Accounting for *cis*-regulatory constraint prioritizes genes likely to affect species-specific traits

Alexander L. Starr[1], David Gokhman[1,2], Hunter B. Fraser[1,3]

1. Department of Biology, Stanford University, Stanford, CA, USA
2. Current address: Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel
3. Correspondence: hbfraser@stanford.edu

Abstract

Measuring differences in allele specific gene expression in hybrids between two different species is a powerful way to detect cis-regulatory changes underlying adaptation. However, these studies frequently identify thousands of differentially expressed genes. This makes it difficult to differentiate candidate genes most likely to explain interspecies phenotypic differences from genes whose expression levels diverge with little evolutionary constraint. Here we outline a simple strategy that leverages population-scale allele-specific RNA-seq data to identify genes that have constrained *cis*-regulation within species yet show divergence between species. By applying this strategy to data from human/chimpanzee hybrid cortical spheroids we identify signatures of lineage-specific selection on the expression of genes related to cellular proliferation, speech, glucose metabolism, and glycan degradation. In addition, we highlight human-derived differences in the expression of *CUX1* and *EDNRB* that may contribute to differences in neural connectivity and brain size between humans and other primates.

Introduction

Changes in gene expression are thought to play a major role in the evolution of complex traits (Reilly and Noonan 2016; Romero, Ruvinsky, and Gilad 2012; King and Wilson 1975; Fraser 2013). As a result, comparing gene expression between species can enable the identification of molecular changes underlying differences in organismal phenotypes. However, obtaining accurate comparisons of gene expression between species is challenging due to confounding factors like age, environmental effects, differential cell type abundances, differences in developmental timing, and batch effects (Kelley and Gilad 2020; Housman and Gilad 2020; Y. Zhu et al. 2018a). The use of interspecies hybrids overcomes these issues through the measurement of allele-specific expression (ASE) (Gokhman et al. 2021; Agoglia et al. 2021). In hybrids, the genomes of both species share the same nucleus and are exposed to identical environments, so there are no confounding environmental, batch, compositional, or developmental effects. This approach has been successfully applied in many species and advanced our understanding of the evolution of gene regulation and its role in establishing phenotypic variation (Hu et al. 2022; Mack, Campbell, and Nachman 2016; Combs et al. 2018; X. Zhang and Borevitz 2009). Furthermore, the recent development of human/chimpanzee allotetraploid “hybrid” cells and organoids enables detailed, accurate quantification of differences in gene expression between humans and our closest living relatives (Gokhman et al. 2021; Agoglia et al. 2021).

Hybrids also enable the separation of *cis* and *trans* components of interspecies differences in gene expression (Gokhman et al. 2021; Agoglia et al. 2021). The *cis*-component is caused by differences in regulatory elements such as promoters or enhancers that only affect the expression of a nearby gene or genes on the same DNA molecule. The *trans*-component stems from changes in diffusible molecules such as transcription factors that can regulate gene

expression throughout the genome. In hybrids the genomes of the two species are exposed to the same *trans* factors. As a result, allele-specific differences in gene expression can only be explained by *cis*-regulatory differences. In addition, using ASE instead of gene expression levels from non-hybrid samples to identify differentially expressed genes (referred to as AS-DEGs when hybrids are used) enables elimination of many confounding factors. This not only increases the signal-to-noise ratio, but also disentangles important, potentially evolutionarily significant gene-specific *cis*-regulation from broad *trans*-acting changes (Prud'homme, Gompel, and Carroll 2007; Wittkopp and Kalay 2012; Fraser 2013).

While the resulting list of AS-DEGs from hybrids is likely more accurate and isolates the *cis*-regulatory component, there are often thousands of AS-DEGs which makes it difficult to prioritize candidate genes and pathways that may have played a major role in evolution. Differential expression p-values and fold changes are commonly used to rank genes in comparative RNA-seq studies. However, many AS-DEGs with moderate to large fold changes likely have relatively low constraint on their expression levels. Furthermore, some large fold changes could result in no phenotypic changes (e.g. in pseudogenes) or could result in phenotypic changes that did not evolve adaptively. In both cases, the expression of the gene evolves under limited constraint and the large fold change alone is not sufficient to determine the importance of the gene in the evolution of the parental species. For example, consider a gene whose expression varies by two-fold between species. If this gene also varies by two-fold between individual members of the same species, it is unlikely to account for any species-specific phenotypes. In contrast, a gene that is under strong stabilizing selection—with little variation in expression within species but with a two-fold change between species–is more likely to have contributed to phenotypic divergence between species. This idea of comparing gene expression divergence between species to diversity within species has been applied in several studies, but its utility has been limited by small sample sizes and the confounding factors

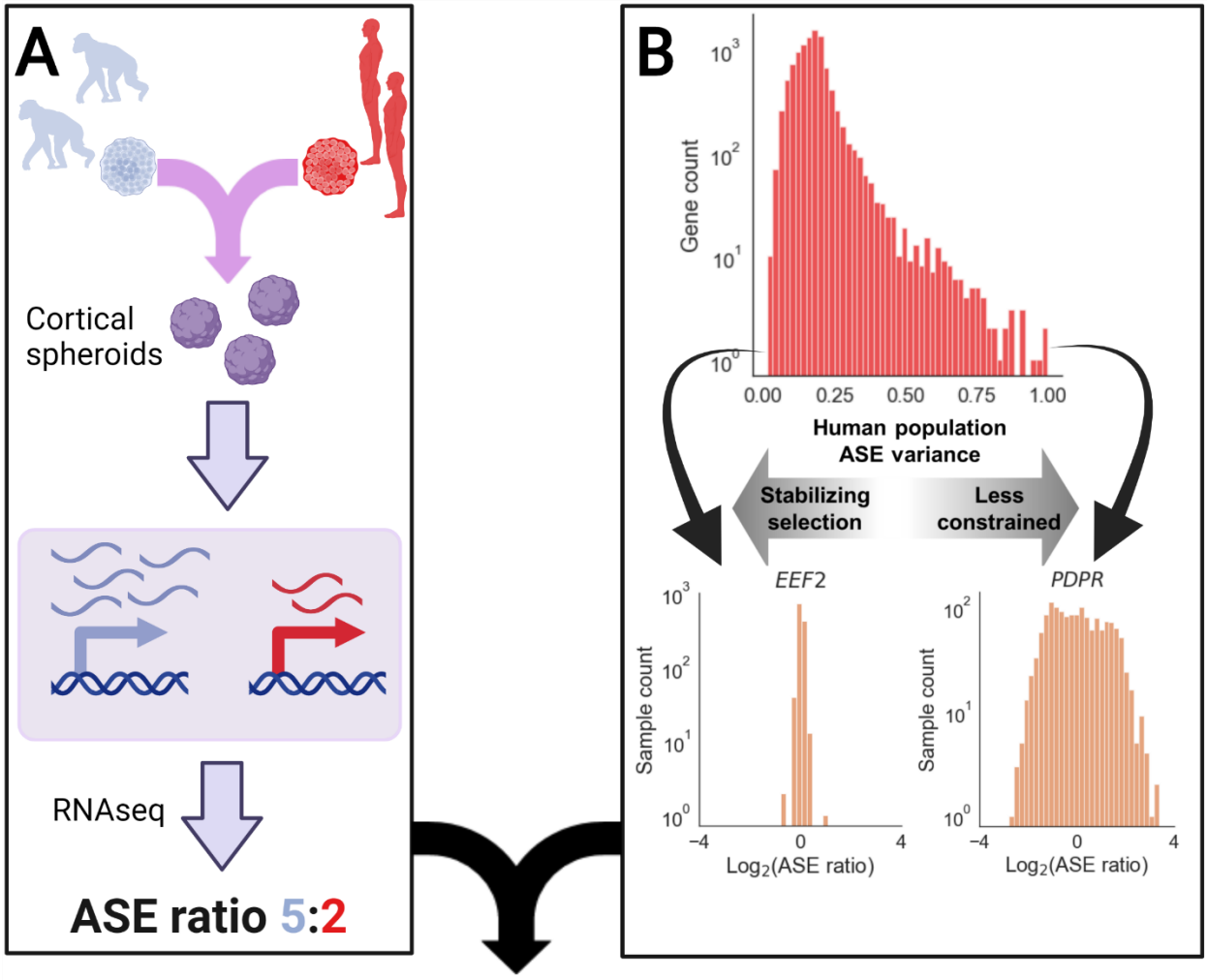
inherent to any between-species comparison (Blekhman et al. 2008; Gilad et al. 2006; Roy et al. 2013).

Here, we describe a method that leverages population-scale ASE data to approximate constraint on *cis*-regulation of gene expression. This method ranks AS-DEGs identified from interspecies hybrids in a way that is likely more able to prioritize adaptive, functionally significant changes. We apply this method to ASE data from human/chimpanzee hybrid cortical spheroids that recapitulate the gene expression patterns of the developing cerebral cortex. We identify lineage-specific selection on the expression of genes related to speech, glucose metabolism, cellular proliferation, and glycan degradation (Gokhman et al. 2021; Agoglia et al. 2021). In addition, we highlight divergence in the expression of *CUX1* and *EDNRB* that may have played an important role in human brain evolution.

Results

In a typical ASE pipeline, AS-DEGs are identified by comparing the RNA-seq read counts from the allele from species 1 and the allele from species 2 and ranked using a p-value for differential expression (Fig. 1A). Various enrichment tests and knowledge from the literature can then be used to identify interesting trends and prioritize candidate genes. However, these previous methods do not consider within-species variation in gene expression levels (Fig. 1B). If some genes have highly variable expression even within a single species, then differences of a similar (or smaller) magnitude between species are unlikely to explain species-specific traits (e.g. PDPR in Fig. 1B). Conversely, differential expression of genes with highly constrained expression in at least one species are more likely to be responsible for differences in organismal phenotypes between species. For example, ZNF331 and RPS16 have similar fold change magnitudes in human/chimpanzee hybrid cortical spheroids (Fig. 1C). However, the fold change

for ZNF331 lies well within the distribution of fold changes between alleles in the human population whereas the fold changes for RPS16 are nearly outside the human population distribution (Fig. 1C). This indicates that the expression level of RPS16 is much more constrained and that its differential *cis*-regulation between human and chimpanzee is more likely to have phenotypic consequences.


