FEATURED ARTICLE

INVESTIGATION

G3 Genes | Genomes | Genetics

Signatures of Recent Positive Selection in Enhancers Across 41 Human Tissues

Jiyun M. Moon,* John A. Capra,*,†,‡ Patrick Abbot,* and Antonis Rokas*,†,‡,1

*Department of Biological Sciences, †Vanderbilt Genetics Institute, and ‡Departments of Biomedical Informatics and Computer Science, Vanderbilt University, Nashville, TN, 37235

ORCID ID: 0000-0002-7248-6551 (A.R.)

ABSTRACT Evolutionary changes in enhancers are widely associated with variation in human traits and diseases. However, studies comprehensively quantifying levels of selection on enhancers at multiple evolutionary periods during recent human evolution and how enhancer evolution varies across human tissues are lacking. To address these questions, we integrated a dataset of 41,561 transcribed enhancers active in 41 different human tissues (FANTOM Consortium) with whole genome sequences of 1,668 individuals from the African, Asian, and European populations (1000 Genomes Project). Our analyses based on four different metrics (Tajima's *D*, F_{ST}, H12, nS_L) showed that ~5.90% of enhancers showed evidence of recent positive selection and that genes associated with enhancers under very recent positive selection are enriched for diverse immune-related functions. The distributions of these metrics for brain and testis enhancers were often statistically significantly different and in the direction suggestive of less positive selection compared to those of other tissues; the same was true for brain and testis enhancers that are tissue-specific compared to those that are tissue-broad and for testis enhancers associated with tissue-enriched and non-tissue-enriched genes. These differences varied considerably across metrics and tissues and were generally in the form of changes in distributions' shapes rather than shifts in their values. Collectively, these results suggest that many human enhancers experienced recent positive selection throughout multiple time periods in human evolutionary history, that this selection occurred in a tissue-dependent and immune-related functional context, and that much like the evolution of their protein-coding gene counterparts, the evolution of brain and testis enhancers has been markedly different from that of enhancers in other tissues.

KEYWORDS
enhancers
human evolution
recent positive selection
immunity
testis
brain

Enhancers are *cis*-acting DNA segments that, either independently of or in concert with other regulatory elements, control spatial, temporal and quantitative aspects of gene expression (Ong and Corces 2011; Rubinstein and de Souza 2013; Long *et al.* 2016). While the precise architecture of enhancers is still debated (Long *et al.* 2016), a typical enhancer contains multiple transcription factor binding sites (TFBS) arranged in specific order and distance from one another. Enhancers facilitate initiation of gene transcription by helping to recruit RNA Polymerase II, general transcription factors, and additional components of the transcriptional machinery to the gene's promoter (Pennacchio *et al.* 2013; Rubinstein and de Souza 2013). Multiple enhancers, each with its own repertoire of TFBS, can regulate the activities of a gene in a tissue-specific manner or across distinct developmental stages, enabling enhancers to alter its expression patterns in a particular context without affecting expression of other genes (Wray 2007; Sholtis and Noonan 2010). Enhancers in the human genome are more numerous than protein-coding genes (Pennacchio *et al.* 2013), facilitating the induction of diverse gene expression programs in different spatial and temporal contexts (Long *et al.* 2016).

Changes in gene regulation have long been thought to play a major role in the adaptive evolution of human traits (King and Wilson 1975; Carroll 2005). One reason for regulatory regions in general, and enhancers in particular, as preferential targets of selection is that,

doi: https://doi.org/10.1534/g3.119.400186
Manuscript received March 13, 2019; accepted for publication June 13, 2019; published Early Online June 18, 2019.

Supplemental material available at FigShare: https://doi.org/10.25387/g3.8230757.
[1]Corresponding Author: Vanderbilt University, VU Station B 351634, Nashville, TN 37235 ; E-mail: antonis.rokas@vanderbilt.edu

compared to protein-coding regions, they tend to be modularly organized (Sholtis and Noonan 2010). This modular organization means that mutations in enhancers are less likely to have pleiotropic effects and more likely to contribute to phenotypic evolution (Carroll 2005; Wray 2007; Rebeiz and Tsiantis 2017). In the context of human evolution, several studies suggest that evolutionary changes in enhancers might have played a major role in the acquisition of human-specific traits. For example, a considerable proportion of regions that has experienced accelerated evolution in the human lineage are developmental enhancers active in the brain (Capra *et al.* 2013), and *cis*-regulatory regions of genes with neurological and nutritional roles have been shown to exhibit evidence of accelerated evolution in the human lineage (Rockman *et al.* 2005; Haygood *et al.* 2007). Accelerated evolution is one signature of positive selection, and several studies suggest that recent selection might also have preferentially acted on human *cis*-regulatory regions (Wray 2007). For instance, Rockman *et al.* found that the promoter region of a gene that encodes the precursor of an opioid neuropeptide (*PDYN*) exhibits a significant degree of population differentiation between human populations, especially between Europeans and East Asians, which is suggestive of local adaptation (Rockman *et al.* 2005). Similarly, patterns of variation in the promoter region of the *LCT* gene that confers lactase persistence in Africans are consistent with the action of selective sweeps (Tishkoff *et al.* 2007). Finally, SNPs with evidence of recent positive selection are more likely to be associated with expression of nearby genes than random SNPs, and this enrichment is strongest for Yorubans (Kudaravalli *et al.* 2008).

More broadly, genome-wide studies have found that, compared to protein-coding regions, enhancers are enriched for variants that are statistically associated with various human diseases (*e.g.*, inflammatory diseases, metabolic diseases) (Andersson *et al.* 2014; Karnuta and Scacheri 2018). Moreover, recent studies have argued that human population-level differences in transcriptional responses to infection likely resulted from local adaptation, further supporting the idea that selection on enhancers has contributed to recent human evolution (Nédélec *et al.* 2016). In addition, multiple events, such as migrations and shifts in cultural practices, have occurred in different time periods during recent human evolution (Karlsson *et al.* 2014), likely introducing novel selective agents. Given the role of enhancers in contributing to phenotypic differences, it is likely that enhancers experienced selection events in response to such selective pressures.

Each human tissue serves a particular physiological function so it is reasonable to hypothesize that the genetic elements active in each tissue are influenced by distinct selective pressures that shape their evolutionary rates (Wray 2007; Gu and Su 2007). For example, both gene expression and protein-coding sequence divergence patterns might be expected to be more conserved in developmentally constrained tissues (*e.g.*, nervous tissues) than in developmentally relaxed tissues (*e.g.*, testis or endocrine tissues, such as pancreas). Alternatively, certain functions of tissues (*e.g.*, reproductive processes for testis) might have experienced increased levels of positive selection (Khaitovich 2005). Early studies examining patterns of divergence in gene expression and protein-coding sequence evolution among mammals found that both were lowest for nervous tissues (*e.g.*, brain, cerebellum) and greatest for testis (Wray 2007; Khaitovich 2005; Brawand *et al.* 2011). More recent examination of a larger number of human tissues has provided further support for this hypothesis. For example, the correlation between expression levels of genes and the strength of purifying selection (as assessed by dN/dS) is strongest for the brain, while this correlation is lowest for liver, placenta and testis (Kryuchkova-Mostacci and Robinson-Rechavi 2015).

Hitherto, studies investigating the variation of evolutionary patterns among tissues have mainly focused on inter-species divergence of gene expression and of protein-coding sequences. However, studies comprehensively quantifying levels of selection on enhancers at multiple evolutionary time periods during recent human evolution and how enhancer evolution varies across human tissues are lacking. Furthermore, it is plausible, if not likely, that distinct selective pressures are also acting on regulatory elements in different tissues, further contributing to the global differences in gene expression patterns among the tissues described above (Ong and Corces 2011; Rubinstein and de Souza 2013).

To examine the influence of selection on enhancers at multiple evolutionary time periods during recent human evolution and how enhancer evolution varies across human tissues, we used a dataset of 41,561 enhancers active in 41 human tissues and genotype data of 1,668 individuals from three different super-populations from the 1000 Genomes Project to calculate signatures of recent selection that happened at different time points and via different modes (*i.e.*, hard *vs.* soft selective sweeps). We found that on average, 5.90% of enhancers exhibit evidence of recent positive selection and that their putative target genes are enriched for immunity-related functions. Furthermore, we found that the distributions of positive selection metrics for enhancers expressed in the testis and brain exhibit were often statistically different and in the direction of lower positive selection than those of numerous other tissues. We found the same trend for the distributions of positive selection metrics for enhancers active in only brain or testis *vs.* enhancers that were also active in other tissues as well as when we studied enhancers associated with tissue-enriched and non-tissue-enriched genes in brain and testis. Our results suggest that human enhancers, in particular ones associated with immune-related genes, have experienced different modes of recent positive selection at different periods of recent human evolution. Furthermore, patterns of selection on human enhancers differ between tissues, including for brain and testis, a pattern reminiscent of their protein-coding counterparts.

METHODS

Genetic Variation Data

To examine signatures of natural selection on human enhancers, we used the whole genome sequence data for 1,668 individuals from Phase 3 of the 1000 Genomes Project (The 1000 Genomes Project Consortium *et al.* 2015) (for information on data sources, see File S1). The 1,668 individuals represent three major human populations (661 Africans, 503 Europeans, and 504 East Asians). We used these three populations because their demographic histories have been most confidently estimated and excluded the other two (Admixed Americans and South Asians) because their demographic histories are more complex.

FANTOM Enhancer Data

To study patterns of recent evolution on human enhancers active in different tissues, we used the enhancers detected in 41 human tissues compiled by the Functional Annotation of the Mammalian Genomes (FANTOM) project (Figure 1 & File S1) (Andersson *et al.* 2014) based bidirectional transcription identified via cap analysis of gene expression (CAGE) in diverse human cell and tissue types (The FANTOM Consortium and the RIKEN PMI and CLST (DGT) *et al.* 2014). We further classified the enhancers into 'tissue-specific' enhancers that are detected in only a single tissue, and 'tissue-broad' enhancers that are detected in two or more tissues. Of the tissue-broad enhancers, only a small proportion (833 out of 6,757 enhancers; 12.33%) belonged only to one organ system; 39.35% (2,659 enhancers), 18.12% (1,224 enhancers), 10.14% (685 enhancers), 6.16% (416 enhancers), 4.51% (305 enhancers),

2.86% (193 enhancers), 2.77% (187 enhancers), 1.75% (118 enhancers), and 2.03% (137 enhancers) of tissue-broad enhancers belonged to two, three, four, five, six, seven, eight, nine, and ten organ systems, respectively. As genomic regions located on sex chromosomes exhibit patterns of genetic variation that differ from those observed in autosomal genomic regions (Schaffner 2004), we restricted our analyses to only autosomal enhancers.

