenomeIDs_summary_Table.csv` in the July 2013 RoadMap data compendium. Whenever a tissue type contained both fetal and adult cell lines, we further subdivided that tissue type into "Adult" and "Fetal." We then computed a pleiotropy number for each enhancer by counting the number of distinct tissue type labels in the cell lines where that enhancer is annotated as active (state 6, 7 or 12 in the honeybadger model). Fetal and adult tissue types are counted as distinct tissues for the purpose of this computation.

Testing for depletion of archaic variation relative to matched control variation

To estimate the strength of background selection experienced by human genomic loci, B-statistic values ranging on a scale from 1 to 1000 were downloaded at `http://www.phrap.org/software_dir/mcvicker_dir/bkgd.tar.gz`. We quantized these values by rounding them down to the nearest multiple of 50 B-statistic units, then lifted them over from hg18 coordinates to hg19 coordinates. Each SNP in the SGDP data was assigned the B statistic value of the closest site annotated by McVicker, et al.

Our tests for depletion of archaic variation are computed relative to non-archaic control SNPs that have the same joint distribution of allele frequency and B statistic as the SNPs annotated as archaic in origin. See the next section, "Detailed sampling procedure for matched control SNPs", for more information on how these matched control sets are obtained.

Assume that $\mathcal{A}$ is a set of A archaic SNPs and $\mathcal{C}$ is a set of $2 \times A$ matched controls. To test whether archaic variation of this type is enriched or depleted in a set $\mathcal{G}$ of genomic regions, we start by counting the number A_G of archaic SNPs contained in $\mathcal{G}$ and the number C_G of control SNPs contained in $\mathcal{G}$. We say that this type of archaic variation is depleted from $\mathcal{G}$ if the odds ratio $(A_G/(A - A_G))/(C_g/(2A - C_g))$ is less than 1.

To assess the significance of any enrichment or depletion we measure, we ask whether the corresponding log odds ratio $\log(A_G) + \log(2A - C_G) - \log(A - A_G) - \log(C_G)$ is more than two standard errors away from zero. The standard error of this log odds ratio is $\sqrt{1/A_G + 1/(2A - C_G) + 1/C_G + 1/(A - A_G)}$. In each forest plot presented in the manuscript, this formula was used to draw error bars that span two standard errors in each direction.

Detailed sampling procedure for matched control SNPs

For each archaic SNP set (Neanderthal 1, Neanderthal 2, Denisovan 1, and Denisovan 2) and each population p, we counted the number $A_p(b, c)$ of alleles with B-statistic value b and derived allele count c in population p, counting the allele as archaic if all derived alleles were annotated as present on archaic haplotypes in the relevant call set of population p. We then counted the number $N_p(b, c)$ of non-archaic alleles with B-statistic b and derived allele count c. In order for a SNP to count as non-archaic, none of its derived alleles could be present on a haplotype from population p that was called as archaic in either call set

1 or call set 2. A set $\mathcal{C}_p(b,c)$ of $2 \times A_p(b,c)$ control SNPs was then sampled uniformly at random without replacement from the $N_p(b,c)$ control candidate SNPs. In the rare event that $N_p(b,c) < 2 \times A_p(b,c)$, the control set was defined to be the entire set $N_p(b,c)$ and an extra $2 \times A_p(b,c) - N_p(b,c)$ SNPs from the control set were chosen unformly at random to be counted twice in all analyses.

Several analyses in the paper are performed on a merged set of archaic variation compiled across populations. To form the archaic SNP set $\mathcal{A}(b,c)$, we merged together the archaic SNP sets $\mathcal{A}_p(b,c)$ across all populations p. For each site where the derived allele was present in two or more populations, it was randomly assigned one population of origin. This population assignment process yielded new archaic allele counts $A'_p(b,c)$ that might be less than the counts $A_p(b,c)$ due to the deletion of duplicate SNPs. For each population p, we sampled $2 \times A'_p(b,c)$ control SNPs from population p as before and merged all of these control sets together to obtain a merged control set $\mathcal{C}(b,c)$. In the unlikely event that a single control allele is sampled in two or more populations, this control SNP will simply be counted two or more times during downstream analyses.