A
Cortical spheroids
RNAseq
ASE ratio 5:2
B
Gene count
Human population ASE variance
Stabilizing selection
Less constrained
EEF2
PDPR
Sample count
Log2(ASE ratio)

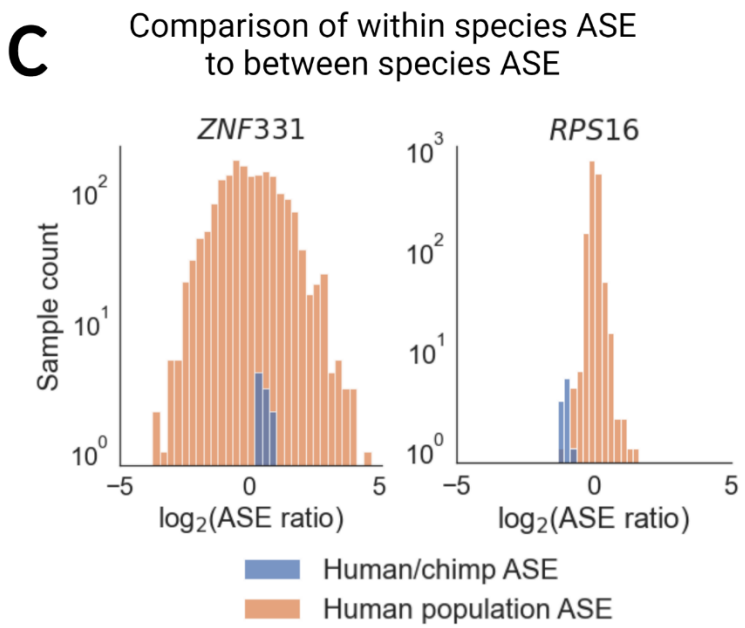

C
Comparison of within species ASE to between species ASE
ZNF331
RPS16
Sample count
log2(ASE ratio)
Human/chimp ASE
Human population ASE

Figure 1: A) Outline of a typical ASE pipeline. Hybrids are generated and RNA-seq is used to determine the relative expression of each allele. The ASE ratio is computed as the ratio of species-specific read counts between the two alleles. **B)** The distribution of the variance in ASE ratio for each gene in the GTEx data. Insets in orange show two genes at the extreme ends (EEF2 with low variance suggesting strong stabilizing selection, and PDPR with high variance suggesting less constraint on gene expression). **C)** Schematic of incorporation of the interspecies ASE and population level ASE. ZNF331 has a wide range of ASE values, and the human/chimpanzee ASE is well within the population distribution whereas human-chimp ASE in RPS16 is on the edge of the population distribution indicating greater potential for functional relevance. For both **B** and **C**, only GTEx brain samples were used.

To systematically apply this concept, we compute the distribution of ASE in one population (in this case primarily all tissues in GTEx for humans) and use the Mann-Whitney U test to compare the population level ASE distribution to the interspecies ASE distribution for that gene (see Methods)(Castel et al. 2020). The p-value reflects how divergent the ASE of the gene is between species compared to its divergence within species. We then use those p-values to separately rank genes with DESeq2 false discovery rate (FDR) less than or equal to 0.1 and greater than 0.1 such that genes with FDR less than or equal to 0.1 are always ranked higher than genes with FDR greater than 0.1. This method prioritizes genes based on whether the interspecies difference in ASE is of greater magnitude than would be expected from the intra-species ASE distribution while minimizing false positives. The ranked list is then used in enrichment analyses and to identify of top candidates (Benjamini and Hochberg 1995; Love, Huber, and Anders 2014).

To minimize any allelic bias in the population distribution, we normalize the median to 1. The resulting rankings are robust to choosing either the reference or alternative allele as the

numerator for the population distribution (spearman's rho = 0.999 $p < 10^{-300}$). Notably, we minimize any batch effects between GTEx and the human/chimpanzee hybrid dataset by focusing on ASE since direct comparisons of expression levels between data sets are not involved. For example, if some technical factor (e.g. sequencing platform or RNA isolation method) caused a gene to show a spurious 2-fold higher expression in GTEx samples compared to the hybrid data, this would cancel out in the GTEx ASE calculation.

First, we tested whether the variance in ASE in the human population is a reasonable proxy for constraint on gene expression. We compared the variance of the GTEx ASE distribution for each gene to its probability of haploinsufficiency score (pHI), a measure of sensitivity to a 50% reduction in gene dosage (Collins et al., 2022.). We observe a significant negative correlation (Spearman's rho = -0.28, $p < 10^{-170}$) between the two measures (Supp. Fig. 1) indicating that the variance of the population-scale ASE distribution provides a reasonable proxy for constraint on expression levels. This correlation remains significant when only five (as opposed to ten) reads from each allele for a gene in a sample are required for inclusion (Spearman's rho = -0.27, $p < 10^{-163}$). Furthermore, as pHI is computed from copy number variation, there is no circularity in reasoning when comparing pHI to the variance of the GTEx ASE distribution (Collins et al., 2022). While this indicates that variation in ASE is sufficient to approximate evolutionary constraint on gene expression, the strength of this relationship should not be interpreted as a quantitative estimate of how well ASE represents constraint.

GTEx contains data from a wide range of adult tissues, some of which may not reflect gene expression constraint in cortical spheroids, which mimic fetal development (Paşca et al. 2015; Ferraro et al. 2020). To test this, we used an ASE dataset generated from neural progenitors, neurons, and fetal cortical wall which resembles the developmental stage of cortical spheroids more closely than GTEx samples (Aygün et al. 2021). There is a strong correlation between the

rankings generated by comparing to the fetal cortex dataset and GTEx (Spearman's rho = 0.92, $p < 10^{-170}$ for each of day 50, day 100, and day 150 after initial differentiation of cortical spheroids; Supp Fig. 2). Due to the higher number of genes that have sufficient data in GTEx, we focus on it throughout but verify results for specific genes in the fetal cortex data when possible.

To highlight differences between our method and the traditional method for ranking genes (Fig 1A, C), we computed the difference in ranks between the two methods and used GSEAPY to identify gene sets enriched near the top or bottom of the resulting sorted list (referred to as the difference in ranks list, Fig 2A)(Subramanian et al. 2005). A top-ranking gene would be one that is lowly ranked by the traditional method but highly ranked by the population comparison method. Many gene ontology (GO) categories related to cortical development were enriched near the top of the difference in ranks list including calcium and potassium channel activity, various transcription factor (TF) related terms, and cytokine activity (Fig 2 B, C and Supp. Table 1). This is consistent with previous observations that TFs, ion channel subunits, and important players in canonical signaling pathways tend to have strongly constrained expression compared to other genes(Collins et al., n.d.). Many genes that drive the enrichment of the term "RNA polymerase II regulatory region sequence-specific DNA binding" are haploinsufficient and play key roles in neurodevelopment including *MEF2C*, *NEUROD2*, and *CUX1* (Supp. Fig. 3A) (Platzer et al. 2018; Paciorkowski et al. 2013; Runge et al. 2021). Differences in the expression of haploinsufficient genes are more likely to have phenotypic consequences than differences in the expression of genes for which loss of one copy has no clear phenotypic effect. In particular, there is strong evidence that both increases and decreases in CUX1 expression alter behavior in humans, which implies that the CUX1 expression divergence between humans and chimpanzees is likely to have phenotypic consequences (Box 1) (Doan et al. 2016; Platzer et al. 2018).