Human Protein Atlas Data

To compare the patterns of recent evolution between enhancers associated with genes that exhibit higher expression levels in one tissue compared to others and those that are not in our tissues of interest, the brain and testis, we downloaded the list of tissue-enriched genes for the brain and testis tissues from the Human Protein Atlas database (Uhlen *et al.* 2015) (version 88; download date: January 17th, 2018): the Human Protein Atlas Database defines a gene as tissue-enriched if its mRNA levels in a given tissue are at least fivefold higher compared to its levels of expression in all other tissues. We defined 'non-tissue-enriched' genes as those genes that are not labeled as 'tissue-enriched' according the to the Protein Atlas database.

Identifying Signatures of Recent Positive Selection

To identify signatures of recent positive selection on human enhancers, we calculated four metrics aimed to capture genetic signatures left from the action of selection at different time ranges during recent human evolution (Sabeti *et al.* 2006; Vitti *et al.* 2013; Moon *et al.* 2018) using previously described methodology (Moon *et al.* 2018). Briefly, the four metrics that we used can be divided into two categories: those that are designed to detect hard selective sweeps (Tajima's D, Weir & Cockerham's F_{ST} and nS_L) and those designed to detect soft selective sweeps (H12). Furthermore, the three tests for hard selective sweeps are most sensitive to selection events that occurred at different time points in human history: Tajima's D can detect selection that happened approximately 250,000 - 200,000 years ago; F_{ST} can identify signatures of local selection that occurred following the out-of-African migration approximately 75,000 - 50,000 years ago; finally, nS_L has power to detect signatures left by selection events that occurred approximately 20,000 - 10,000 years ago (Sabeti *et al.* 2006; Moon *et al.* 2018). Following the methods of Moon *et al.* (Moon *et al.* 2018) for each enhancer region in every tissue we calculated Tajima's D using the R package *PopGenome*, version 2.1.6 (Pfeifer *et al.* 2014) and the weighted Weir & Cockerham's fixation index (F_{ST}) (Weir & Cockerham 1984) among all

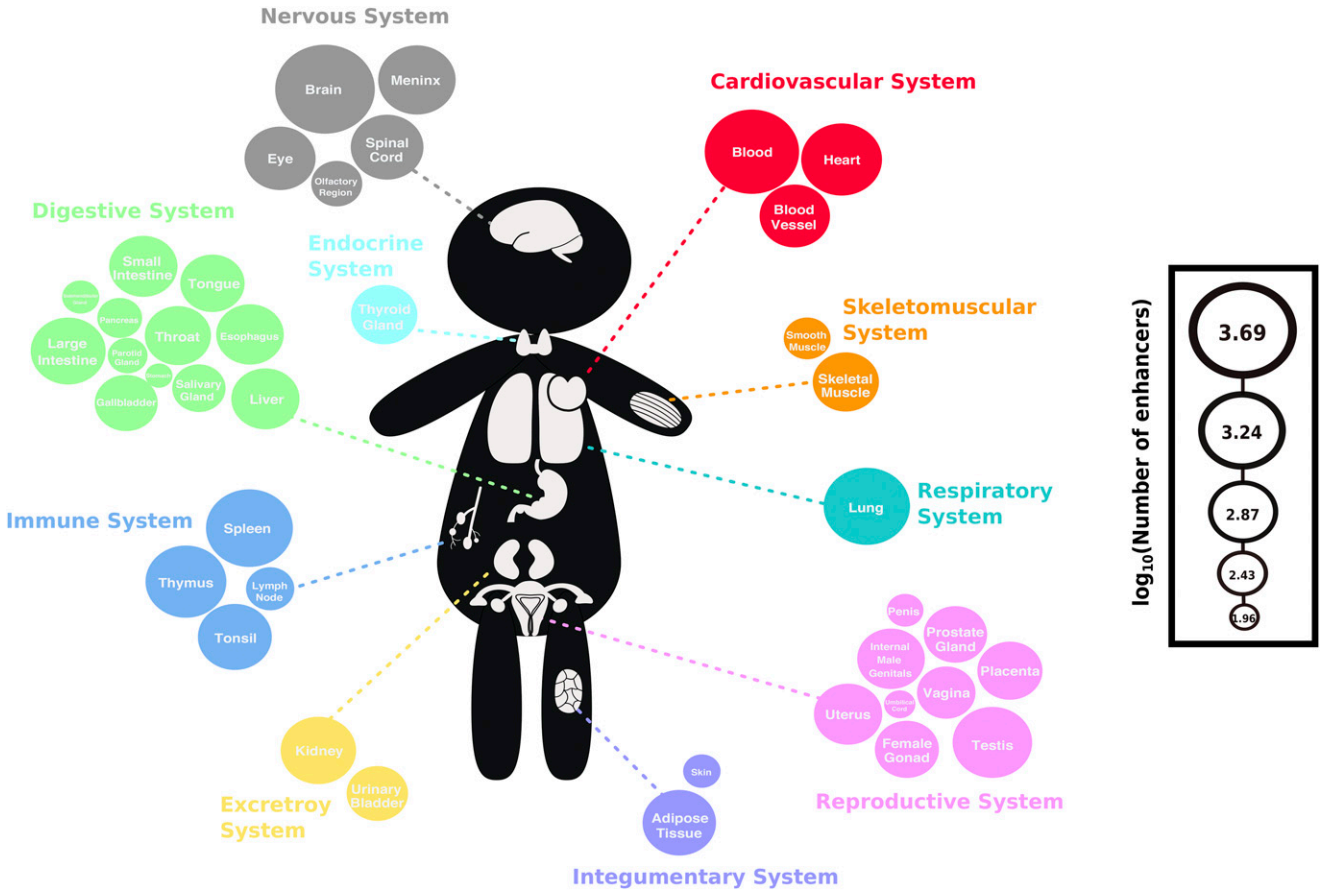

Figure 1 Visual summary of the FANTOM5 enhancer data set used in this study Each circle represents a tissue and all 41 tissues are grouped into non-overlapping organ systems and are color-coded accordingly. The size of the circle is proportional to the $\log_{10}$ of the number of enhancers active within each tissue (Table S1: from highest to lowest: brain: 4,883, blood: 3,657, spleen: 2,429, lung: 2,366, thymus: 1,741, heart: 1,737, testis: 1,621, meninx: 1,447, kidney: 1,294, large-intestine: 1,266, tonsil: 1,193, adipose-tissue: 1,161, spinal-cord: 1,115, eye: 1,019, blood-vessel: 982, small-intestine: 880, uterus: 877, liver: 875, internal-male-genitalia: 846, esophagus: 845, thyroid-gland: 795, skeletal-muscle: 784, throat: 782, placenta: 735, tongue: 708, female-gonad: 693, prostate-gland: 667, gallbladder: 654, urinary-bladder: 602, vagina: 512, salivary-gland: 387, olfactory-region: 338, lymph-node: 292, smooth-muscle: 270, pancreas: 243, parotid-gland: 184, skin: 161, submandibular-gland: 159, penis: 154, umbilical-cord: 115, stomach: 92).

three populations using *VCFtools*, version 0.1.13 (Danecek *et al.* 2011). In addition, we calculated the nS_L statistic (number of segregating sites by length; (Ferrer-Admetlla *et al.* 2014)) using *Selscan*, version 1.2.0 (Szpiech and Hernandez 2014). For each enhancer, we calculated nS_L for the region extending 50 kilobases (kb) upstream and downstream using the default settings of *Selscan*, except for the minor allele frequency (MAF) cut-off value, for which we used 0.01. We calculated this metric for each of the three populations, as haplotype-based methods are known to detect selection events that occurred after the out-of-African migration (Voight *et al.* 2006; Sabeti *et al.* 2006). We used the maximum absolute un-standardized nS_L values calculated over the entire window to represent each enhancer region. Finally, for each enhancer region we also calculated the H12 index using a custom script, based on Garud *et al.*'s original script (Garud *et al.* 2015). The values of all enhancers in the 41 human tissues for all metrics can be found in File S2.

To test for statistical differences in the distributions of each of the four metrics between any two tissues, we also carried out pairwise Kolmogorov-Smirnov tests (two-sided). Visualization of the data (Figure S1) revealed that assumptions of normality and equal variance could not be upheld, and therefore, we chose the non-parametric Kolmogorov-Smirnov test which makes no assumptions regarding the shape of the distributions and carried out this test in the R programming environment. We also conducted this test to examine whether there are any significant differences between the distributions of patterns of recent evolution between tissue-specific and tissue-broad enhancers, as well as between enhancers that are associated with tissue-enriched and non-tissue-enriched genes. All statistical tests were followed by *post hoc* Bonferroni corrections in R. The results of the statistical tests for all metrics can be found Tables S22-27, 28-33.

Simulations of Neutral Evolution

The action of non-adaptive processes and demographic changes, such as genetic drift and population expansion, can produce patterns of variation that these tests may mistake for positive selection (Sabeti *et al.* 2006). To determine the likelihood of the empirical values being generated by the action of selection, we carried out simulations of neutral evolution based on a model of recent human demographic history and compared the observed values to expected values under neutrality. Simulations of neutrality were conducted using *SLiM*, version 2.4.1 (Messer 2013; Haller and Messer 2017). Following Gravel *et al.*, we used previously calculated demographic parameters for the three populations included in our study (Gravel *et al.* 2011; Messer 2013). Detailed information on the parameters used for the simulations can be found in the Supplementary File S1.

We simulated the genotypes of 661, 503, and 504 individuals corresponding to the numbers of individuals from each of the three populations analyzed. Since there are 41,561 autosomal enhancers in the FANTOM dataset and anywhere from 92 (stomach) to 4,883 (brain) enhancers in any given tissue (Table S1), carrying out individual simulations for each enhancer was computationally prohibitive. To reduce the computational load of our simulations, we focused on simulating the mutational profile of an average enhancer for a given tissue under a model of neutral evolution. Specifically, to carry out the neutral simulations for enhancers in a given tissue, we used the average length of all the autosomal enhancers found in that tissue (these values can be found in File S3) and a fixed recombination rate of $1x10^{-8}$. For each tissue, we carried out 10,000 simulations. We next used these simulated sequences to calculate Tajima's *D* values, Weir & Cockerham's F_{ST}, and H12 as described above. We calculated an empirical *p*-value for a given enhancer for any metric (including for nS_L, the simulation process of which is described below) as the proportion of simulated sequences that obtained scores equal to or more extreme than the observed value. We used a *p*-value of 0.05 as cutoff for significance; enhancers with *p*-values lower than 0.05 were considered to significantly deviate from neutral evolution and to have experienced selection.

Recombination Rate Interpolation and Simulation of nS_L Values

As variation in recombination rates can affect the values of nS_L (Ferrer-Admetlla *et al.* 2014), we carried out a separate set of neutral simulations for the calculation of nS_L values by incorporating average recombination rates for each tissue. To do this, we used the genome-wide genetic map curated by the HapMap II consortium (International HapMap Consortium *et al.* 2007), which provides pre-computed recombination rates (cM/Mb) for the variants included in the HapMap II project. More specifically, we carried out linear interpolation to infer the recombination rates of the variants in the 1000 Genomes Project dataset that are not included in the HapMap Phase II dataset. Next, we calculated the average recombination rate (cM/Mb) for the region spanning 50 kb up and downstream of each enhancer in a given tissue, and then used the average of all recombination rates of the enhancer regions in a given tissue as the recombination rate parameter (converted to probability of crossovers per bp) for the neutral simulations. The average recombination rates, as used in our *SLiM* simulations, for each tissue can be found in File S4. Using these calculated recombination rates, we then created 2,500 enhancers that span 50 kb upstream and downstream of the average length of all autosomal enhancers found in that given tissue and calculated nS_L values on these simulated enhancers using *Selscan* as described above. As with the actual enhancers, we only included variants with MAF greater than 0.01 in the calculations and used the maximum absolute un-standardized values to represent a given simulated window.