To obtain sets of distal archaic SNPs and controls, we must be careful about how we subtract call set 2 from call set 1. We want to sample control SNPs such that no control SNP is part of call set 2 for any archaic species in any population. To achieve this, the set of distal Denisovan SNPs $\mathcal{A}^{(D1-2)}(b,c)$ is defined as the set of all SNPs that are present in Denisovan call set 1 $\mathcal{A}^{(D1)}(b,c)$ but absent from both the Denisovan call set 2 $\mathcal{A}^{(D2)}(b,c)$ and the Neanderthal call set 2 $\mathcal{A}^{(N2)}(b,c)$. To generate the corresponding control set $\mathcal{C}^{(D1-2)}(b,c)$, we first look within each population to generate the superset of matched control SNPs $\mathcal{N}_p^{(D1-2)}(b,c)$. $\mathcal{N}_p^{(D1-2)}(b,c)$ is defined as the set of all SNPs present in population p in Denisovan call set 1 ($\mathcal{N}_p^{(D1)}(b,c)$) but absent from the population-merged sets of nonarchaic variants from Neanderthal Set 2 ($\mathcal{N}^{(N2)}(b,c)$) plus Denisovan Set 2 ($\mathcal{N}^{(D2)}(b,c)$). Once we have the population-specific candidate control sets $\mathcal{N}_p^{(D1-2)}(b,c)$, we randomly assign each archaic SNP from $\mathcal{A}^{(D1-2)}(b,c)$ to one of the populations where the derived allele is called as archaic, obtaining population-specific call sets $\mathcal{A}_p'^{(D1-2)}(b,c)$ that each contain $A_p'^{(D1-2)}(b,c)$ SNPs. As described earlier, we sample $2 \times A_p'^{(D1-2)}(b,c)$ control SNPs uniformly at random from each set $\mathcal{N}_p^{(D1-2)}(b,c)$ and merge these control sets together to obtain a merged set of distal controls $\mathcal{C}_p^{D1-2}$.

Quantifying singleton enrichment in the 1000 Genomes site frequency spectrum

Let $\mathcal{G}$ be a set of enhancers or other genomic regions. To test whether $\mathcal{G}$ is under stronger purifying selection than its immediate genomic neighborhood, we compared its site frequency spectrum (SFS) to the SFS of a region set $\mathcal{G}'$ defined as follows: $\mathcal{G}$ can always be defined as a collection of genomic intervals $\{(g_1^{(1)}, g_1^{(2)}), \ldots, (g_n^{(1)}, g_n^{(2)})\}$, where each $(g_k^{(1)}, g_k^{(2)})$ is a pair of genomic coordinates delineating a piece of DNA contained entirely within the set $\mathcal{G}$. We define $\mathcal{G}'$ to be the collection of genomic intervals $\{(2 \times g_k^{(1)} - g_k^{(2)}, g_k^{(k)})\}$, i.e. the set of intervals immediately adjacent on the left to the intervals that make up $\mathcal{G}$. (We are slightly abusing notation here by failing to note that different chromosomes have different coordinate systems).

We computed folded site frequency spectra for $\mathcal{G}$ and $\mathcal{G}'$ using the African individuals in the 1000 Genomes Phase 3 VCF, excluding SNPs that do not pass the VCF's default quality filter. Let S_G and $S_{G'}$ be the numbers of singletons that fall in into the regions $\mathcal{G}$ and $\mathcal{G}'$, respectively, and let N_G and $N_{G'}$ be the numbers of non-singleton variants that fall into these regions. We say that $\mathcal{G}$ is enriched for singletons if the odds ratio $(S_G/N_G)/(S_{G'}/N_{G'})$ is greater than 1. To assess the significance of any enrichment or depletion, we use the fact that the standard error of this binomial test is $\sqrt{1/S_G + 1/N_G + 1/S_{G'} + 1/N_{G'}}$. All singleton enrichment plots in this manuscript contain error bars that span 2 standard errors above and below the estimated odds ratio.

References

[1] Green, R. E. *et al.* A draft sequence of the Neandertal genome. *Science* **328**, 710–722 (2010).

[2] Prüfer, K. *et al.* The complete genome sequence of a Neanderthal from the Altai mountains. *Nature* **505**, 43–49 (2014).

[3] Sankararaman, S. *et al.* The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* **507**, 354–357 (2014).

[4] Vernot, B. & Akey, J. M. Resurrecting surviving neandertal lineages from modern human genomes. *Science* **343**, 1017–1021 (2014).

[5] Sankararaman, S., Mallick, S., Patterson, N. & Reich, D. The combined landscape of Denisovan and Neanderthal ancestry in present-day humans. *Current Biology* **26**, 1241–1247 (2016).