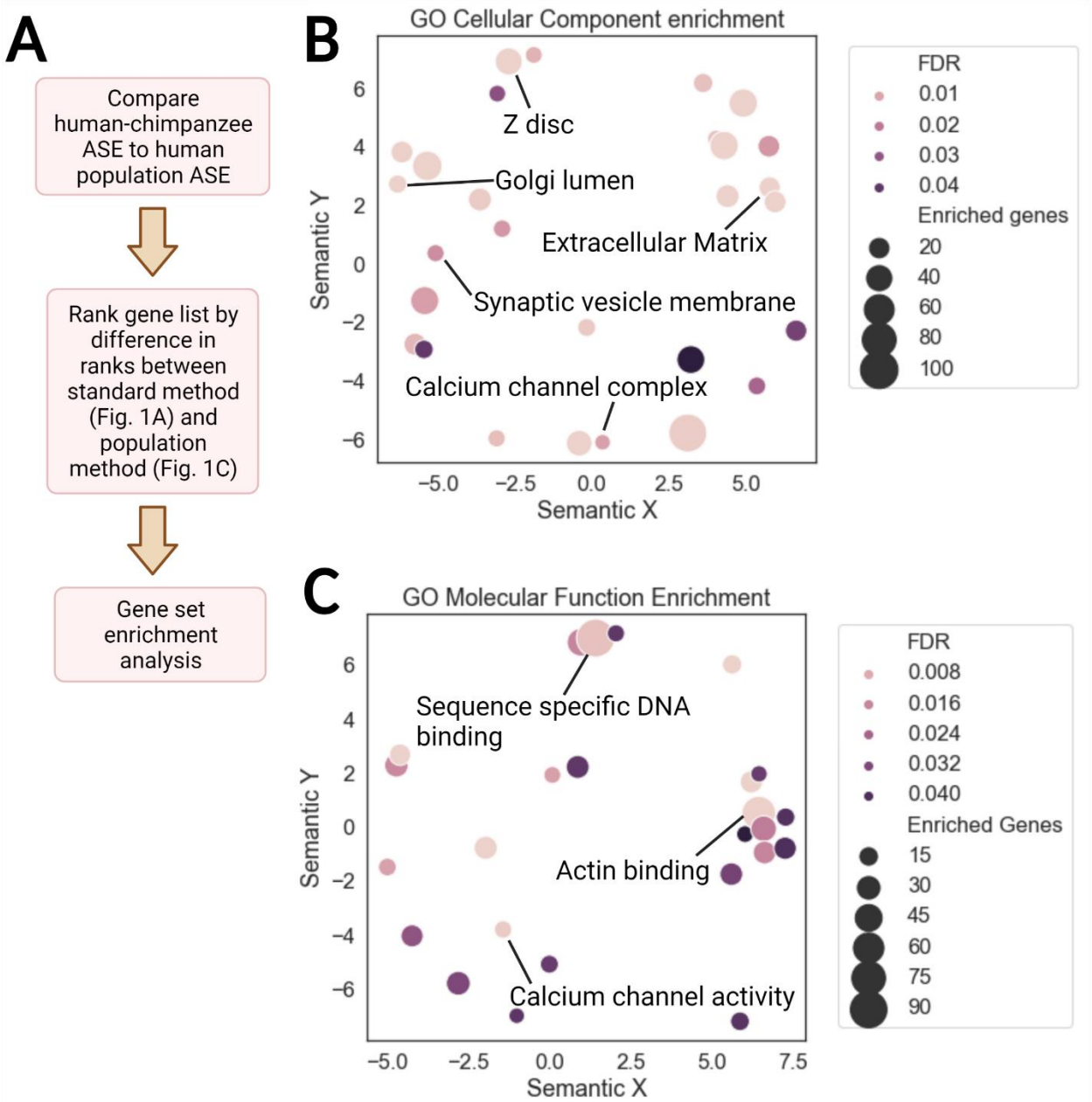


Figure 2: A) Pipeline for gene set enrichment analysis. Gene set databases tested include Gene Ontology (GO) Cellular Component, GO Molecular Function, Kyoto Encyclopedia of Genes and Genomes (KEGG) human gene sets, Human Phenotype Ontology (HPO), and REACTOME (Köhler et al. 2021; Carbon et al. 2009; Jassal et al. 2019; Kanehisa et al. 2016; Köhler et al. 2014). **B)** Summary of GO Cellular Component enrichments across all time points with false discovery rate (FDR) < 0.05. REVIGO in conjunction with a custom python script was

used to generate the plot. The axes are derived from multidimensional scaling and measure semantic similarity. Each circle represents a GO term and circles near each other contain similar genes in the corresponding gene set. Labeled gene sets are those particularly relevant to cortical development. **C)** Summary of GO Molecular Function enrichments across all time points with FDR < 0.05.

Box 1: Recent work has linked a mutation in a human accelerated region (HAR) that likely increases *CUX1* expression to autism spectrum disorder (Doan et al. 2016). *CUX1* is expressed at a lower level in humans than chimpanzees across time points in both hybrid and parental cortical spheroids (e.g. $\log_2$ fold-change of -0.93, FDR < 0.005 in parentals, mean $\log_2$ fold-change of -0.74, FDR < 0.023 in hybrids at day 150 of differentiation) and the per-sample ASE ratios are well outside the human fetal cortex population ASE distribution (Supp. Fig. 3 B-E). This reduced expression in humans is surprising considering that a recent massively parallel reporter assay (MPRA) found that the HAR linked to *CUX1* should increase expression in humans (Girskis et al. 2021). As expression is lower in humans, haploinsufficiency of *CUX1* might provide a reasonable model of the phenotypic consequences of this change. *CUX1* haploinsufficiency in humans leads to delayed development of speech and motor skills (Platzer et al. 2018). One aspect of this condition is that individuals with one *CUX1* allele often close the developmental gap over time (i.e. cognitive impairments and delays disappear with age) (Platzer et al. 2018). This “catch up” phenotype is very rare; and may even be specific to *CUX1* haploinsufficiency (Platzer et al. 2018). Interestingly, humans develop more slowly than other great apes (known as neoteny) but eventually “catch up,” reminiscent of the *CUX1* phenotype. While changes in *CUX1* expression may have played a role in causing human neoteny, investigation of the development of layer II-III cortical neurons and behavior of *CUX1* haploinsufficient mice will be required to explore this further.

Applying the Sign Test to Constrained Differentially Expressed Genes

The sign test is designed to detect lineage-specific selection based on systematic up- or down-regulation of genes that deviates from neutral expectation (Kita et al. 2017; Fraser 2011).One important requirement of the test is that the expression divergence of every gene being tested should be driven by independent genetic differences. For example, a single mutation in a *trans*-acting factor could cause a whole pathway to be down-regulated, but this would only count as one genetic difference. For this reason, hybrids are ideal for applying the sign test to gene expression, since *cis*-regulation of genes is typically independent (except in the case of some neighboring genes that share *cis*-elements, which can be accounted for by using a minimum distance threshold between genes). For example, in the human/chimpanzee hybrid, if a particular pathway contains significantly more genes with higher expression from the human allele than the chimpanzee allele (significance determined by a binomial test), that would provide evidence of selection on that pathway in the human and/or chimpanzee lineage.

To prioritize enrichments, we used the sign test to identify gene sets in which highly ranked genes (i.e. those with a difference in ranks between the population comparison method and the traditional method above the cutoff found by GSEAPY) show a systematic bias in directionality. Three gene sets are significant by the binomial test at an FDR cutoff of 0.1: *Neuronal Cell Body*, *Cytokine Activity* and *Cytoskeleton* (Supp. Table 2, FDR = 0.085 for all three terms). Notably, the *Cytokine Activity* category is dominated by genes that promote stem cell proliferation which may have consequences for the proliferation of human and chimpanzee radial glia (although we caution against overinterpretation of this result given the low fold changes of most genes in this category, Box 2).

Box 2: The *Cytokine Activity* category is enriched near the top of the difference in ranks list with 15 out of 16 genes driving the enrichment displaying chimpanzee-biased expression (Fig. 3A, B). It is dominated by genes that generally promote neural stem cell proliferation, which is unexpected considering the directionality bias and the higher proliferative capacity of human neural stem cells. *EDN1* had the strongest chimpanzee bias (mean $\log_2$ fold-change = -0.92, FDR < 0.025 at Day 150, Fig 3B). As *EDN1* primarily signals through *EDNRB* in the brain(Adams et al. 2020), we also investigated the expression of *EDNRB*. Surprisingly, *EDNRB* is one of the most strongly human-biased genes across all timepoints in both hybrid and parental cortical spheroids (mean $\log_2$ fold-change = 4.46, FDR < 0.005 at Day 150 in hybrids, $\log_2$ fold-change = 2.85, FDR < 0.0005 in parental samples). Human/chimpanzee ASE generally exceeds ASE found in human populations for *EDNRB* (Fig 3C) although this is not the case for *EDN1* (Mann Whitney U Test comparing EDNRB distribution to human population distribution p = 0.00027 for *EDNRB*, Fig 3D). The human-biased *EDNRB* expression and chimpanzee-biased *EDN1* expression is generally consistent across timepoints in hybrid cortical spheroids (Fig 3E, F).

Changes in *EDNRB* expression appear to be human-derived with respect to gorillas and macaques (although orangutans may have independently acquired similar expression to humans in early-stage brain organoids) (Supp. Fig. 4 A, B). We were unable to confidently determine whether changes in *EDN1* expression were human- or chimpanzee-derived using currently available data (data not shown). Next, we analyzed single-cell RNA-seq data generated from 1 month, 2 month, and 4 month-old human and chimpanzee brain organoids to identify the cell types driving increased EDNRB expression. We found that a previously identified radial glial cell (RGC) cluster was characterized by high *EDNRB* expression with non-zero expression in over 50% of cells (Kanton et al. 2019). Furthermore, this cluster had higher expression than any chimpanzee cluster (Mann-Whitney U test, $p < 10^{-16}$, Supp. Fig. 5-6).

Overall, our results suggest that a subpopulation of human radial glia have much higher *EDNRB* expression than chimpanzee radial glia.

EDNRB haploinsufficiency reduces proliferation of cerebellar granule precursor cells and chemical inhibition of *EDNRB* signaling reduces proliferation of mouse radial glia (Vidovic et al. 2008; Shinohara et al. 2004). Based on this, the change in *EDNRB* expression may have promoted human brain expansion by increasing the proliferation of the subpopulation of radial glia that express *EDNRB*. In addition, a recent study identified a population of caudal late interneuron progenitor (CLIP) cells marked by expression of *EDNRB* and *PTGDS* along with caudal ganglionic eminence markers (Eichmüller et al. 2022). As both *EDNRB* and *PTGDS* have strongly human-biased ASE it would be interesting to investigate if this population of cells exists in chimpanzee brain organoids and if it may have expanded in humans.