Functional Enrichment Analyses and Semantic Similarity Calculations

To gain insight into the functions of genes associated with enhancers that show evidence of selection, we next carried out functional enrichment analyses. To associate enhancers with putative target genes, we used the transcription start site (TSS)-enhancer mapping file generated by Andersson *et al.* (Andersson *et al.* 2014) (file location information can be found in File S1), which was compiled by measuring the pairwise correlation between enhancer activity and transcription level of putative target genes. We carried out functional enrichment analyses on the list of these putative target genes using the R package *TopGO*, version 2.32.0 (Alexa and Rahnenfuhrer 2016). For each enrichment analysis within a given tissue, we used the set of all putative target genes associated with all the autosomal enhancers in that tissue as the background. In short, the GO terms associated with the putative target genes of enhancers with significant evidence of recent positive selection in a tissue were compared against the GO terms of all putative target genes of all autosomal enhancers in that tissue. Detailed description of the files and commands used can be found in File S1. In short, we used the 'weight' algorithm that compares the significance scores of the connected nodes to explicitly account for the hierarchy of the gene ontology tree and carried out analyses for the three general ontologies: Biological Process (BP), Molecular Function (MF), and Cellular Component (CC). Subsequent corrections for multiple comparisons were carried out by calculating the FDR-adjusted *p*-values in the R statistical environment; only those GO terms with an FDR-adjusted *p*-value less

than 0.05 were retained for further analyses. All *TopGO* enrichment analysis results for all the metrics can be found in Files S5-21.

To quantify the overall similarity of the patterns of gene functional enrichment among different tissues, we calculated pairwise semantic similarity of the GO terms between any two tissues using a graph-based method (Wang *et al.* 2007), which takes into consideration the topology of the GO graph structure (*i.e.*, the location of the GO terms in the graph and the relationship of the terms to their ancestor terms), as implemented in the R package *GOSemSim*, version 2.4.1 (Yu *et al.* 2010): this analysis was carried out for only those tissues that have 5 or more significant GO terms. To determine if the semantic similarity scores thus calculated significantly deviated from random expectations, we randomly sampled the same number of GO terms for each tissue from the pool of all GO terms associated with a given tissue 1,000 times and calculated the pairwise semantic similarity score; scores equal or greater than the 95th percentile value of the distribution of semantic similarity scores were considered significant. The calculated semantic similarity values, along with the 95th percentile values obtained from the 1,000 randomly sampled GO terms for each tissue pair, can be found in Tables S7-15.

Transposable Element (TE) Data and Analyses

TE-derived sequences often contribute to the origin of new enhancers and the modification of regulatory networks (Lynch *et al.* 2015; Simonti *et al.* 2017). To determine if the enhancers with evidence of recent positive selection harbor higher proportions of TEs than those with nonsignificant deviations from neutral expectations, we used BEDOPS, version 2.4.35 (Neph *et al.* 2012) to overlap the enhancer regions with the RepeatMasker-annotated regions. We downloaded the RepeatMasker annotation track for the hg19 assembly from the UCSC Genome Browser (Kent *et al.* 2002; Casper *et al.* 2018; Raney *et al.* 2014). The location of the RepeatMasker annotation file and the specific commands used for BEDOPS can be found in Supplementary File S1. To test for statistical differences in the proportions of TEs between enhancers with evidence of recent selection and those that do not exhibit such significant deviation from neutral expectations, we carried out a 2x2 chi-squared test with Yate's continuity correction using R. The results of the chi-squared tests can be found in Tables S16-21.

GWAS Catalog Data Annotation

Previous studies have shown that many variants in human *cis*-regulatory regions are significantly associated with a broad range of complex traits and diseases (Lee and Young 2013; Andersson *et al.* 2014; Karnuta and Scacheri 2018). To investigate whether brain and testis enhancers that exhibit evidence of recent positive selection show enrichment of variants in the Genome-Wide Association Study (GWAS) catalog, we downloaded the NHGRI-EBI GWAS catalog, version1.0.2 (MacArthur *et al.* 2017) (download date: June 2nd, 2018) and first compared the SNPs that reside in the enhancer regions with GWAS catalog SNPs. We also included SNPs that are outside the enhancer regions but in complete linkage disequilibrium (LD) with the SNPs within the enhancer regions. We used Plink, version 1.0.9 (Chang *et al.* 2015) to obtain a list of SNPs that are in complete LD ($r^2 = 1.0$) with those that lie within the enhancers of interest ('complete LD SNPs') and carried out the same analyses as described above. Detailed description of the Plink analyses can be found in Supplementary File S1. For subsequent analyses, we combined the GWAS hits of the SNPs that reside within the enhancers and those that are in complete LD with SNPs inside the enhancers and considered them together. We also carried out the same analyses for brain and testis enhancers that exhibit nonsignificant deviations from neutral expectations. To determine if there is a statistical difference in the enrichment of variants in the GWAS hits between enhancers that exhibit evidence of recent positive selection and those that do not, we carried out 2x2 chi-squared tests with Yate's continuity correction using R. The full list of all the GWAS hits that overlap with all the variants considered, as well as the chi-squared test results, can be found in Tables S36-37 and S38-39.

Data Availability

All the supplementary tables, figures and files from this study are in the G3 figshare repository. Supplemental material available at FigShare: https://doi.org/10.25387/g3.8230757.

RESULTS

Enhancers experienced positive selection at different time periods during recent human history

Calculation of four metrics of selection (Weir & Cockerham's F_{ST}, Tajima's D, H12 and nS_L) for all 41,561 enhancers in 41 tissues (these data can be found in File S2) revealed that an average 5.90% of enhancers have experienced recent positive selection considering all metrics and tissues (Figure 2). The proportions of enhancers that exhibit significant deviations from neutral expectations for each metric varied (Figure 2; Table S2). Specifically, greater fractions of enhancers show evidence of recent positive selection according to H12 and Tajima's D metrics (H12: from 6.57 to 16.67%; Tajima's D: 6.75–13.92%) than other metrics (F_{ST}: 1.91–5.56%; nS_L in Africans: 1.12–7.08%; nS_L in Europeans: 2.25–5.39%; nS_L in East Asians: 2.25–5.93%). Furthermore, for 33/41 tissues examined, the H12 metric had the highest proportion of enhancers exhibiting evidence of recent selection. In addition, there were no significant differences in lengths between enhancers with significant and non-significant evidence of recent positive selection. The only exceptions were: kidney (nS_L in East Asians; adjusted p-value = 0.042), lung (nS_L in Europeans; adjusted p-value = 0.023), meninx (nS_L in Africans; adjusted p-value = 0.019), parotid-gland (Tajima's D; adjusted p-value = 0.002), skin (nS_L in Africans; adjusted p-value = 0.024), throat (nS_L in Africans; adjusted p-value = 0.0003), thyroid-gland (nS_L in Africans; adjusted p-value = 0.0003), tongue (nS_L in Africans; adjusted p-value = 0.002), tonsil (nS_L in East Asians; adjusted p-value = 0.047), urinary-bladder (nS_L in Africans; adjusted p-value = 0.045), and vagina (nS_L in Africans; adjusted p-value = 0.044) (Tables S3-8). In short, these results imply that human enhancers experienced selection at different time points in human history, with substantial evidence for soft selective sweeps in which multiple haplotypes increase in frequency within a population (Pennings and Hermisson 2006; Messer and Petrov 2013). However, we caution against overly interpreting differences in the proportions of enhancers across metrics as differences in the action of selection across different time periods since they likely vary in power to detect selection.

Enhancers that experienced positive selection 20,000–10,000 years ago putatively regulate the activities of genes with immunity-related functions

To study the functions of the enhancers that show evidence of recent positive selection, we carried out enrichment analysis on the putative target genes associated with them (these results can be found in Files S5-21). We found that there was little functional enrichment among the enhancers identified by F_{ST}, Tajima's D and H12 (Table 1). Analyzing all enhancers irrespective of their tissues produced similar results, with no significant GO terms for the F_{ST} metric and very few GO terms for the Tajima's D and H12 metrics (these results can be found in the

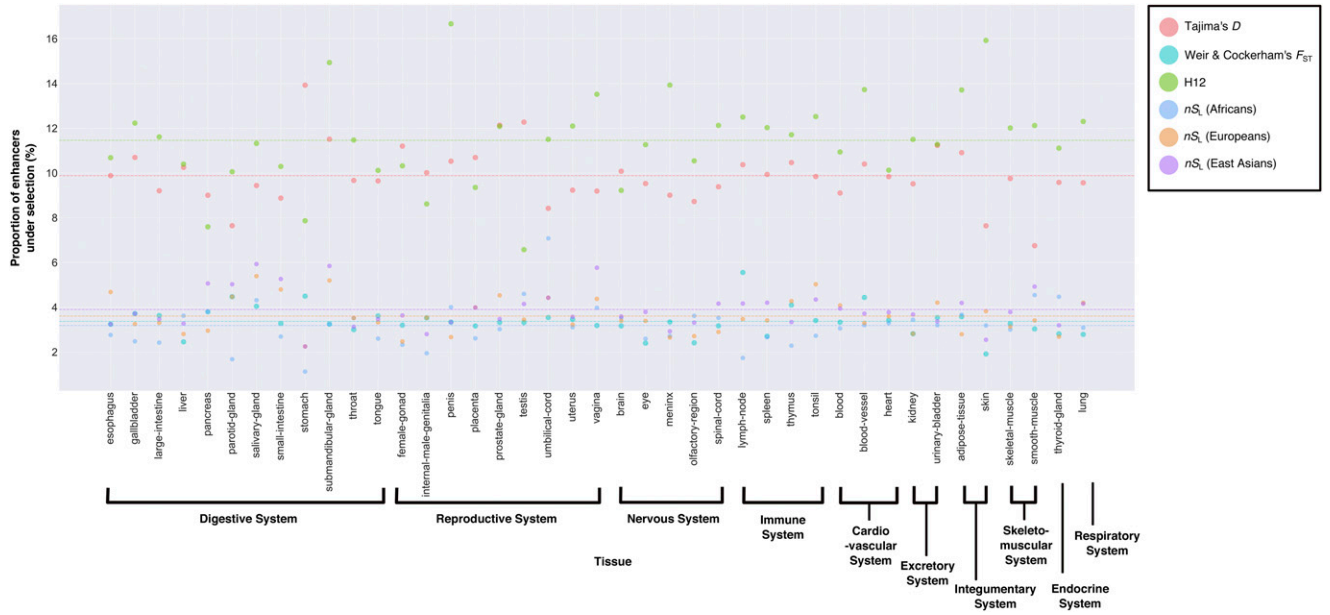

Figure 2 Proportions of enhancers exhibiting significant deviations from neutral expectations for different recent positive selection metrics across 41 tissues Each dot represents the proportion of enhancers (Y-axis) that exhibited significant deviation from the neutral model (*i.e.*, p-value < 0.05) in a given tissue (X-axis). Differently colored dots correspond to the different selection metrics used (Salmon: Tajima's D; Teal: F_{ST}; Green: H12; Blue Gray: nS_L (Africans); Orange: nS_L (Europeans); Purple: nS_L (East Asians)). For each metric, the p-value for the observed value for an enhancer (*i.e.*, the likelihood under neutral expectations of obtaining a value as or more extreme as the observed value) was assessed by comparing to 10,000 simulated values calculated on sequences generated from the neutral simulations. Differently colored horizontal lines correspond to the average of all proportions calculated for all 41 tissues according to the different selection metrics used (Salmon: Tajima's D; Teal: F_{ST}; Green: H12; Blue Gray: nS_L (Africans); Orange: nS_L (Europeans); Purple: nS_L (East Asians)). Note the 1) differences in the proportion of enhancers exhibiting significant deviations from neutral expectations across tissues in any given metric, and 2) similarly, differences in the proportion of enhancers with significant deviations from the neutral model across all metrics in any given tissue.