[6] Vernot, B. *et al.* Excavating Neanderthal and Denisovan DNA from the genomes of Melanesian individuals. *Science* 10.1126/science.aad9416 (2016).

[7] Petr, M., Pääbo, S., Kelso, J. & Vernot, B. Limits of long-term selection against Neanderthal introgression. *Proc Natl Acad Sci USA* **116**, 1639–1644 (2019).

[8] King, M. & Wilson, A. Evolution at two levels in humans and chimpanzees. *Science* **188**, 107–116 (1975).

[9] Enard, W. *et al.* Intra- and interspecific variation in primate gene expression patterns. *Science* **296**, 340–343 (2002).

[10] Wray, G. The evolutionary significance of *cis*-regulatory mutations. *Nature Reviews Genetics* **8**, 206–216 (2007).

[11] McCoy, R., Wakefield, J. & Akey, J. Impacts of Neanderthal-introgressed sequences on the landscape of human gene expression. *Cell* **168**, 916–927 (2017).

[12] Dannemann, M., Prüfer, K. & Kelso, J. Functional implications of Neanderthal introgression in modern humans. *Genome Biology* **18**, 61 (2017).

[13] Hahn, M. Detecting natural selection on *cis*-regulatory dna. *Genetica* **129**, 7–18 (2007).

[14] Castellano, S. *et al.* Patterns of coding variation in the complete exams of three Neanderthals. *Proc Natl Acad Sci USA* **111**, 6666–6671 (2014).

[15] Wong, W. & Nielsen, R. Detecting selection in noncoding regions of nucleotide sequences. *Genetics* **167**, 949–958 (2004).

[16] Torgerson, D. *et al.* Evolutionary processes acting on candidate *cis*-regulatory regions in humans inferred from patterns of polymorphism and divergence. *PLoS Genetics* **5**, e1000592 (2009).

[17] Ward, L. & Kellis, M. Evidence of abundant purifying selection in humans for recently acquired regulatory functions. *Science* **337**, 1675–1678 (2012).

[18] Smith, J., McManus, K. & Fraser, H. A novel test for selection on *cis*-regulatory elements reveals positive and negative selection acting on mammalian transcriptional enhancers. *Molecular Biology and Evolution* **30** (2013).

[19] Arbiza, L. *et al.* Genome-wide inference of natural selection on human transcription factor binding sites. *Nature Genetics* **45**, 723–729 (2013).

[20] Huerta-Sánchez, E. *et al.* Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* **512**, 194–197 (2014).

[21] Prabhakar, S. *et al.* Human-specific gain of function in a developmental enhancer. *Science* **321**, 1346–1350 (2008).

[22] Capra, J., Erwin, G., McKinsey, G., Rubenstein, J. & Pollard, K. Many human accelerated regions are developmental enhancers. *Phil Trans R Soc B* **368**, 20130023 (2013).

[23] Rinker, D. *et al.* Neanderthal introgression reintroduced functional alleles lost in the human out of Africa bottleneck. *bioRxiv preprint* doi.org/10.1101/533257 (2019).

[24] The ENCODE Project Consortium. The ENCODE (Encyclopedia of DNA Elements) Project. *Science* **306**, 636–640 (2004).

[25] Roadmap Epigenomics Consortium. Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).

[26] Silvert, M., Quintana-Murci, L. & Rotival, M. Impact and evolutionary determinants of Neanderthal introgression on transcriptional and post-transcriptional regulation. *The American Journal of Human Genetics* **104**, 1241–1250 (2019).

[27] Mallick, S. *et al.* The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538**, 201–206 (2016).

[28] McVicker, G., Gordon, D., Davis, C. & Green, P. Widespread genomic signatures of natural selection in hominid evolution. *PLoS Genetics* **5**, e1000471 (2009).

[29] Enard, D. & Petrov, D. Evidence that RNA viruses drove adaptive introgression between Neanderthals and modern humans. *Cell* **175**, 360–371 (2018).

[30] Quach, H. *et al.* Genetic adaptation and Neandertal admixture shaped the immune system of human populations. *Cell* **167**, 643–656 (2016).

[31] Nédélec, Y. *et al.* Genetic ancestry and natural selection drive population differences in immune responses to pathogens. *Cell* **167**, 657–669 (2016).