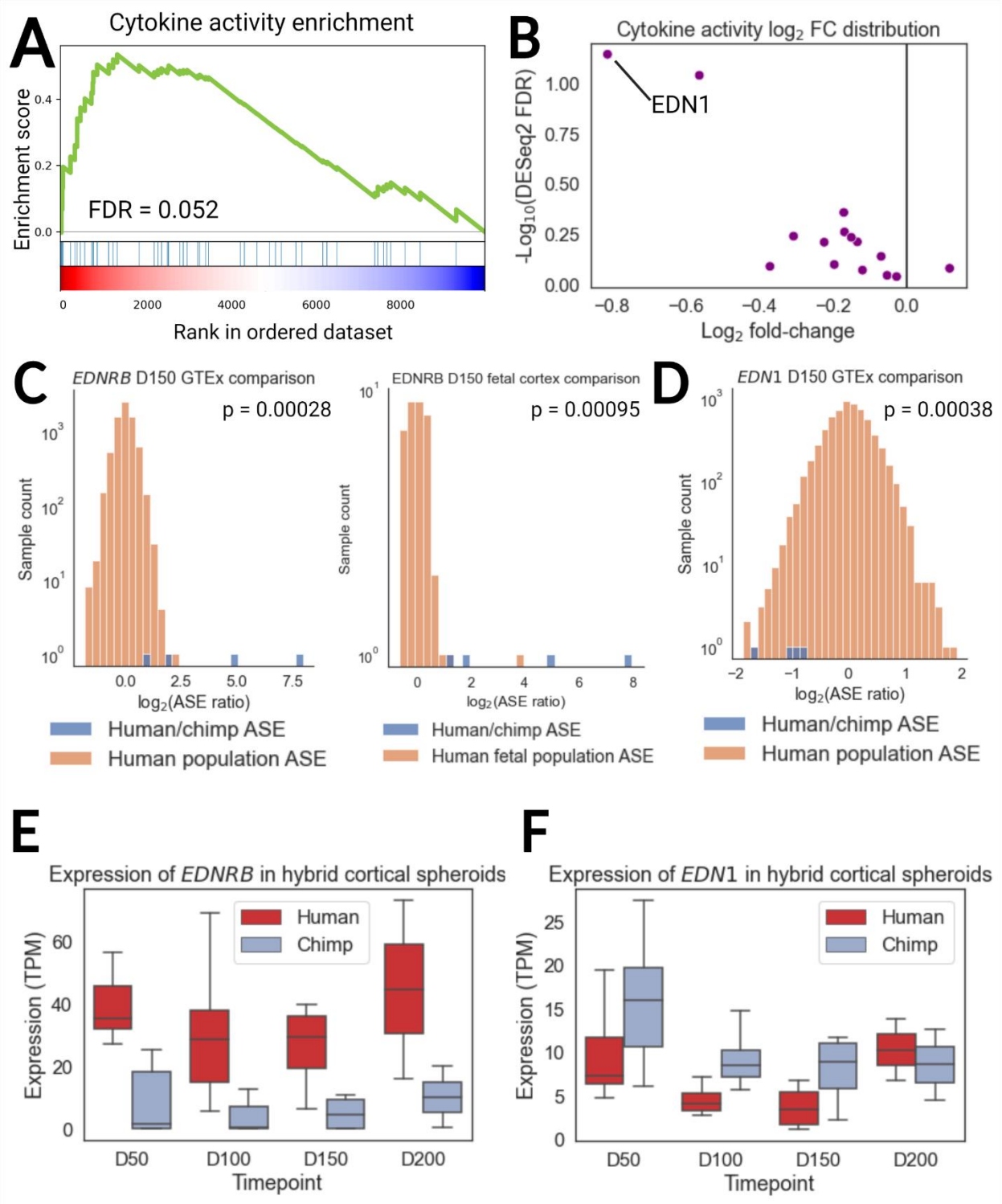


Figure 3: A) Summary of Cytokine Activity enrichment. Each blue line represents a gene in the gene set and the green curve is the cumulative enrichment score. Top ranking genes could be human or chimpanzee biased in expression. **B)** Volcano plot showing expression levels of

genes driving Cytokine Activity enrichment. The log_2 fold-change (log_2 FC, which refers to the value computed by DESeq2) is chimpanzee-biased for 15 out of 16 genes. Mean log_2 FC between chimpanzee-referenced and human-referenced log_2 FC values is shown. Negative log_2 FC indicates chimpanzee-biased expression. **C)** Comparison of human/chimpanzee EDNRB ASE to GTEx and fetal cortical within human population ASE. Raw ASE ratios (as opposed to the value derived from DESeq2) are indicated by "ASE Ratio". The human/chimpanzee ASE is significantly outside of the human ASE distribution. P-values are from the Mann-Whitney U Test comparing the distribution of human population ASE to human/chimpanzee ASE. **D)** Comparison of human/chimpanzee EDN1 ASE to GTEx and human population ASE. Fetal cortical ASE is not shown due to insufficient data. P-value is the same as in C. **E)** ASE of EDNRB across timepoints in cortical spheroids. Expression from the human allele is consistently higher than expression from the chimpanzee allele. **F)** ASE of EDN1 across timepoints in cortical spheroids. Expression from the chimpanzee allele is consistently higher than expression from the human allele except at D200.

Thus far, we have primarily sought to demonstrate the differences between our new method and the traditional method for ranking genes. However, another approach is to focus only on the new ranking rather than this difference. In this approach, we rank genes by the metric outlined in Fig. 1 so that highly ranked genes with human-biased expression are at the top of the list and highly ranked genes with chimpanzee-biased expression are at the bottom of the list. In effect, this results in a test for directionally-biased cis-regulatory divergence that exceeds the cis-regulatory variation between most human alleles present in the GTEx population. Enrichment testing with GSEAPY identified several enriched gene sets at an FDR cutoff of 0.25 including *Other glycan degradation* (human-biased), *Spastic Dysarthria* (human-biased), and *Gluconeogenesis* (chimpanzee-biased) (Fig 4 A-F). In all three cases, there were zero genes

showing expression bias in the other direction at an identical rank cutoff ($p = 0.00012$ for *gluconeogenesis*, $p = 0.031$ for *Other glycan degradation*, and $p = 0.00195$ for *Spastic Dysarthria* by binomial test), suggesting lineage-specific selection on genes with constrained expression in these gene sets (Fraser 2011).

All three of these enriched categories may influence human-specific phenotypes. Gluconeogenesis siphons oxaloacetate from the TCA cycle and eventually produces glucose (Gkini and Namba 2022). Many of the gene expression changes driving the gluconeogenesis enrichment appear to be human-derived compared to other great apes (Supp. Fig. 7). Decreased gluconeogenesis in the human lineage would likely enable increased flux through other anabolic pathways and the TCA cycle possibly increasing the availability of oxaloacetate for anabolic pathways that promote proliferation. Spastic dysarthria is a condition in which patients speak in a characteristic slow, regular, monotone manner (Clark et al. 2014). Loss of function of genes in this category (as opposed to pathogenic gain of function mutations) are associated with spastic dysarthria and show systematic human bias, which may be connected to the human capacity for speech (Fig 4D). Finally, six genes that were ranked very highly by our method and all have DESeq2 FDR < 0.05 drive the “Other glycan degradation” enrichment (Box 3).

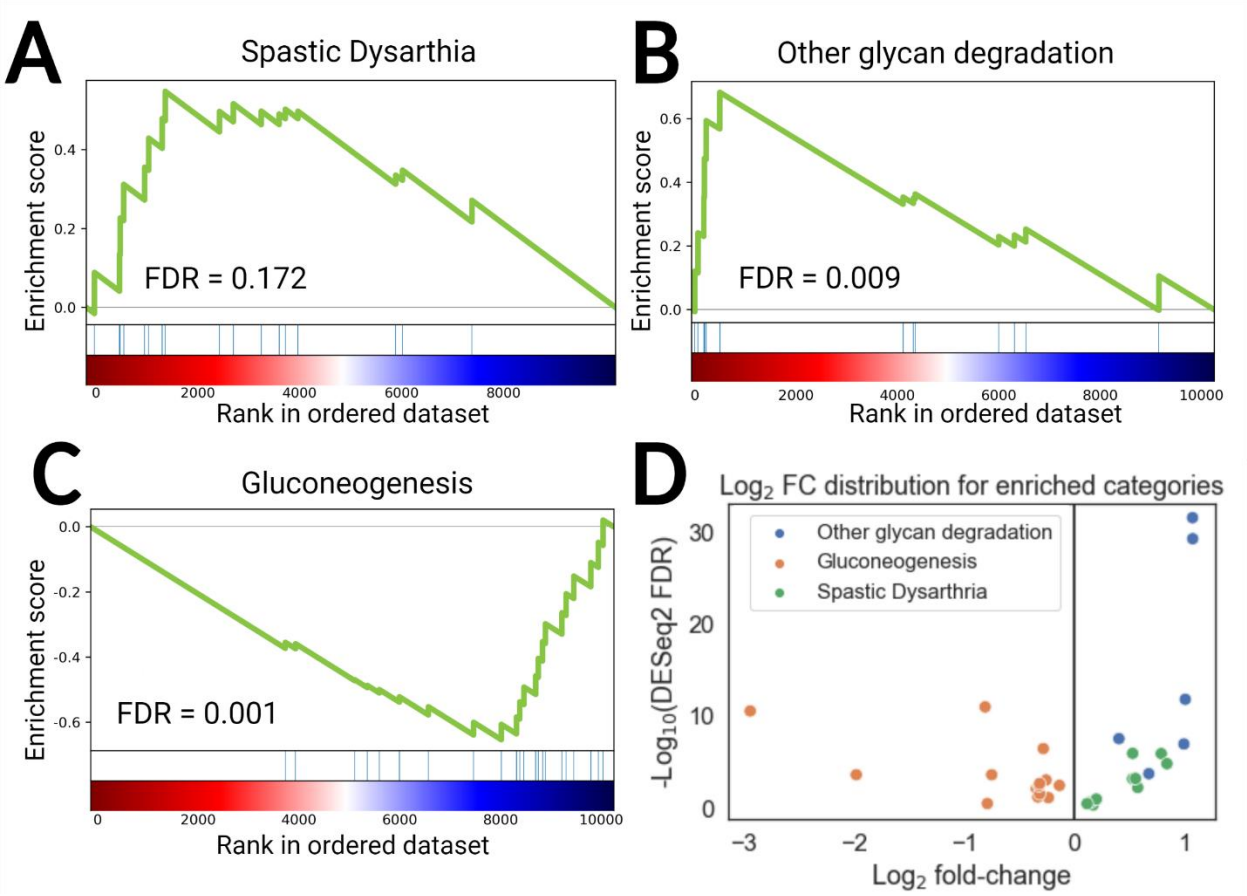


Figure 4: A) Summary of *Spastic dysarthia* enrichment (from Human Phenotype Ontology)(Köhler et al. 2014; 2021). In A, B, and C, each blue line represents a gene in the gene set and the green curve is the cumulative enrichment score. **B)** Summary of *Other glycan degradation* enrichment (from KEGG). **C)** Summary of *Gluconeogenesis* enrichment (from REACTOME). **D)** Volcano plot summarizing of Log2 fold-changes for genes driving the enrichments for *Spastic Dysarthria*, *Other glycan degradation*, and *Gluconeogenesis*. All genes are human-biased for *Other glycan degradation* and *Spastic Dysarthria* whereas all genes are chimpanzee-biased for *Gluconeogenesis*.