Tables S9-12). In contrast, many tissues had multiple significant enriched GO terms among the nS_L hits, the majority of which were immunity-related (Tables 1, 2, 3, and 4). Importantly, analyses on the putative target genes of enhancers with no evidence of recent positive selection according to the nS_L metric showed that none of the tissues possessed any significantly enriched GO terms: the only exception was a single Cellular Component term in the parotid-gland for nS_L metric in Europeans (GO:0044444: cytoplasmic part, adjusted p-value = 0.047). These results suggest that approximately 20,000–10,000 years ago, enhancers that putatively regulate the activities of immunity-related genes in multiple tissues underwent selection, a finding consistent with the hypothesis that human populations faced novel local selective pressures as they moved into new environments (Balaresque *et al.* 2007; Fumagalli *et al.* 2011).

To determine if the pattern of enrichment for immunity-related functions among the enhancers with significant evidence for positive selection according to the nS_L metric was shared across tissues, we quantified the semantic similarity (SS) of the functionally enriched terms among tissues (Figures S2a-i). The pairwise SS values among tissues were consistently very high (BP: Africans: 0.687 to 1.000; Europeans: 0.786 to 1.000; MF: Africans: 0.779 to 1.000; Europeans: 0.715 to 1.000; East Asians: 0.902 to 1.000; CC: Africans: 0.873 to 1.000; Europeans: 0.864 to 1.000; East Asians: 0.853 to 1.000): one exception to this general trend were the SS scores calculated on the Biological Process terms in East Asians (from 0.208 to 1.000) (Figure S2c). To determine if the range of SS values we observed were unusually high, we compared each pairwise SS value to the 95th percentile value of the SS values calculated on the 1,000 randomly sampled GO terms for the same pair of tissues (these values, as well as the empirical SS values, can be found in Tables S13-21). We found that most of the empirical SS values were higher than the 95th percentile values of the SS values calculated on the randomly sampled GO terms, with the exception of pairwise comparisons involving liver for the Biological Process terms in East Asians (Figure S2c). In other words, most pairs of tissues were associated with semantically similar GO terms, suggesting that patterns of functional enrichment were very similar across tissues. The sole exception to this

Table 1 Number of tissues with one or more significantly enriched GO Terms

Metric	# of Biological Process GO terms	# of Molecular Function GO terms	# of Cellular Compartment GO terms
Tajima's D	4	2	4
F_{ST}	2	2	2
H12	0	1	1
nS_L (AFR)	18	18	18
nS_L (EUR)	20	20	19
nS_L (EAS)	19	19	19

Table 2 Most frequently-occurring GO terms among enhancers with evidence of recent positive selection by nS_L across tissues in Africans

BP GO Terms	# of tissues	MF Go Terms	# of tissues	CC GO Terms	# of tissues
Antigen processing and presentation of exogenous peptide antigen	18	MHC class II receptor activity	18	*trans*-Golgi network membrane	18
T cell co-stimulation	18	MHC class II protein complex binding	18	Integral component of luminal side of endoplasmic reticulum membrane	18
Interferon-gamma-mediated signaling pathway	17	Peptide antigen binding	15	Endosome membrane	18
T cell receptor signaling pathway	16	TAP1 binding	13	ER to Golgi transport vesicle membrane	17
Antigen processing and presentation of exogenous peptide antigen via MHC class I TAP dependent	15	Peptide antigen-transporting ATPase activity	12	Lysosomal membrane	16
MHC class II protein complex assembly	14	TAP2-binding	10	TAP complex	14
				MHC protein complex	11

were 13 comparisons involving the liver, which had three additional biological terms in East Asians (GO:0071294: cellular response to zinc ion, adjusted p-value = 0.005; GO:0046597: negative regulation of viral entry into host cell, adjusted p-value = 0.005; GO:007126: cellular response to cadmium ion, adjusted p-value = 0.014), which were found only in the liver. Overall, our analyses suggest that the enrichment for immunity-related functions found in putative target genes of enhancers with evidence of recent positive selection by the nS_L metric is not confined to specific tissues, but represents a general trend.

Positively selected enhancers are not enriched for transposable element origins

To examine whether enhancers that exhibit statistically significant signatures of recent positive selection tend to have arisen from TEs, we compared the proportions of TE overlap between enhancers that exhibit evidence of recent positive selection *vs.* enhancers with no evidence of recent positive selection across all 41 tissues (Tables S22-27). Of the 41 tissues examined, only the lung (χ^2 = 12.859, adjusted p-value = 0.002) and the female gonad (χ^2 = 8.632, adjusted p-value = 0.020) displayed statistically significant differences in terms of the proportions of TE-overlapping regions between the two groups of enhancers for the H12 metric (Figures S3a-f). Moreover, in these two tissues, enhancers with evidence of recent positive selection showed lower proportions of TE overlap than enhancers with no such evidence. Overall, these results suggest that enhancers that have arisen from TEs were not preferential targets of recent positive selection.

Enhancers active in testis and brain show different patterns of recent evolution

To test the hypothesis that signatures of selection during recent human history differed across enhancers active in different tissues, we compared the distributions of the four metrics across all 820 pairs of the 41 tissues (these results can be found in Tables S28-33). The number of tissue pairs exhibiting different distributions for Tajima's D was 22 / 820, for F_{ST} was 5 / 820, for H12 was 56 / 820, and for nS_L was 2 / 820 in Africans, and 0 / 820 in Europeans and East Asians (Figures 3a-f). All tissue pairs exhibiting significantly different distributions involved either brain or testis, with the exception of two nS_L comparisons (Africans: Adipose-tissue and Large-intestine: p-value = 0.014; Adipose-tissue and Thymus: p-value = 0.007). Specifically, brain enhancers exhibited significantly different distributions for Tajima's D, F_{ST}, and H12 compared to enhancers from 5, 2, and 23 other tissues (Figures 3a-c); similarly, testis enhancers exhibited significantly different distributions for Tajima's D, F_{ST} and H12 compared to enhancers from fifteen, three, and thirty other tissues (Figures 3a-c). We note that the high fractions of enhancers under selection in brain and testis are unlikely to be solely due to the high number of enhancers found in these tissues; specifically, there are several other tissues that have high number of enhancers (Table S1) and yet fail to exhibit statistically significant pairwise comparisons with other tissues.

In general, the significant differences in the distributions of these metrics between testis or brain and other tissues were manifested as differences in the magnitude of the peaks of the distributions and/or differences in shapes of the distributions rather than from shifts in the range of the distributions. For all metrics, most pairs of tissues with significant differences had very subtle shifts in the distributions, and in cases where the shifts were more noticeable, brain and testis enhancers' distributions tended to be shifted to the left (for F_{ST}; Figure S5) or the right (for Tajima's D; Figure S4), both in the direction suggestive of lower levels of positive selection for the brain or testis enhancers. The nature of the differences in the peaks of the magnitude varied for each metric: for F_{ST}, the peaks for both brain and testis were consistently higher compared to those of the other tissues. In contrast, for Tajima's D, the peaks for brain and testis were consistently lower than the other tissues being compared; the sole exception was the comparison between brain and meninx and urinary-bladder, in which the peaks were almost equal or slightly higher for the brain, respectively. In addition, for Tajima's D and H12, there were differences in the shapes of their distributions (Figures S4; S6-7). These results imply that enhancers active in the brain and testis experienced different selective pressures during recent human history compared to enhancers active in other tissues.

To further determine if brain and testis stood out compared to other tissues in terms of the patterns of recent evolution of their enhancers, we also ranked the tissues according to the median values of each metric (Figure S8) and the proportion of enhancers that exhibit statistically

Table 3 Most frequently-occurring GO terms among enhancers with evidence of recent positive selection by nS_L across tissues in Europeans

BP GO Terms	# of tissues	MF GO Terms	# of tissues	CC GO Terms	# of tissues
Interferon gamma-mediated signaling pathway	18	MHC class II receptor activity	18	Integral component of luminal side of endoplasmic reticulum membrane	19
Antigen processing and presentation of exogenous peptide antigen	17	MHC class II protein complex binding	18	ER to Golgi transport vesicle membrane	18
T cell co-stimulation	17	Peptide antigen binding	15	*trans*-Golgi network membrane	18
T cell receptor signaling pathway	14	TAP1 binding	14	Endosome membrane	17
MHC class II protein complex assembly	13			Lysosomal membrane	15
				Endocytic vesicle membrane	12
				TAP complex	12
				MHC protein complex	12

significant evidence of recent positive selection for each metric (Figure S9). Overall, we did not find any general patterns regarding the recent evolution of brain and testis, in terms of either the median values of the metrics of recent selection (Figure S8) or the proportion of enhancers that have putatively been under selection (Figure S9). Exceptions to the general trend were that testis and brain ranked 1st and 2nd, respectively, for median values of the H12 metric (Figure S8c) and that testis ranked 2nd in the proportion of enhancers with evidence of recent selection according to Tajima's D and nS_L in Africans (Figures S9a & d).