[32] Dannemann, M., Andrés, A. & Kelso, J. Introgression of Neanderthal- and Denisovan-like haplotypes contributes to adaptive variation in human toll-like receptors. *American Journal of Human Genetics* **98**, 22–33 (2016).

[33] Skov, L. *et al.* Detecting archaic introgression using an unadmixed outgroup. *PLoS Genetics* **14**, e1007641 (2018).

[34] Browning, S., Browning, B., Zhou, Y., Tucci, S. & Akey, J. Analysis of human sequence data reveals two pulses of archaic Denisovan admixture. *Cell* **173**, 53–61 (2018).

[35] Hubisz, M., Williams, A. & Simple, A. Mapping gene flow between ancient hominids through demography-aware inference of the ancestral recombination graph. *bioRxiv preprint* doi.org/10.1101/687368 (2019).

[36] Sawyer, S. & Hartl, D. Population genetics of polymorphism and divergence. *Genetics* **132**, 1161–1176 (1992).

[37] Fish, A., Chen, L. & Capra, J. Gene regulatory enhancers with evolutionarily conserved activity are more pleiotropic than those with species-specific activity. *Genome Biology and Evolution* **9**, 2615–2625 (2017).

[38] Stone, J. & Wray, G. Rapid evolution of *cis*-regulatory sequences via local point mutations. *Molecular Biology and Evolution* **18**, 1764–1770 (2001).

[39] MacArthur, S. & Brookfield, J. Expected rates and modes of evolution of enhancer sequences. *Molecular Biology and Evolution* **21**, 1064–1073 (2004).

[40] Nord, A. *et al.* Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. *Cell* **155**, 1521–1531 (2013).

[41] Harris, K. & Nielsen, R. The genetic cost of Neanderthal introgression. *Genetics* **203**, 881–891 (2016).

[42] Juric, I., Aeschbacher, S. & Coop, G. The strength of selection against Neanderthal introgression. *PLoS Genetics* **12**, e1006340 (2016).

[43] Schumer, M. *et al.* Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science* **360** (2018).

[44] Ruff, C., Trinkaus, E. & Holliday, T. Body mass and encephalization in Pleistocene *Homo*. *Nature* **387**, 173–176 (1997).

[45] Churchill, S. Bioenergetic perspectives on Neanderthal thermoregulatory and activity budgets. In Hublin, J., Harvati, K. & Harrison, T. (eds.) *Neanderthals Revisited: New Approaches and Perspectives*, 113–133 (Springer, Dordrecht, 2006).

[46] Gulsuner, S. *et al.* Spatial and temporal mapping of *de novo* mutations in schizophrenia to a fetal prefrontal cortical network. *Cell* **154**, 518–529 (2013).

[47] Turner, T. *et al.* Genomic patterns of de novo mutation in simplex autism. *Cell* **171**, 710–722 (2017).

[48] Short, P. *et al.* *De novo* mutations in regulatory elements in neurodevelopment disorders. *Nature* **555**, 611–616 (2018).

[49] Steinrücken, M., Spence, J., Kamm, J., Wieczorek, E. & Song, Y. Model-based detection and analysis of introgressed Neanderthal ancestry in modern humans. *Molecular Ecology* **27**, 3873–3888 (2018).

[50] Khaitovich, P. *et al.* Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* **309**, 1850–1854 (2005).

[51] Coolon, J., McManus, C., Stevenson, K., Graveley, B. & Wittkopp, P. Tempo and model of regulatory evolution in *Drosophila*. *Genome Research* **24**, 797–808 (2014).

[52] Turner, L., White, M., Tautz, D. & Payseur, B. Genomic networks of hybrid sterility. *PLoS Genetics* **10**, e1004162 (2014).

[53] Mack, K., Campbell, P. & Nachman, M. Gene regulation and speciation in house mice. *Genome Research* **26**, 451–461 (2016).

[54] Lewis, J., van der Burg, K., Mazo-Vargas, A. & Reed, R. ChIP-Seq-Annotated *Heliconius erato* genome highlights patterns of *cis*-regulatory evolution in lepidoptera. *Cell reports* **16**, 2855–2863 (2016).