Box 3: Glycans are polysaccharide polymers involved in a wide range of biological processes. Interestingly, loss of function of three of the six human-biased glycan degradation genes (MANBA, MAN2B2, and MAN2B1) is associated with intellectual disability (Alkhayat et al. 1998; Cathey et al. 2019; Blomqvist et al. 2019). Genes related to glycan biosynthesis are generally expressed at much lower levels in the brain than other tissues in adult (Williams et al. 2022). To test whether this is also true for these genes in fetal tissues, we compared mean expression in the brain to mean expression in 11 other organs (Cao et al. 2020). Four out of six genes (MAN2B1, MAN2B2, FUCA2, and ENGASE) showed lower expression in the brain than in all or all but one of the other 11 tissues (Supp. Fig. 8). 14% of all genes have lower expression in the brain than all or all but one of the 11 other organs in the fetal single cell dataset. Due to this, we used the binomial test and determined that the four out of six glycan degradation genes that show this pattern is more than expected by chance (p-value = 0.0048). Combined with the high ranking given to these genes by our method, this implies that these genes may generally be constrained to be lowly expressed in the brain and that the increased expression of these genes in humans compared to chimps may lead to differences in their glycome.

Discussion

Here we presented a method incorporating population-scale ASE data as a proxy for constraint on expression. This ranking method helps reveal candidate genes and signatures of selection that may explain phenotypic differences between humans and chimpanzees. Although the phenotypic consequences of these differences remain to be determined, our finding of polygenic lineage-specific selection on several gene sets suggests that these changes must have some phenotypic effects in order to be under natural selection. Collectively, these findings more than double the number of known cases of lineage-specific polygenic selection on gene expression

between humans and chimpanzees (the two previous examples being Hedgehog signaling and astrocyte-related genes) (Gokhman et al. 2021; Agoglia et al. 2021).

Looking ahead, we anticipate that our method will become a useful tool for identifying functionally significant gene expression changes in human/chimpanzee comparisons, as well as other species with large gene expression data sets, e.g. comparing ASE in *Arabidopsis* interspecies hybrids to ASE within *A. thaliana* (H. Zhang et al. 2020). The strong correlation between the combined GTEx rankings and the fetal cortex rankings (Supp Fig 2) suggests that comparison to the GTEx population distribution will be useful for other cell types and organoids (although GTEx includes many brain samples, which may increase the correlation with other brain-related datasets). However, population-scale datasets that more closely match the samples of interest will likely provide a more accurate estimate of the population ASE distribution. As human ASE datasets expand and diversify, such stage- and tissue-matched comparisons will become possible and enable deeper insight into the genetic basis of human-specific traits.

Materials and Methods:

Read Alignment and RNA-seq Data Processing

Data from hybrid cortical spheroids was mapped as previously described(Agoglia et al. 2021). We used the ASE $\log_2$ fold-change ($\log_2$ FC) values available in the supplemental tables of Agoglia et al. Throughout, for hybrids the mean $\log_2$ FC between human genome mapped and chimpanzee genome mapped reads is stated as well as the highest p-value. Additional data

was downloaded from GSE127898, GSE106245, GSE153076, and phs000755.v2.p1 and mapped separately for each dataset to the respective species' genome (PanTro6 for chimpanzee, hg38 for human, mmul10 for rhesus macaque, Gorgor6 for gorilla, and PonAbe3 for orangutan) (Field et al. 2019; Y. Zhu et al. 2018b; Benito-Kwiecinski et al. 2021; Khrameeva et al. 2020). We used STAR v2.5.4 with arguments: -outSAMattributes MD NH -outFilterMultimapNmax 1 -sjdbGTFfile -sjdbOverhang N where N is 1 less than the read length used for each respective dataset(Dobin et al. 2013). For paired end reads, we used Picard to remove duplicates with argument: DUPLICATE_SCORING_STRATEGY = RANDOM(Broad Institute 2019). We used HT-Seq with the following arguments: -t exon -i gene_name -m intersection-strict -r pos to count reads overlapping gene bodies (Givanna H Putri 2021). Transcripts per million (TPM) was computed as previously described (Li et al. 2010). We used the likelihood ratio test in DESeq2 to test for differential expression in the downloaded datasets with RIN and sex included as covariates for the Khrameeva et al. dataset (A. Zhu, Ibrahim, and Love 2019; Love, Huber, and Anders 2014). We binarized RIN values as high if greater than or equal to 7.5 and low otherwise as we do not generally expect the expression level of genes to scale linearly with RIN.

Comparison of Population and Interspecies ASE Distributions

GTEx data was downloaded from https://www.gtexportal.org/home/datasets (accessed 08/15/2021, although the exact file downloaded is no longer available) and the fetal cortex ASE data was kindly provided by the Stein laboratory (Castel et al. 2020; Aygün et al. 2021). To preprocess the GTEx data, we split the file into two files each containing read counts from 1 of the alleles with a custom R script. To more accurately model cortical spheroids, we pooled fetal

cortical wall, neural progenitor, and neuron counts per individual in the fetal cortex dataset. If a gene had multiple SNPs with reads, we summed those reads to generate one read count per individual per gene. For each gene in each sample, we added a pseudocount to each gene and computed the ASE ratio as the ratio of counts from allele 1 (the reference) to counts from allele 2. The distribution for each gene was then normalized so that the median was 1. This normalization removes any bias toward the reference or alternative allele. Notably, flipping the sign of each value in the GTEx ASE distribution had minimal effect on the rankings (Spearman's rho = 0.999, $p < 10^{-300}$ for D50, D100, and D150) indicating a lack of reference or alternative bias in these corrected values. All samples with at least 10 counts (not including the pseudocount) from each allele for a sample were included in the ASE population distribution. Notably the rankings and our results are robust to requiring at least 5 counts from each allele instead of 10 (Spearman's rho = 0.996, $p < 10^{-300}$). To filter out genes that are lowly expressed in cortical spheroids (and thus are unlikely to have important effects on cortical development), we removed genes with an average number of counts from the chimpanzee and human alleles less than 25 (i.e. mean of human and chimpanzee read count less than 25 and mean of chimpanzee read count less than 25). In addition, we filtered out any genes showing mapping bias (listed in the supplemental tables of Agoglia et al.) as well as genes on chromosomes 18 and 20 as parts of these chromosomes were duplicated in some cortical spheroid samples. After filtering, we computed the interspecies ASE distribution in an identical manner to the population ASE distribution except that we did not require 10 counts from each allele and did not normalize the medians. We compared the log2(ASE Ratio) interspecies distribution to the population distribution using the Mann-Whitney U Test (a nonparametric test robust to the distribution of data) and used the resulting p-values to rank genes as described below.

Generation of Gene Rankings and Enrichment Analysis

We tested gene sets from the Gene Ontology Cellular Component and Molecular Function categories, the Human Phenotype Ontology, KEGG, and REACTOME using the same version as in Gokhman et. Al (Carbon et al. 2009; Jassal et al. 2019; Kanehisa et al. 2016; Köhler et al. 2014; Gokhman et al. 2021; Köhler et al. 2021). Regardless of which ranking was used, we used GSEAPY preranked with the following arguments: processes=4, permutation_num=1000, seed=6, min_size = 10, max_size = 300 to test for enrichment (Subramanian et al. 2005). Following the authors suggestion, we considered any category with an FDR below 0.25 to be nominally enriched(Subramanian et al. 2005). We required that ASE data be available from at least 50 individuals in GTEx for a gene to be included in the ranking. We conducted gene set enrichment analysis with four different rankings. First, we ranked genes by the Mann-Whitney U Test p-value with the lowest p-value receiving the highest rank. To reduce false positives at the top of the list, we separately ranked genes with DESeq2 FDR less than or equal to 0.1 and greater than 0.1. We then concatenated the list so that genes with FDR less than or equal to 0.1 were always ranked higher than genes with FDR greater than 0.1 (referred to as the MWU ranking). This consensus gene ranking was then used in GSEAPY preranked with the rankings used as the score that GSEAPY uses to sort the list. All results highlighted in the text replicated when using 0.05 as a cutoff instead of 0.1. We used REVIGO in conjunction with a custom python script to generate the plots shown in figure 2(Supek et al. 2011).

To compare to the traditional method, we also ranked genes by the DESeq2 derived FDR and used that in GSEAPY preranked (referred to as the DESeq2 ranking). To highlight differences between the two methods, we computed the difference in ranks between the two methods by subtracting the DESeq2 ranking from the MWU ranking and sorting the list on those rankings for use in GSEAPY. In this context, highly ranked genes are likely those that show relatively mild

gene expression changes but have more constrained expression (Supp. Fig. 9). We performed the sign test on all non-redundant categories with FDR < 0.25 and greater than 10 ledge genes (i.e. genes driving the enrichment) across the three organoid time points for both the GTEx and fetal cortex datasets. We used the average of the DESeq2 $\log_2$ FC from mapping to the human allele and from mapping to the chimpanzee allele as input for the binomial test to identify gene sets with significantly more human-biased or chimpanzee-biased changes than expected by chance. We considered any gene set with Benjamini-Hochberg corrected FDR < 0.1 to be significant. Finally, we ranked the genes using a signed version of the MWU ranking. More specifically, genes were effectively ranked by the $\log_{10}$(MWU p-value) multiplied by the sign of mean DESeq2 $\log_2$ fold-change so that top ranked genes with negative L2FC are at the bottom of the list and top ranked genes with positive L2FC are at the top of the list. This ranking was then used in GSEAPY preranked. All statistical tests (Mann Whitney U Test, Binomial Test, correlations) were performed in python using the implementation in scipy.