Evolution of tissue-specific enhancers differs from that of tissue-broad enhancers in the brain and testis

We next compared the patterns of recent evolution between tissue-specific and tissue-broad enhancers for each tissue. We found that the number of tissues with statistically significant differences in the distributions of the metrics between enhancers active in a single tissue and those active in multiple tissues was as follows, sorted from highest to lowest: 20/41 for H12, 8/41 for Tajima's D, 8/41 for F_{ST} and 1/41 for nS_L in Africans and East Asians (Figures 4a & S10-15). Tissue-specific and tissue-broad enhancers in the brain and testis show statistically significant differences in the patterns of recent evolution for all metrics except nS_L (Tajima's D: brain: 1.882e-07; testis: 9.523e-11; F_{ST}: brain: 6.074e-07; testis: 1.639e-07; H12: brain: $<$ 2.2e-16; testis: $<$ 2.2e-16 (Figures 4b-d). Regarding the nature of the differences between the two groups of enhancers in these tissues, we found contrasting patterns for H12 compared to Tajima's D and F_{ST}: for Tajima's D and F_{ST}, we found that the interquartile values of the tissue-specific enhancers were lower (or higher for Tajima's D) than those of the tissue-broad enhancers in brain and testis, suggestive of less positive selection in tissue-specific enhancers in these two tissues. In contrast, values for the H12 metric were consistently higher for the tissue-specific enhancers compared to the tissue-broad enhancers in brain and testis. These results suggest that, in general, enhancers with different breadth of tissue activity did not differ in their patterns of recent evolution. Nevertheless, enhancers active only in the brain do show significant differences compared to those active in multiple tissues (including brain) and the same is true for testis, raising the possibility that the recent evolution of brain- and testis-specific enhancers may have been different from that of other tissue-specific enhancers.

Table 4 Most frequently-occurring GO terms among enhancers with evidence of recent positive selection by nS_L across tissues in East Asians

BP GO Terms	# of tissues	MF GO Terms	# of tissues	CC GO Terms	# of tissues
Antigen processing and presentation of exogenous peptide antigen	17	Peptide antigen binding	19	Integral component of luminal side of endoplasmic reticulum membrane	19
T cell co-stimulation	17	MHC class II receptor activity	18	*trans*-Golgi network membrane	18
interferon-gamma-mediated signaling pathway	16	MHC class II protein complex binding	18	Lysosomal membrane	18
T cell receptor signaling pathway	15	TAP1 binding	12	ER to Golgi transport vesicle membrane	17
MHC class II protein complex assembly	14			Endocytic vesicle membrane	16
Antigen processing and presentation of exogenous peptide antigen via MHC class I TAP dependent	11			TAP complex	12
				Endosome membrane	10
				MHC II protein complex	10

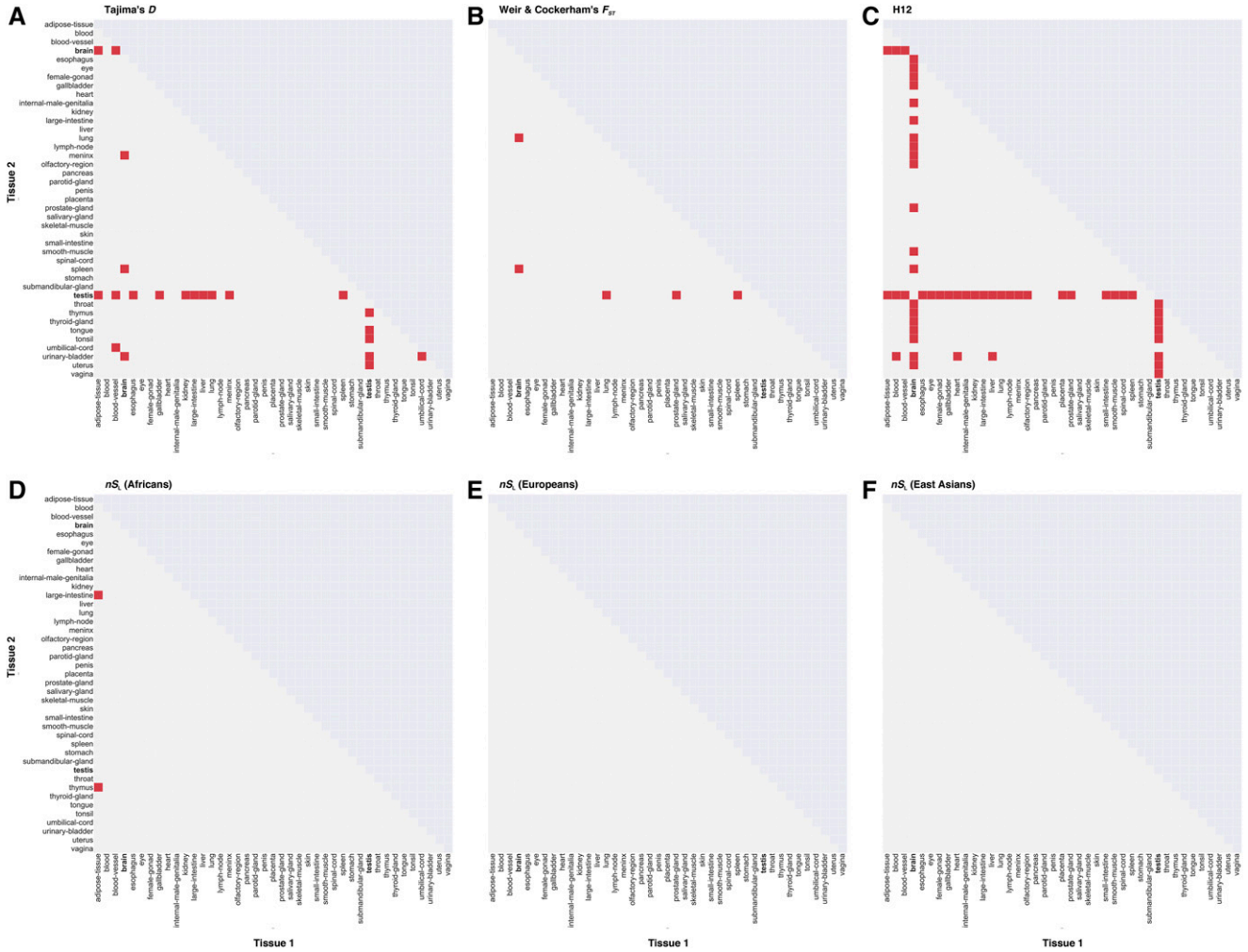

Figure 3 Pairwise comparisons of the patterns of recent evolution among enhancers from 41 tissues Each graph shows the results of the pairwise Kolmogorov-Smirnov tests carried out between all pairs of 41 tissues (shown on the X- and Y-axes) for all recent positive selection metrics. Each cell represents a pairwise comparison between two specific tissues: red-filled cells represent pairwise comparisons that are statistically significant (*i.e.*, adjusted *p*-value < 0.05) and empty cells represent non-significant pairwise comparisons.

The recent evolution of testis enhancers associated with tissue-enriched genes differs from the evolution of enhancers associated with non-tissue-enriched genes

To determine whether there are differences in the patterns of recent evolution between enhancers associated with tissue-enriched genes *vs.* those associated with non-tissue-enriched genes, we compared the distributions of the four metrics between brain and testis enhancers stratified by the breadth of expression of their target genes (these results can be found in Tables S40-41). The only significant differences that we found were for H12 (adjusted *p*-value =3.667e-06) and Tajima's *D* (adjusted *p*-value = 0.004) in the testis (Figure 5). More specifically, we found that the distribution of Tajima's *D* values of the enhancers associated with testis-enriched genes were shifted to the left (*i.e.*, more negative values suggestive of stronger positive selection) compared to the enhancers associated with non-testis-enriched genes (Figure 5a). The difference between the same groups of enhancers for the H12 metric was comparatively subtle, with the main difference being that the peak of the distribution of the enhancers associated with the non-testis-enriched genes was higher than that of the enhancers associated with testis-enriched genes (Figure 5b). These results imply that enhancers associated with genes with enriched expression in the testis have experienced different degrees of selection in the form of both hard and soft selective sweeps compared to enhancers associated with genes that do not show enriched expression in the same tissue.

Brain and testis enhancers under recent positive selection are not significantly enriched for variants associated with complex human traits and diseases

To examine if genetic variants in brain and testis enhancers showing signatures of recent selection are associated with particular complex human traits and diseases, we queried the NHGRI-EBI GWAS catalog (MacArthur *et al.* 2017). We found that for all metrics, most or all enhancers did not harbor SNPs that are associated with any human traits (brain: Tajima's *D*: 372/421; F_{ST}: 131/150; H12: 430/437; nS_L(Africans): 150/167; nS_L (Europeans): 144/161; nS_L (East Asians): 152/170; testis: Tajima's *D*: 143/163; F_{ST}: 47/52; H12: 102/103; nS_L (Africans): 64/72; nS_L (Europeans): 48/54; nS_L (East Asians): 60/65). In the few instances in which SNPs within enhancers showing evidence of selection were associated with human traits, these associations were with traits such as Alzheimer's disease (brain: Tajima's *D*), obesity-related traits (brain: Tajima's *D*), and type 2 diabetes (brain: nS_L (Africans)) (Tables 5 & 6; the full list of overlapping GWAS hits,