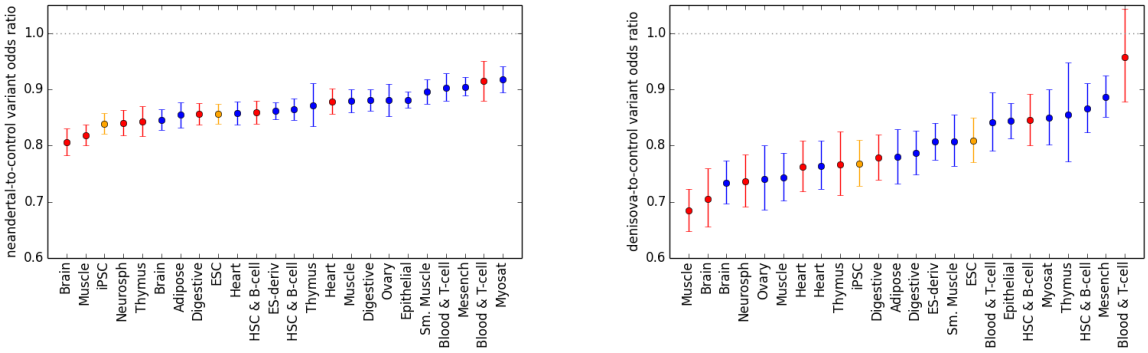

Figure 3– Figure Supplement 1 These plots show the data from Figure 3 with Neanderthal and Denisovan odds ratios on separate plots for clarity.