Single Cell RNA-seq Data Processing and Analysis

Single cell data from human and chimpanzee organoids and associated metadata were downloaded from E-MTAB-7552 (Kanton et al. 2019). We used SCANPY to read in the counts matrix and filter the data so that only data from 1 month, 2 month, and 4 month old organoids remained (Wolf, Angerer, and Theis 2018). We used a two-sided Mann Whitney U Test to compare *EDNRB* log2(counts per million) between the “RGC early 2” cluster and all chimpanzee clusters with and without cells with 0 *EDNRB* counts included. Mean counts by cell type and tissue for fetal human expression were downloaded from GSE156793 (Cao et al. 2020). We compared cerebrum counts to all other organs except eye and cerebellum due to

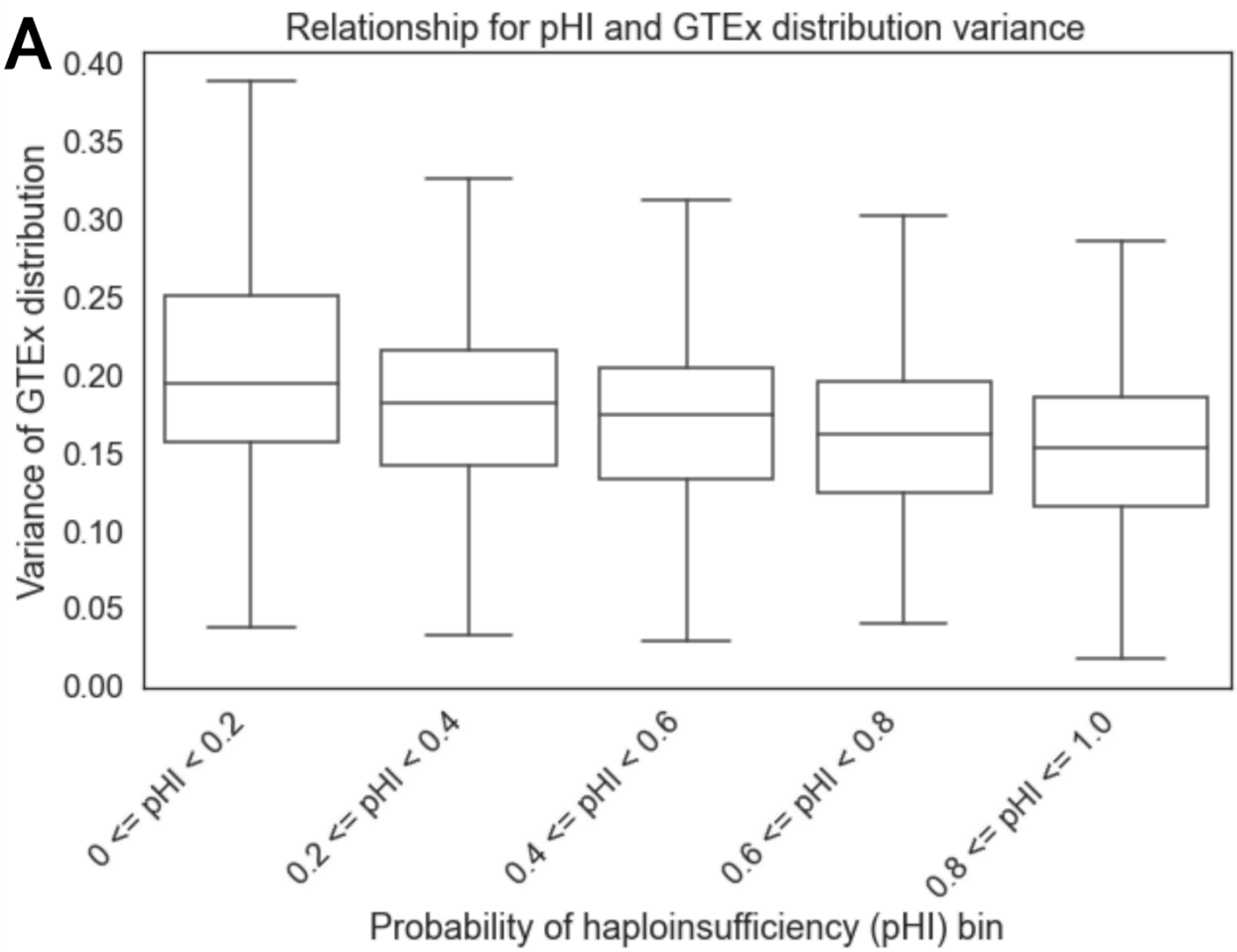
their similarity to the cerebrum and used a binomial test to determine if more tissues exhibited higher expression than brain than expected by chance.

Code Availability: All code used to analyze data and generate figures can be found at https://github.com/astarr97/Z_Score.

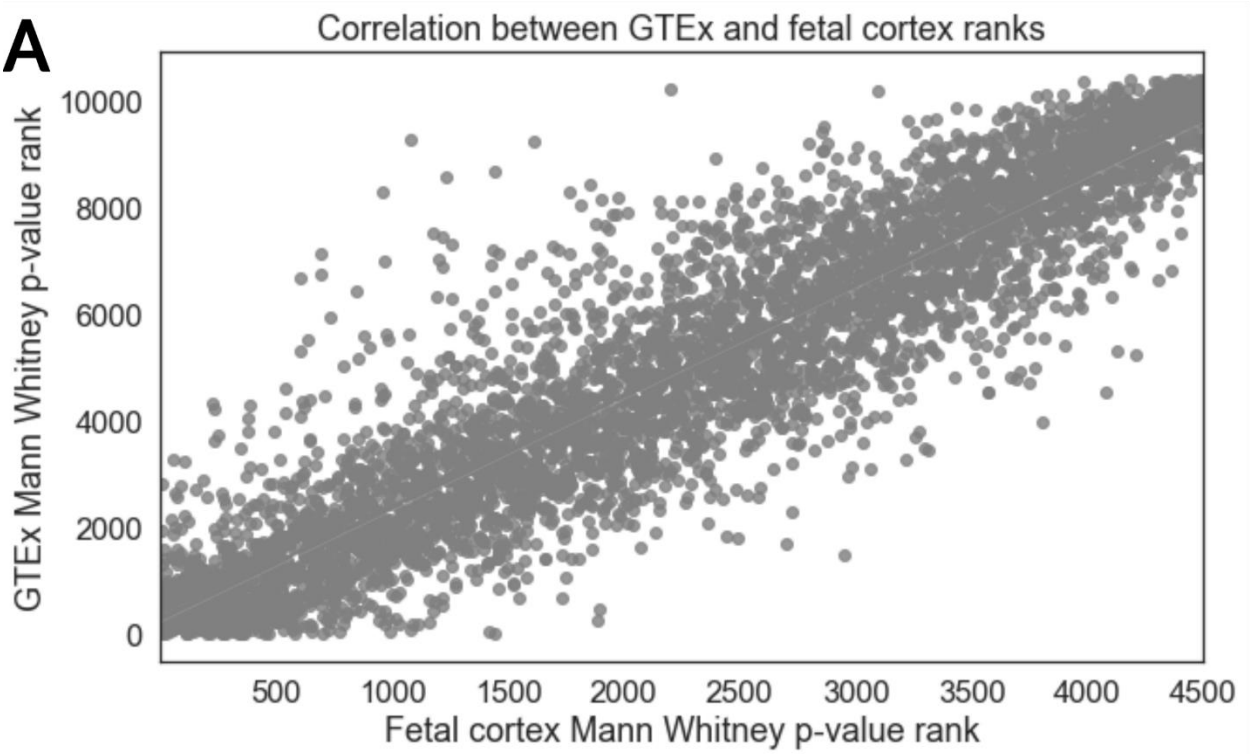
Data Availability: No new data was generated during this study. Accessions and citations for all publicly available datasets have been provided in the Materials and Methods section.

Acknowledgements: We thank Jason Stein and Nil Aygun for providing processed fetal cortex, neuron, and neural progenitor ASE calls. We thank members of the Fraser lab for helpful discussion. Illustrations were created with BioRender.