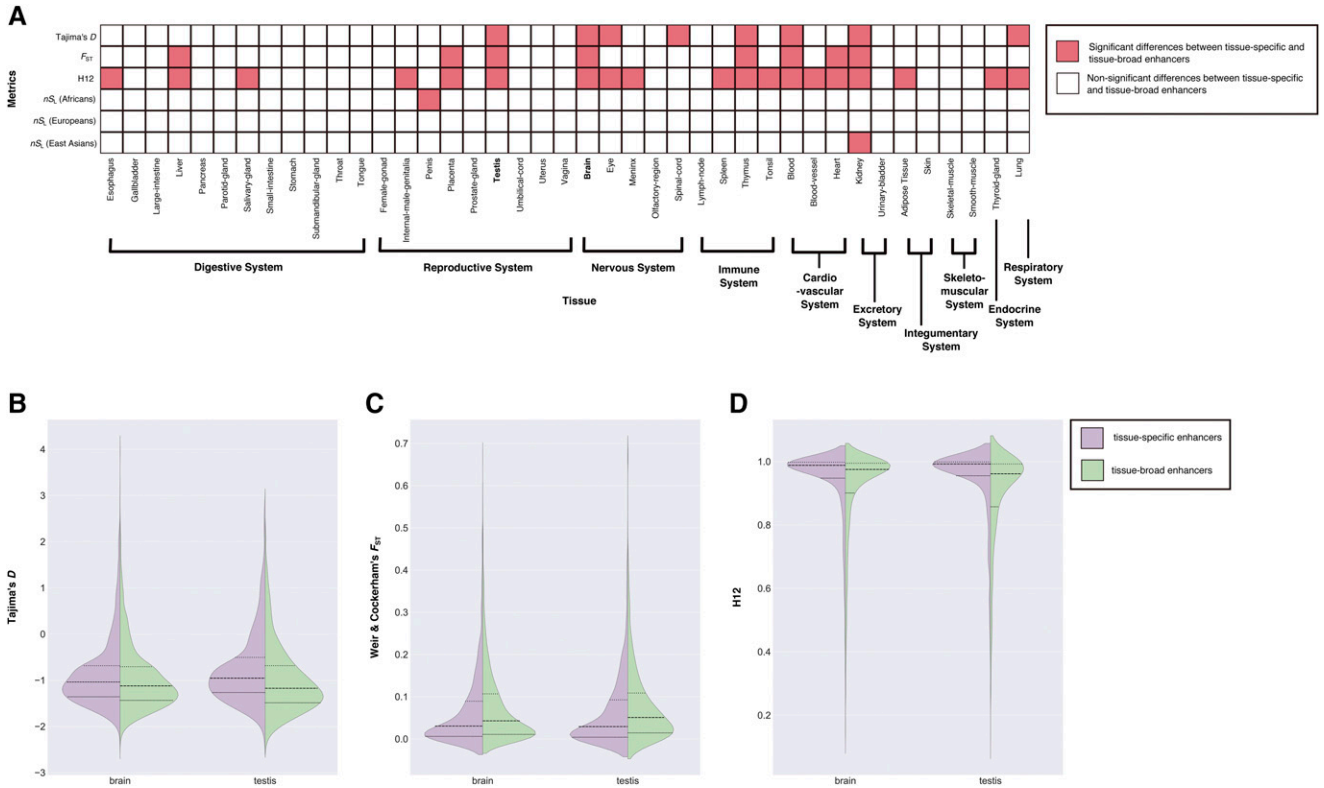

Figure 4 Comparisons of patterns of recent evolution between tissue-specific and tissue-broad enhancers (a) The grid panel depicts which tissues exhibit significantly different distributions of metrics between tissue-specific and tissue-broad enhancers. Each cell represents a comparison between tissue-specific and tissue-broad enhancers in a given tissue: pink-filled cells represent the comparisons that are statistically significant (*i.e.*, adjusted *p*-value < 0.05) and empty cells represent non-significant comparisons. (b–d) Violin plots depicting the distributions of the metrics for tissue-specific and tissue-broad enhancers for brain and testis for (b) Tajima's *D*, (c) Weir & Cockerham's F_{ST}, and (d) H12. In all cases, the distributions of tissue-specific and tissue-broad enhancers were significantly different (see also panel a).

including variants in LD with those inside enhancers, can be found in the Tables S42-43). We next compared the proportions of enhancers that overlapped with GWAS hits between enhancers that exhibit evidence of recent positive selection and enhancers that do not and did not find any significant difference between them. More specifically, we saw no statistically significant differences in the proportions of overlap with GWAS hits between enhancers with or without evidence of recent positive selection: the only exception was the H12 metric in the brain (χ^2= 30.552, adjusted *p*-value = 1.950e-07), which showed depletion of GWAS hits among enhancers with evidence of recent positive selection (These results can be found in Tables S44-45). These results show that overall, enhancer SNPs under recent positive selection are not preferentially associated with specific human traits or complex human diseases. However, we note that it is also possible that such lack of enrichment for GWAS hits among enhancers with significant evidence of recent selection is due to the fact that such enhancers are less likely to harbor potentially deleterious SNPs that have reached high frequencies within human populations.

DISCUSSION

In this study, we examined signatures of recent positive selection in 41,561 enhancers active in 41 human tissues. We found that approximately 5.90% of enhancers across tissues and metrics exhibit significant evidence of recent positive selection, b) the putative target genes of enhancers that experienced selection in the last 20,000 – 10,000 years ago are enriched for immunity-related functions, and c) enhancers active in the brain and testis exhibited significantly different patterns of recent evolution compared to enhancers in other tissues.

Our first key result is that the proportions of enhancers in each tissue that exhibited signatures of selection vary across the different metrics (Figure 2). Since the origin of modern humans approximately 250,000 - 200,000 years ago, numerous episodes, including the out-of-Africa migration approximately 75,000 - 50,000 years ago, the Eurasian split that occurred approximately 45,000 - 36,000 years ago, and the Agricultural Revolution approximately 20,000 - 10,000 years ago, exposed the human populations to novel selective pressures (*e.g.*, pathogens, dietary changes) (Sabeti *et al.* 2006; Karlsson *et al.* 2014). Adaptations in response to such changes would have likely altered the allele frequencies in both the genic and regulatory regions. As the different metrics of selection that we employed in our study are sensitive to distinct types of genomic signatures (Sabeti *et al.* 2006; Voight *et al.* 2006), one possible interpretation is that within any given tissue, the differences in percentages of enhancers under selection between metrics reflect temporal variation in the action of selection during human history.

There are two caveats to interpreting the results of Figure 2 to represent temporal variation in the action of selection. The first caveat is that these metrics likely have different power to detect selection. For instance, Tajima's *D*, a metric that detects excess of rare alleles in a region of interest, is most sensitive to selective sweeps that have resulted in complete fixation of the target locus but has limited power otherwise (Sabeti *et al.* 2006; Ferrer-Admetlla *et al.* 2014). In contrast, nS_L, a

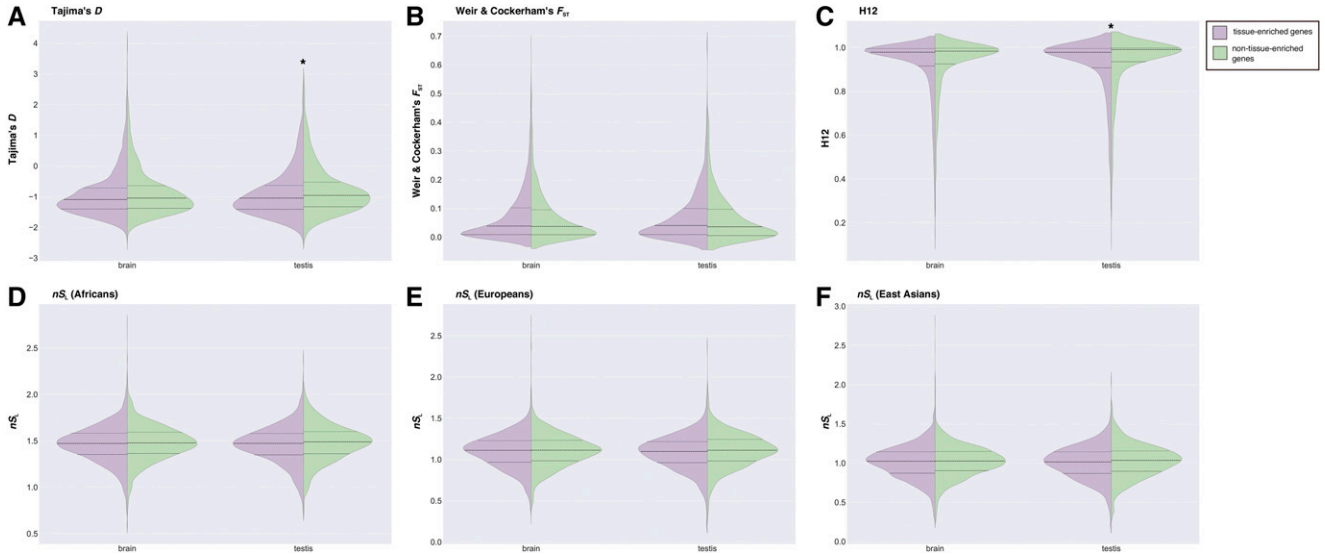

Figure 5 Comparisons of patterns of recent evolution between enhancers associated with tissue-enriched genes and enhancers associated with non-tissue-enriched genes in brain and testis The violin plots depicting the distributions of the metrics for enhancers associated with tissue-enriched and non-tissue-enriched genes for brain and testis for: (a) Tajima's *D* (b) Weir & Cockerham's F_{ST} (c) H12 (d) nS_L (Africans) (e) nS_L (Europeans) (f) nS_L (East Asians). Plots with asterisks above them indicate significant pairwise comparisons.

haplotype-based metric, best detects ongoing or incomplete hard selective sweeps resulting in intermediate allele frequencies (*i.e.*, allele frequencies = 60–80%) and rapidly loses power as allele frequencies increase to 100% (Ferrer-Admetlla *et al.* 2014). Thus, the observed differences in the proportions of enhancers with evidence of recent positive selection across the metrics also reflect the disparity in power of the metrics to successfully detect incidences of selection that happened at a particular time period and cannot be directly compared to each other. The second caveat is that our inferences rely on an explicit neutral demographic model whose key parameter values are fixed rather than drawn from distributions (due to the substantially increased computational burden associated with the latter choice). Given the very large number of enhancers examined, it is possible that not accounting for neutral variation may have increased our numbers of false positives in a metric-dependent fashion (Harris *et al.* 2018). Nevertheless, the estimates of demographic parameters we used are consistent with those derived from other data sources (Gravel *et al.* 2011), and the inference of selection using large genomic segments by some of our metrics (*e.g.*, nS_L) makes them less susceptible to this potential type of error. Distinguishing whether our results are best explained by temporal variation in selection or by our experimental design choices and differences in the power of our metrics could be achieved via more realistic simulations in which selective events are introduced at specific time points and result in a fixed proportion of regions being selected. This is an important future research direction that has potential to shed light into the tempo of selection in the course of recent human evolution.

Our second key result is that the putative target genes of enhancers exhibiting significant evidence of recent positive selection according to the nS_L metric are enriched for immunity-related functions (Tables 2-4). Haplotype-based metrics such as nS_L are known to be sensitive to signatures resulting from selection that occurred approximately 20,000 - 10,000 years ago, which corresponds to the incidence of the Agricultural Revolution (Voight *et al.* 2006). The advent of farming practices, communal living in settlements, and migrations of farmers across the globe resulted in an increase of numbers and densities of humans in any given location, likely facilitating the spread of pathogens (Varki and Gagneux 2009; Page *et al.* 2016; Nielsen *et al.* 2017). In addition, recent studies suggest that selection on *cis*-regulatory regions, such as enhancers, might have been important in driving adaptation of modern human populations to distinct environments, due to their modular organization: change of expression pattern in one temporal or spatial context can often occur without affecting others, which could contribute to phenotypic changes without incurring negative pleiotropic effects (Carroll 2005; Wray 2007). Therefore, it is possible that around 20,000 - 10,000 years ago, enhancers regulating the activities of immunity-related genes underwent selection to allow refined fine-tuning of host defense processes in response to the stronger pressure from pathogens resulting from increased human population sizes.