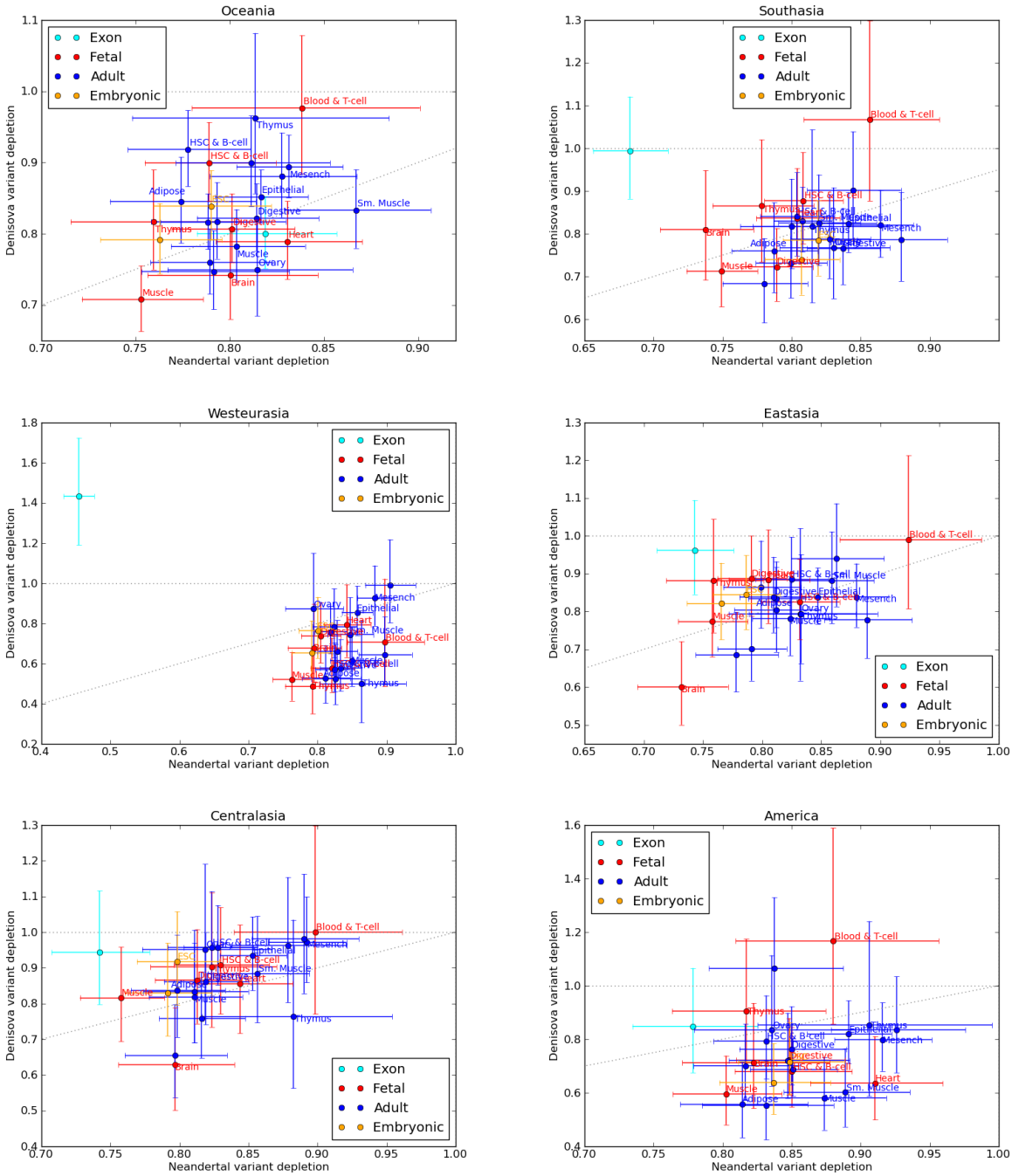

Figure 3– Figure Supplement 2 These plots show the joint distribution of Neanderthal and Denisovan introgression depletion within each SGDP population separately. Although there are differences between populations, particularly since Denisovan introgression is sparse and noisy in general, all show that brain and fetal muscle enhancers are the most depleted of introgression. In most populations the blood & T-cell tissue is least depleted of introgression.

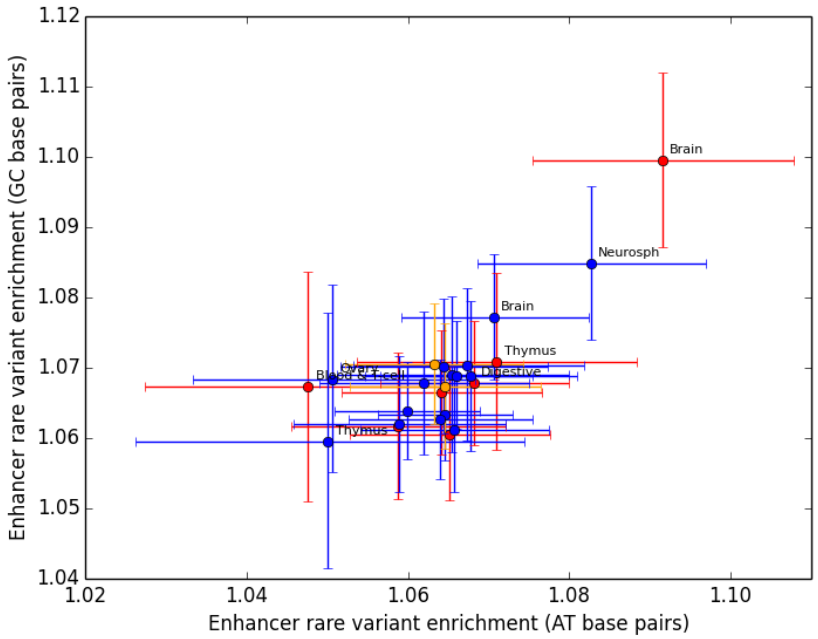

Figure 5– Figure Supplement 1 This figure was generated by partitioning the site frequency spectrum of each enhancer between SNPs that have GC ancestral alleles and SNPs that have AT ancestral alleles. The site frequency spectra of these two classes of sites are expected to be driven in opposite directions by GC biased gene conversion. However, the finding that brain enhancers are enriched for singletons holds up when we restrict to either GC-ancestral SNPs or AT-ancestral SNPs.