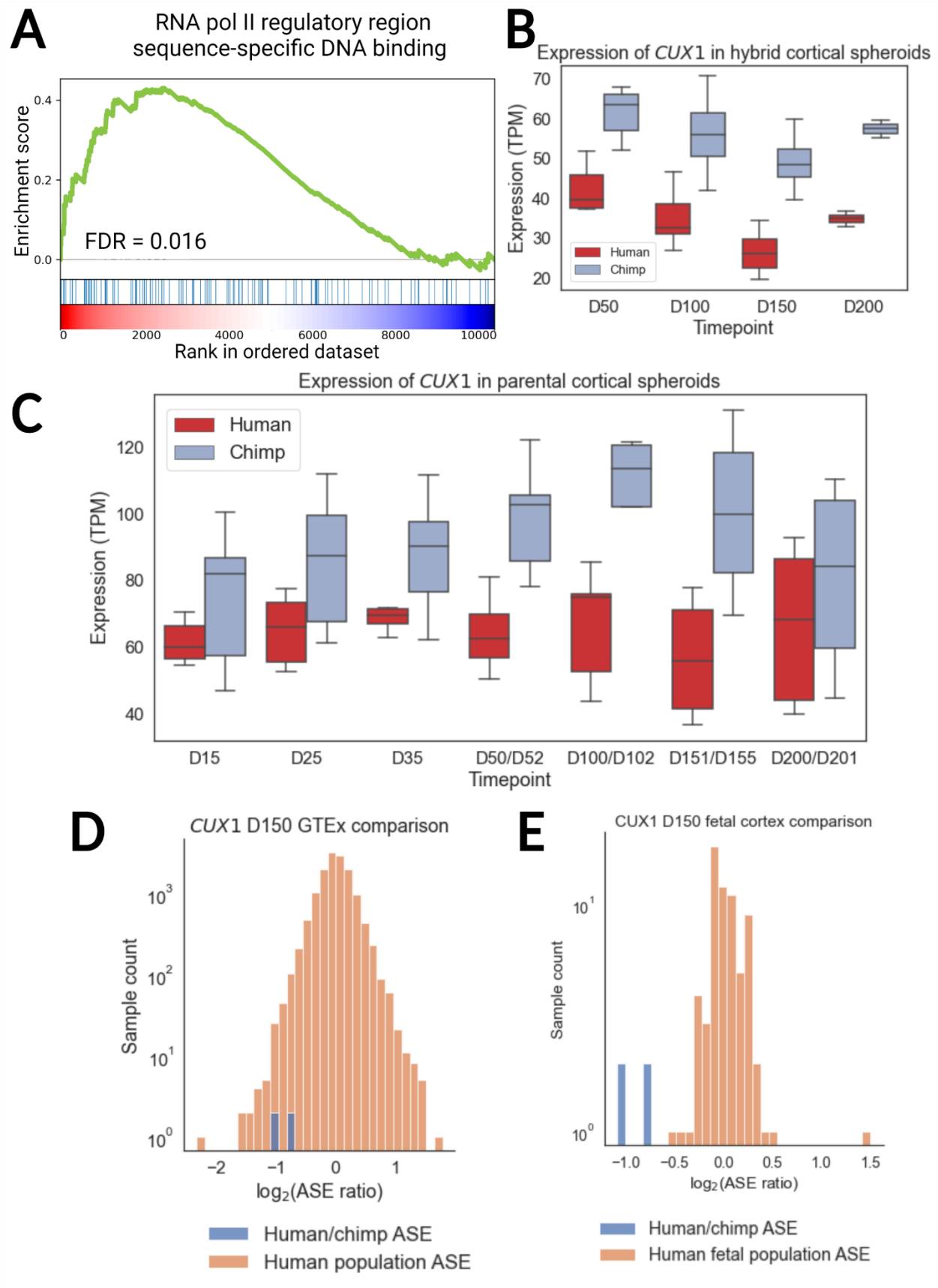
Competing interests: To our knowledge, we have no competing interests.



Supplementary Figure 1: A) Shows the relationship between the variance of the GTEx ASE distribution and the probability of Haploinsufficiency score (a measure of constraint on gene expression) for each gene tested in this manuscript. As the variance of the GTEx distribution increases, the probability of Haploinsufficiency increases. Spearman correlation is -0.28 with $p < 10^{-170}$.

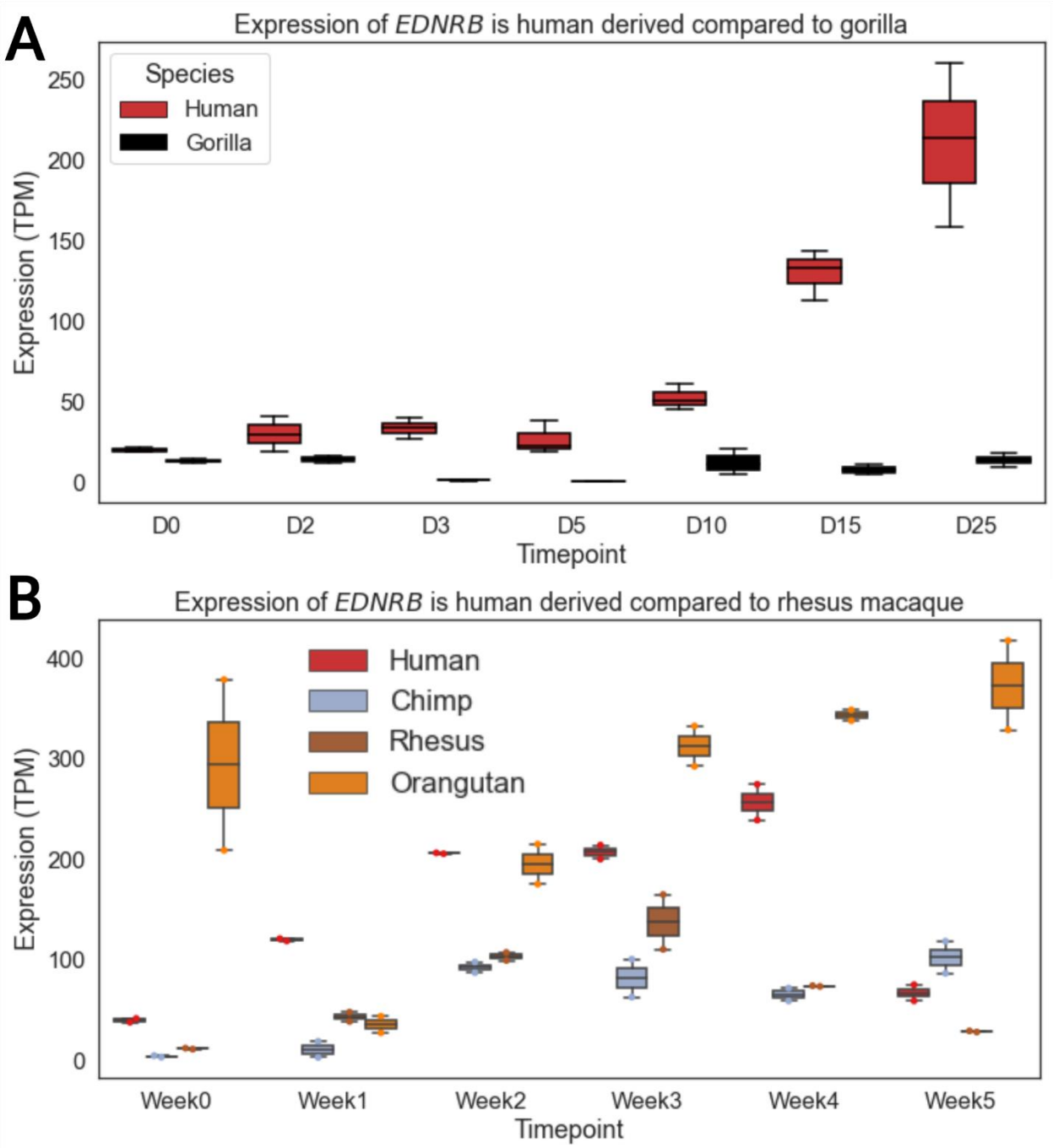


Supplementary Figure 2: A) Correlation between Fetal Cortex Mann Whitney p-value rank and GTEx Mann Whitney p-value rank for genes present in both ranked lists. Spearman's rho = 0.92, $p < 10^{-300}$. See Methods for ranking details.


A
RNA pol II regulatory region
sequence-specific DNA binding
Enrichment score
0.4
0.2
0.0
FDR = 0.016
0
2000
4000
6000
8000
10000
Rank in ordered dataset
B
Expression of CUX1 in hybrid cortical spheroids
Expression (TPM)
70
60
50
40
30
20
Human
Chimp
D50
D100
D150
D200
Timepoint
C
Expression of CUX1 in parental cortical spheroids
Human
Chimp
Expression (TPM)
120
100
80
60
40
D15
D25
D35
D50/D52
D100/D102
D151/D155
D200/D201
Timepoint
D
CUX1 D150 GTEx comparison
Sample count
10^3
10^2
10^1
10^0
−2
−1
0
1
log2(ASE ratio)
Human/chimp ASE
Human population ASE
E
CUX1 D150 fetal cortex comparison
Sample count
10^1
10^0
−1.0
−0.5
0.0
0.5
1.0
1.5
log2(ASE ratio)
Human/chimp ASE
Human fetal population ASE