The third and perhaps most striking result of our study is that brain and testis enhancers exhibited different patterns of recent evolution, typically in the direction of lower levels of positive selection, compared to enhancers of other tissues (Figure 3 & S4-7). For both brain and testis, the high number of enhancers included in these tissues could partly explain our results. Brain, with 4,883 enhancers, has the most enhancers of all the tissues examined, while testis, with 1,621 enhancers, is ranked 7th. The effect of the large number of enhancers included in these tissues is reflected in the differences in the magnitudes of peaks of the distributions of the metrics between brain and testis and other tissues (Figure S4-7). In addition, significant differences in the lengths of enhancers included in the brain and testis compared with other tissues could further contribute to our results (Table S46). Overall, enhancers in testis tend to be shorter than those in other tissues (*i.e.*, distribution of lengths of testis enhancers exhibits subtle shifts to the left compared to other tissues; Figure S16). The same is true for brain enhancers, except when compared with testis enhancers (*i.e.*, the distribution of lengths of enhancers active in the brain are shifted to the right compared to the distribution of lengths of testis enhancers); however, for comparisons with some tissues (*e.g.*, heart, internal male genitalia, small-intestine, blood, and spinal-cord), the shift was more subtle, with the main difference instead occurring in the peaks of the magnitudes or the shapes of the distributions (Figure S17). Nevertheless,

Table 5 Complex human traits and diseases associated with variants within the brain enhancers that exhibit evidence of recent selection

Enhancer ID	rsID	Traits	Metrics
chr2:219271996-219272374	rs921968	Mean corpuscular hemoglobin concentration	F_{ST}
chr3:181418145-181418802	rs34308817	Ankle injury	Tajima's *D*
chr6:30069810-30070038	rs1111180	Eosinophil percentage of granulocytes, Eosinophil percentage of white cells	*n*SL (EUR)
chr6:30923441-30923743	rs17189763	Conotruncal heart defects (maternal effects)	nS_L (EUR)
chr6:32427743-32428120	rs9268831	Response to hepatitis B vaccine	nS_L (AFR, EUR, EAS)
chr6:32427743-32428120	rs9268835	Type 2 diabetes	nS_L (AFR, EUR, EAS)
chr6:32577297-32577935	rs660895	IgA nephropathy, Rheumatoid arthritis, Rheumatoid arthritis (ACPA-negative)	nS_L (AFR, EUR, EAS)
chr7:130720133-130720826	rs10265693	Lung cancer	F_{ST}
chr7:29217860-29218383	rs245914	Psychosis (atypical), Obesity-related traits	Tajima's *D*
chr13:110790027-110791241	rs641862	Obesity-related traits	Tajima's *D*
chr16:22201170-22202123	rs145049847	Alzheimer disease and age of onset	Tajima's *D*
chr16:87856343-87856555	rs76069656	Triglyceride change in response to fenofibrate in statin-treated type 2 diabetes	nS_L (EUR)
chr16:87886258-87886670	rs68149176	Mean corpuscular volume, Mean corpuscular hemoglobin	nS_L (EUR), F_{ST}
chr21:45616099-45616530	rs4456788	Chronic inflammatory diseases (ankylosing spondylitis, Crohn's disease, psoriasis, primary sclerosing cholangitis, ulcerative colitis) (pleiotropy)	nS_L (EAS)

we also observed other types of differences in comparisons of the distributions of these metrics between brain and testis and other tissues: while subtle, some metrics (Tajima's *D*, F_{ST}; Figure S4-5) showed shifts of the distributions of brain and testis enhancers toward weaker signatures of recent positive selection compared to other tissues, suggesting that the significant pairwise differences we see are not solely caused by disparities in the numbers of enhancers or lengths of enhancers within tissues. In addition, there are several other tissues (*e.g.*, lung, spleen, blood) whose numbers of enhancers are nearly on par with those of the brain and more than the testis or whose lengths significantly differed compared to numerous other tissues. However, these tissues did not show significant differences in the distribution of these metrics compared to other tissues. Furthermore, most tissues did not differ significantly in terms of the recombination rates of the enhancers: the only exception was between eye and testis (adjusted *p*-value = 0.05; Table S47). Collectively, these results further suggest that the observed differences in selection are likely biological.

Why do signatures of recent positive selection tend to be lower for brain and testis enhancers? Answering these questions is challenging without additional functional experiments that shed light on the phenotypic effects of the selected variants. In the case of the brain, one possible explanation could be that brain enhancers are under stronger functional constraint (and hence less likely to exhibit positive selection) than enhancers in other tissues. Consistent with this hypothesis, existing evidence suggests that the brain size of modern humans has not changed substantially in the last 250,000 - 300,000 years (Neubauer *et al.* 2018). Similarly, expression patterns of genes expressed in neural tissues (*e.g.*, brain, cerebellum) show low levels of divergence across both species and within humans, suggesting that the overall structure of the neural network is highly conserved (Khaitovich 2005; Brawand *et al.* 2011). Finally, the brain transcriptome exhibits low divergence due to the relatively strong developmental constraints acting on neural tissues (Gu and Su 2007). For instance, a study found that the correlation between expression levels and signals of purifying selection is strongest for brain, suggesting that genes expressed in the brain evolve slowly likely due to stronger functional constraints (Kryuchkova-Mostacci and Robinson-Rechavi 2015).

In the case of testis, our result that enhancers active in the testis often show weaker signatures of positive selection compared to enhancers in other tissues is surprising and harder to reconcile with what is known about how this tissue's regulatory landscape evolved. For example, previous studies have shown that testis exhibits the highest degree of expression pattern divergence between species, likely due to either strong action of positive selection on reproductive processes (Khaitovich 2005; Brawand *et al.* 2011), relaxation of purifying selection in the testis compared to other tissues (Gershoni and Pietrokovski 2014), or both. Unfortunately, our results are not directly comparable with studies measuring gene expression divergence between species due to the difference in the features being compared and the temporal window examined; previous studies have looked at the divergence of the overall transcriptome, which is influenced not just by changes in enhancer sequences but by other types of changes and other regulatory elements as well, whereas we have specifically looked at changes in allele frequency and haplotypes in enhancer regions. In addition, previous

Table 6 Complex human traits and diseases associated with variants within the testis enhancers that exhibit evidence of recent selection

Enhancer ID	rsID	Traits	Metrics
Chr6:30069810-30070038	rs1111180	Eosinophil percentage of granulocytes, Eosinophil percentage of white cells	nS_L (Europeans)
Chr6:30923441-30923743	rs17189763	Conotruncal heart defects (maternal effects)	nS_L (Europeans)

studies have looked at divergence of the expression patterns between species over an evolutionary period spanning several millions of years, whereas we have examined selection events that have occurred in more recent evolutionary periods, on the order of hundreds or tens of thousands of years. Therefore, it is possible that there has been temporal variation (*i.e.*, ancient *vs.* recent) in the occurrence of selection events on these enhancer regions.

More broadly, it is worth emphasizing that most of the 41,561 enhancers examined do not deviate from neutral expectations. Studies examining the divergence of brain and liver transcriptomes across species have previously suggested that tissue transcriptome differences accumulate approximately linearly with evolutionary time (Khaitovich *et al.* 2004), consistent with a broadly neutral model of transcriptome evolution. Quantifying the relative contributions of different evolutionary forces across such a large window of evolutionary time, between tissues with vastly different physiologies, and for a group of genetic elements whose evolution is only now beginning to be understood is, our study illustrates, non-trivial. Further dissection of how adaptive and non-adaptive processes shaped enhancers during the evolution of modern humans will greatly advance our understanding of the molecular evolutionary changes that gave rise to human phenotypic diversity.

ACKNOWLEDGMENTS

This research was supported by the March of Dimes through the March of Dimes Prematurity Research Center Ohio Collaborative, the Burroughs Wellcome Fund, the Guggenheim Foundation, and the National Institutes of Health (R01GM115836, R35GM127087). This work was conducted, in part, using the Advanced Computing Center for Research and Education at Vanderbilt University.

LITERATURE CITED

Alexa, A., and J. Rahnenfuhrer, 2016 Gene set enrichment analysis with topGO, R package version 2.30.1

Andersson, R., C. Gebhard, I. Miguel-Escalada, I. Hoof, J. Bornholdt *et al.*, 2014 An atlas of active enhancers across human cell types and tissues. Nature 507: 455–461. https://doi.org/10.1038/nature12787

Balaresque, P. L., S. J. Ballereau, and M. A. Jobling, 2007 Challenges in human genetic diversity: demographic history and adaptation. Hum. Mol. Genet. 16: R134–R139. https://doi.org/10.1093/hmg/ddm242

Brawand, D., M. Soumillon, A. Necsulea, P. Julien, G. Csárdi *et al.* 2011 The evolution of gene expression levels in mammalian organs. Nature 478: 343–348. https://doi.org/10.1038/nature10532

Capra, J. A., G. D. Erwin, G. McKinsey, J. L. R. Rubenstein, and K. S. Pollard, 2013 Many human accelerated regions are developmental enhancers. Phil. Trans. Roy. Soc. B: Biol. Sci. 368: 20130025. https://doi.org/10.1098/rstb.2013.0025

Carroll, S. B., 2005 Evolution at two levels: On genes and form. PLoS Biol. 3: e245. https://doi.org/10.1371/journal.pbio.0030245

Casper, J., A. S. Zweig, C. Villarreal, C. Tyner, M. L. Speir *et al.*, 2018 The UCSC Genome Browser database: 2018 update. Nucleic Acids Res. 46: D762–D769. https://doi.org/10.1093/nar/gkx1020

Chang, C. C., C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell *et al.*, 2015 Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaSci 4: 7. https://doi.org/10.1186/s13742-015-0047-8

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks *et al.*, 2011 The variant call format and VCFtools. Bioinformatics 27: 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Ferrer-Admetlla, A., M. Liang, T. Korneliussen, and R. Nielsen, 2014 On Detecting Incomplete Soft or Hard Selective Sweeps Using Haplotype Structure. Mol. Biol. Evol. 31: 1275–1291. https://doi.org/10.1093/molbev/msu077

Fumagalli, M., M. Sironi, U. Pozzoli, A. Ferrer-Admettla, L. Pattini *et al.*, 2011 Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. PLoS Genet. 7: e1002355 (erratum: PLoS Genet. 7: 10.1371/annotation/ca428083-dbcb-476a-956c-d7bb6e317cf7). https://doi.org/10.1371/journal.pgen.1002355

Garud, N. R., P. W. Messer, E. O. Buzbas, and D. A. Petrov, 2015 Recent selective sweeps in North American *Drosophila melanogaster* show signatures of soft sweeps. PLoS Genet. 11: e1005004. https://doi.org/10.1371/journal.pgen.1005004

Gershoni, M., and S. Pietrokovski, 2014 Reduced selection and accumulation of deleterious mutations in genes exclusively expressed in men. Nat. Commun. 5: 4438. https://doi.org/10.1038/ncomms5438

Gravel, S., B. M. Henn, R. N. Gutenkunst, A. R. Indap, G. T. Marth *et al.*, 2011 Demographic history and rare allele sharing among human populations. Proc. Natl. Acad. Sci. USA 108: 11983–11988. https://doi.org/10.1073/pnas.1019276108

Gu, X., and Z. Su, 2007 Tissue-driven hypothesis of genomic evolution and sequence-expression correlations. Proc. Natl. Acad. Sci. USA 104: 2779–2784. https://doi.org/10.1073/pnas.0610797104

Haller, B. C., and P. W. Messer, 2017 SLiM 2: Flexible, Interactive Forward Genetic Simulations. Mol. Biol. Evol. 34: 230–240. https://doi.org/10.1093/molbev/msw211