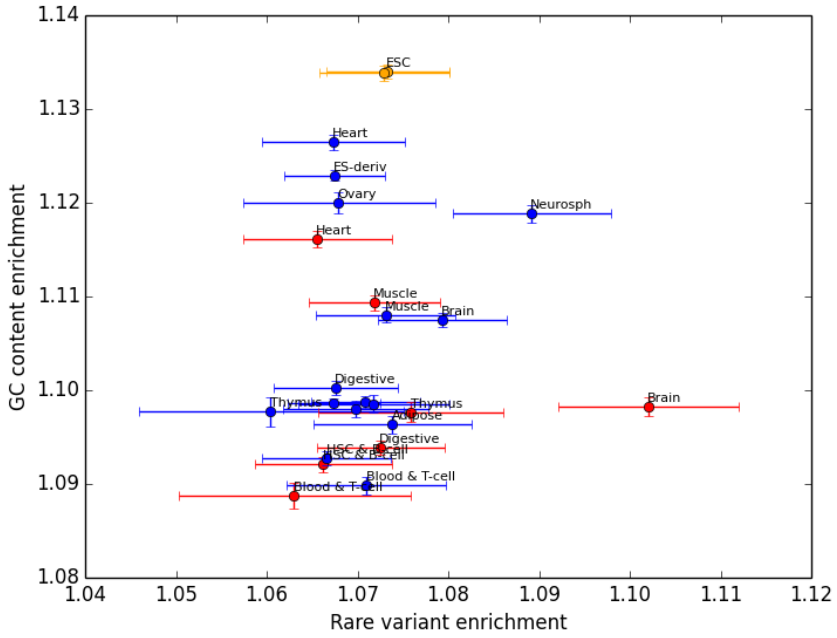

Figure 5– Figure Supplement 2 Enhancers are enriched for GC base pairs compared to adjacent genomic regions, and the degree of this enrichment varies between tissues. However, there is no correlation across tissues between GC content enrichment and the singleton enrichment that we attribute to purifying selection.

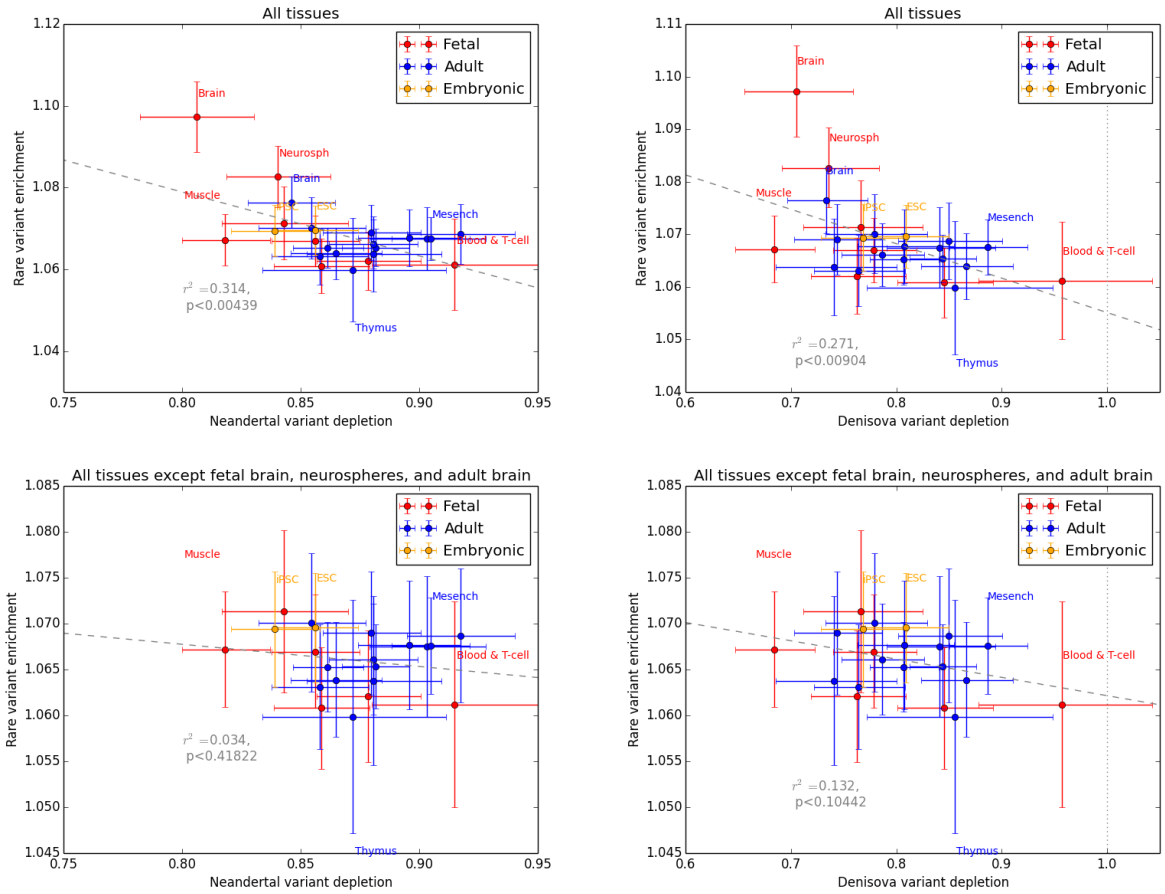

Figure 5– Figure Supplement 3 Although singleton enrichment is negatively correlated between tissues with both Neanderthal and Denisovan variant depletion, the significance of this correlation disappears when all brain related tissues are excluded from the regression.