Harris, R. B., A. Sackman, and J. D. Jensen, 2018 On the unfounded enthusiasm for soft selective sweeps II: Examining recent evidence from humans, flies, and viruses. PLoS Genet. 14: e1007859. https://doi.org/10.1371/journal.pgen.1007859

Haygood, R., O. Fedrigo, B. Hanson, K.-D. Yokoyama, and G. A. Wray, 2007 Promoter regions of many neural- and nutrition-related genes have experienced positive selection during human evolution. Nat. Genet. 39: 1140–1144. https://doi.org/10.1038/ng2104

International HapMap Consortium, Frazer, K. A., D. G. Ballinger, D. R. Cox, D. A. Hinds, L. L. Stuve *et al.*, 2007 A second generation human haplotype map of over 3.1 million SNPs. Nature 449: 851–861. https://doi.org/10.1038/nature06258

Karlsson, E. K., D. P. Kwiatkowski, and P. C. Sabeti, 2014 Natural selection and infectious disease in human populations. Nat. Rev. Genet. 15: 379–393. https://doi.org/10.1038/nrg3734

Karnuta, J. M., and P. C. Scacheri, 2018 Enhancers: bridging the gap between gene control and human disease. Hum. Mol. Genet. 27: R219–R227. https://doi.org/10.1093/hmg/ddy167

Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle *et al.*, 2002 The human genome browser at UCSC. Genome Res. 12: 996–1006. https://doi.org/10.1101/gr.229102

Khaitovich, P., G. Weiss, M. Lachmann, I. Hellmann, W. Enard *et al.*, 2004 A neutral model of transcriptome evolution. PLoS Biol. 2: e132. https://doi.org/10.1371/journal.pbio.0020132

Khaitovich, P., 2005 Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. Science 309: 1850–1854. https://doi.org/10.1126/science.1108296

King, M. C., and A. C. Wilson, 1975 Evolution at two levels in humans and chimpanzees. Science 188: 107–116. https://doi.org/10.1126/science.1090005

Kryuchkova-Mostacci, N., and M. Robinson-Rechavi, 2015 Tissue-specific evolution of protein coding genes in human and mouse. PLoS One 10: e0131673. https://doi.org/10.1371/journal.pone.0131673

Kudaravalli, S., J. B. Veyrieras, B. E. Stranger, E. T. Dermitzakis, and J. K. Pritchard, 2008 Gene expression levels are a target of recent natural selection in the human genome. Mol. Biol. Evol. 26: 649–658. https://doi.org/10.1093/molbev/msn289

Lee, T. I., and R. A. Young, 2013 Transcriptional regulation and its misregulation in disease. Cell 152: 1237–1251. https://doi.org/10.1016/j.cell.2013.02.014

Long, H. K., S. L. Prescott, and J. Wysocka, 2016 Ever-changing landscapes: transcriptional enhancers in development and evolution. Cell 167: 1170–1187. https://doi.org/10.1016/j.cell.2016.09.018

Lynch, V. J., M. C. Nnamani, A. Kapusta, K. Brayer, S. L. Plaza *et al.*, 2015 Ancient transposable elements transformed the uterine regulatory landscape and transcriptome during the evolution of mammalian pregnancy. Cell Reports 10: 551–561. https://doi.org/10.1016/j.celrep.2014.12.052

MacArthur, J., E. Bowler, M. Cerezo, L. Gil, P. Hall *et al.*, 2017 The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. 45: D896–D901. https://doi.org/10.1093/nar/gkw1133

Messer, P. W., 2013 SLiM: Simulating evolution with selection and linkage. Genetics 194: 1037–1039. https://doi.org/10.1534/genetics.113.152181

Messer, P. W., and D. A. Petrov, 2013 Population genomics of rapid adaptation by soft selective sweeps. Trends Ecol. Evol. 28: 659–669. https://doi.org/10.1016/j.tree.2013.08.003

Moon, J. M., D. M. Aronoff, J. A. Capra, P. Abbot, and A. Rokas, 2018 Examination of signatures of recent positive selection on genes involved in human sialic acid biology. G3 (Bethesda) 8: 1315–1325. https://doi.org/10.1534/g3.118.200035

Neph, S., M. S. Kuehn, A. P. Reynolds, E. Haugen, R. E. Thurman *et al.*, 2012 BEDOPS: high-performance genomic feature operations. Bioinformatics 28: 1919–1920. https://doi.org/10.1093/bioinformatics/bts277

Neubauer, S., J.-J. Hublin, and P. Gunz, 2018 The evolution of modern human brain shape. Sci. Adv. 4: eaao5961. https://doi.org/10.1126/sciadv.aao5961

Nédélec, Y., J. Sanz, G. Baharian, Z. A. Szpiech, A. Pacis *et al.*, 2016 Genetic ancestry and natural selection drive population differences in immune responses to pathogens. Cell 167: 657–669.e21. https://doi.org/10.1016/j.cell.2016.09.025

Nielsen, R., J. M. Akey, M. Jakobsson, J. K. Pritchard, S. Tishkoff *et al.*, 2017 Tracing the peopling of the world through genomics. Nature 541: 302–310. https://doi.org/10.1038/nature21347

Ong, C.-T., and V. G. Corces, 2011 Enhancer function: new insights into the regulation of tissue-specific gene expression. Nat. Rev. Genet. 12: 283–293. https://doi.org/10.1038/nrg2957

Page, A. E., S. Viguier, M. Dyble, D. Smith, N. Chaudhary *et al.*, 2016 Reproductive trade-offs in extant hunter-gatherers suggest adaptive mechanism for the Neolithic expansion. Proc. Natl. Acad. Sci. USA 113: 4694–4699. https://doi.org/10.1073/pnas.1524031113

Pennacchio, L. A., W. Bickmore, A. Dean, M. A. Nobrega, and G. Bejerano, 2013 Enhancers: five essential questions. Nat. Rev. Genet. 14: 288–295. https://doi.org/10.1038/nrg3458

Pennings, P. S., and J. Hermisson, 2006 Soft sweeps II—molecular population genetics of adaptation from recurrent mutation or migration. Mol. Biol. Evol. 23: 1076–1084. https://doi.org/10.1093/molbev/msj117

Pfeifer, B., U. Wittelsbürger, S. E. Ramos-Onsins, and M. J. Lercher, 2014 PopGenome: an efficient Swiss army knife for population genomic analyses in R. Mol. Biol. Evol. 31: 1929–1936. https://doi.org/10.1093/molbev/msu136

Raney, B. J., T. R. Dreszer, G. P. Barber, H. Clawson, P. A. Fujita *et al.*, 2014 Track data hubs enable visualization of user-defined genome-wide annotations on the UCSC Genome Browser. Bioinformatics 30: 1003–1005. https://doi.org/10.1093/bioinformatics/btt637

Rebeiz, M., and M. Tsiantis, 2017 Enhancer evolution and the origins of morphological novelty. Curr. Opin. Genet. Dev. 45: 115–123. https://doi.org/10.1016/j.gde.2017.04.006

Rockman, M. V., M. W. Hahn, N. Soranzo, F. Zimprich, D. B. Goldstein *et al.*, 2005 Ancient and recent positive selection transformed opioid cis-regulation in humans. PLoS Biol. 3: e387. https://doi.org/10.1371/journal.pbio.0030387

Rubinstein, M., and F. S. J. de Souza, 2013 Evolution of transcriptional enhancers and animal diversity. Phil. Trans. Roy. Soc. B: Biol. Sci. 368: 20130017. https://doi.org/10.1098/rstb.2013.0017

Sabeti, P. C., S. F. Schaffner, B. Fry, J. Lohmueller, P. Varilly *et al.*, 2006 Positive natural selection in the human lineage. Science 312: 1614–1620. https://doi.org/10.1126/science.1124309

Schaffner, S. F., 2004 The X chromosome in population genetics. Nat. Rev. Genet. 5: 43–51. https://doi.org/10.1038/nrg1247

Sholtis, S. J., and J. P. Noonan, 2010 Gene regulation and the origins of human biological uniqueness. Trends Genet. 26: 110–118. https://doi.org/10.1016/j.tig.2009.12.009

Simonti, C. N., M. Pavličev, and J. A. Capra, 2017 Transposable element exaptation into regulatory regions is rare, influenced by evolutionary age, and subject to pleiotropic constraints. Mol. Biol. Evol. 34: 2856–2869. https://doi.org/10.1093/molbev/msx219

Szpiech, Z. A., and R. D. Hernandez, 2014 selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. Mol. Biol. Evol. 31: 2824–2827. https://doi.org/10.1093/molbev/msu211

The 1000 Genomes Project Consortium, R. A. Gibbs, E. Boerwinkle, H. Doddapaneni, Y. Han, *et al.*, 2015 A global reference for human genetic variation. Nature 526: 68–74.

The FANTOM Consortium and the RIKEN PMI and CLST (DGT)Forrest, A. R., H. Kawaji, M. Rehli, J. K. Baillie *et al.*, 2014 A promoter-level mammalian expression atlas. Nature 507: 462–470. https://doi.org/10.1038/nature13182

Tishkoff, S. A., F. A. Reed, A. Ranciaro, B. F. Voight, C. C. Babbitt *et al.*, 2006 Convergent adaptation of human lactase persistence in Africa and Europe. Nat. Genet. 39: 31–40. https://doi.org/10.1038/ng1946

Uhlen, M., L. Fagerberg, B. M. Hallstrom, C. Lindskog, P. Oksvold *et al.*, 2015 Tissue-based map of the human proteome. Science 347: 1260419. https://doi.org/10.1126/science.1260419

Varki, A., and P. Gagneux, 2009 Human-specific evolution of sialic acid targets: Explaining the malignant malaria mystery? Proc. Natl. Acad. Sci. USA 106: 14739–14740. https://doi.org/10.1073/pnas.0908196106

Vitti, J. J., S. R. Grossman, and P. C. Sabeti, 2013 Detecting natural selection in genomic data. Annu. Rev. Genet. 47: 97–120. https://doi.org/10.1146/annurev-genet-111212-133526

Voight, B. F., S. Kudaravalli, X. Wen, and J. K. Pritchard, 2006 A map of recent positive selection in the human genome. PLoS Biol. 4: e72 (erratum: PLoS Biol. 4: e154 and 5: e147). https://doi.org/10.1371/journal.pbio.0040072

Wang, J. Z., Z. Du, R. Payattakool, P. S. Yu, and C. F. Chen, 2007 A new method to measure the semantic similarity of GO terms. Bioinformatics 23: 1274–1281. https://doi.org/10.1093/bioinformatics/btm087

Weir, B. S., and C. C. Cockerham, 1984 Estimating F-Statistics for the Analysis of Population Structure. Evol. 38: 1358–1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x

Wray, G. A., 2007 The evolutionary significance of cis-regulatory mutations. Nat. Rev. Genet. 8: 206–216. https://doi.org/10.1038/nrg2063

Yu, G., F. Li, Y. Qin, X. Bo, Y. Wu *et al.*, 2010 GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. Bioinformatics 26: 976–978. https://doi.org/10.1093/bioinformatics/btq064

Communicating editor: S. Tishkoff