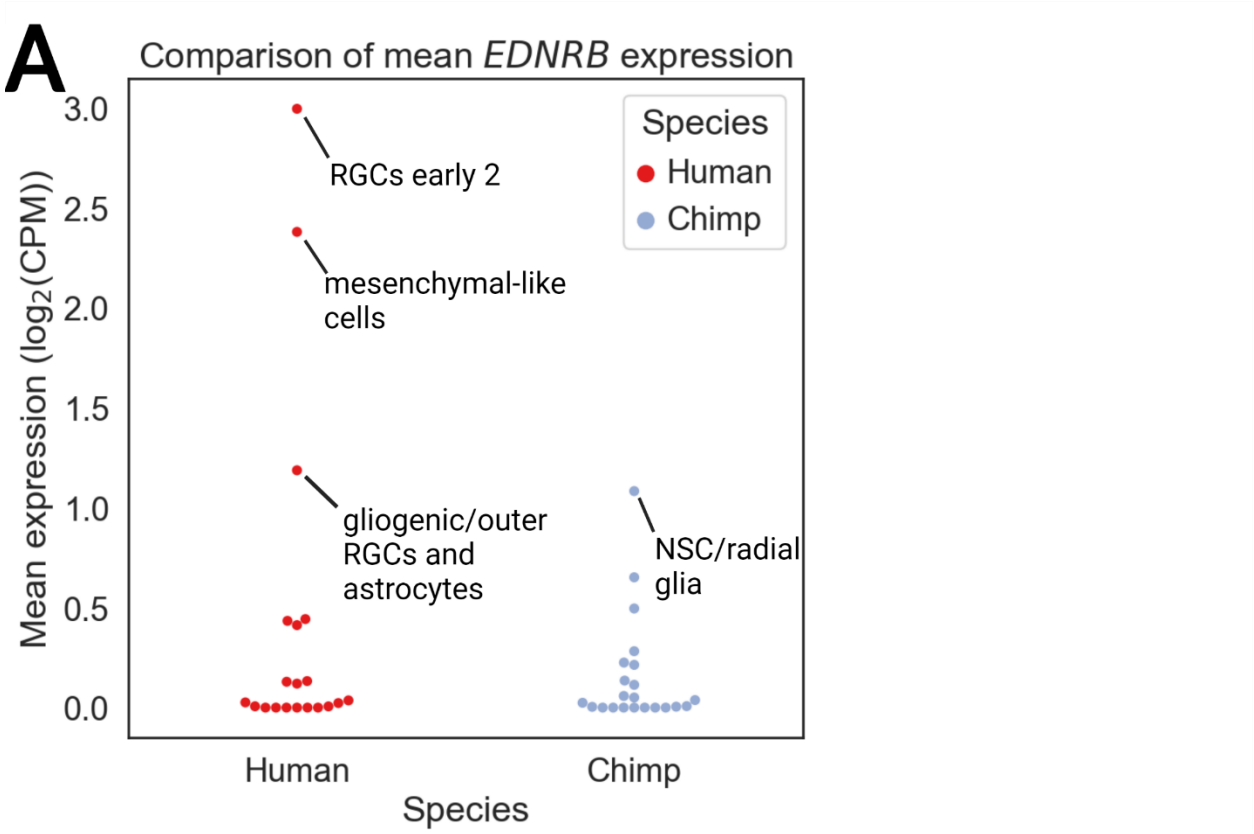
515

Supplementary Figure 3: A) Summary of RNA polymerase II regulatory region sequence-specific DNA binding. Each blue line represents a gene in the gene set and the green curve is the cumulative enrichment score. Genes in the gene set are enriched at the top of the list. **B)** Expression of CUX1 in human/chimpanzee hybrid cortical spheroids (CS). Expression is shown in transcripts per million (TPM). Expression from the chimpanzee allele is consistently higher than expression from the human allele. **C)** Expression of CUX1 in human and chimpanzee parental cortical spheroids (CS). Expression is shown in transcripts per million (TPM). **D)** Comparison of the GTEx ASE distribution for CUX1 to the human/chimpanzee hybrid ASE distribution. **E)** Comparison of the fetal cortex ASE distribution for CUX1 to the human/chimpanzee ASE distribution. The human/chimpanzee hybrid ASE values lie well outside the human fetal cortex distribution.

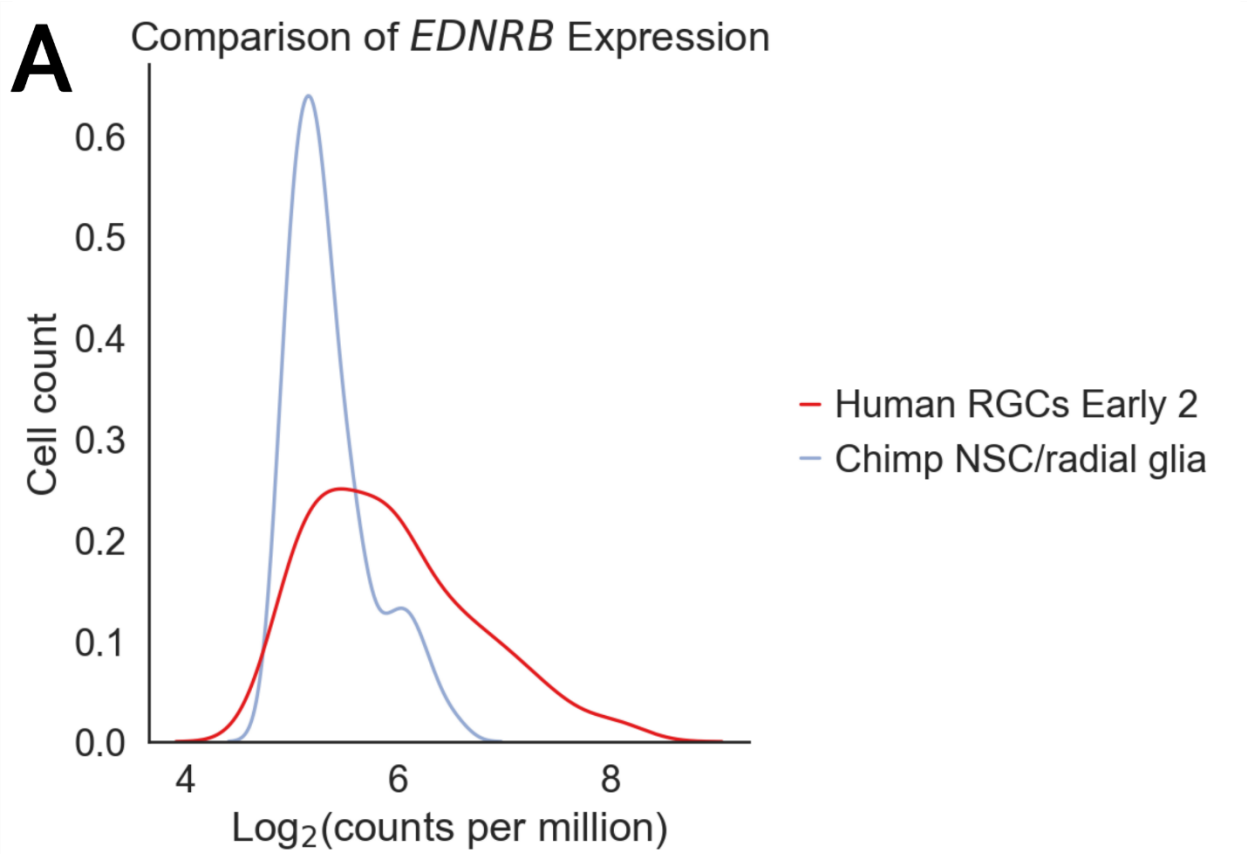


Supplementary Figure 4: A) Comparison of EDNRB expression between early-stage human and gorilla cerebral organoids (Benito-Kwiecinski et al. 2021). Human expression is considerably higher than gorilla expression across timepoints indicating that EDRNB is upregulated in the human lineage (as opposed to downregulated in the chimpanzee lineage). **B)** Comparison of EDNRB expression between early-stage human, chimpanzee, orangutan, and

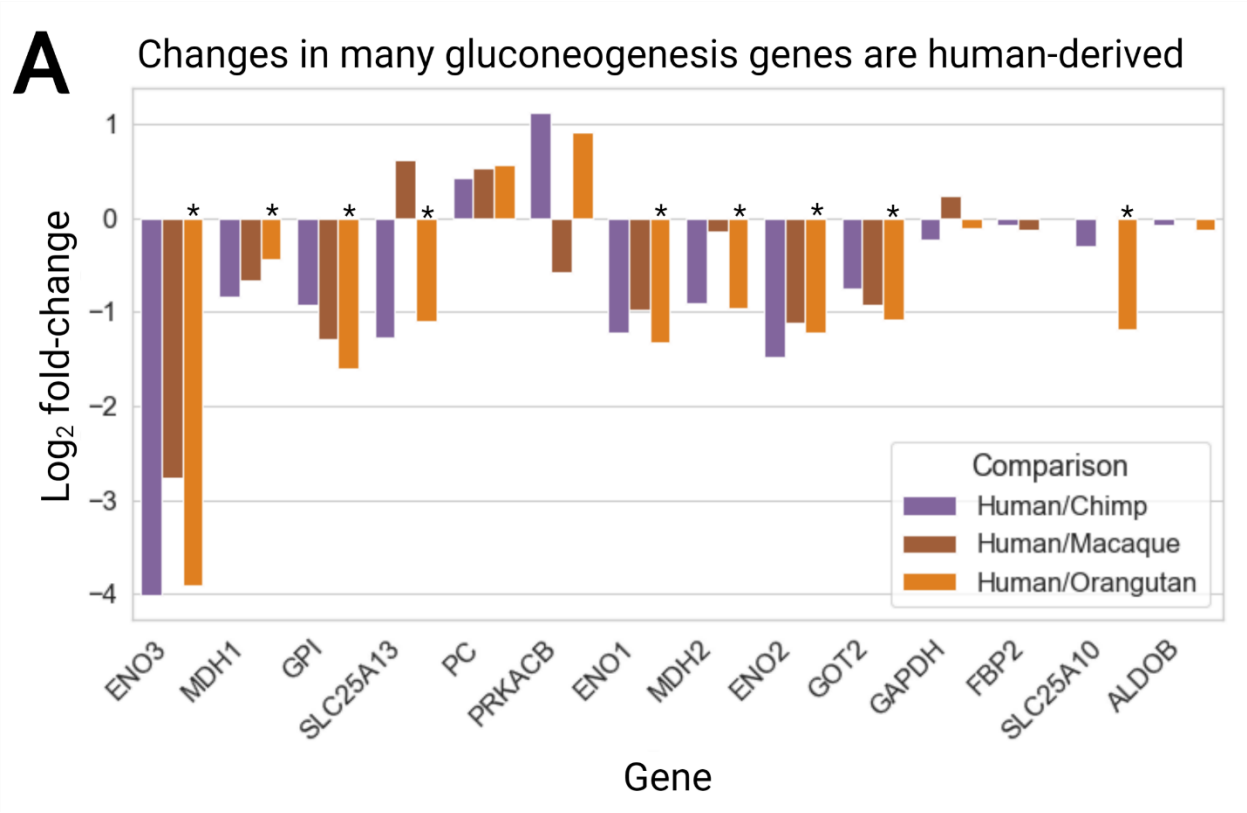
rhesus macaque cerebral organoids (Field et al. 2019). Human expression is considerably higher than rhesus macaque expression across timepoints. However, orangutan EDNRB expression is high as well, indicating an independent increase in expression in the orangutan lineage.



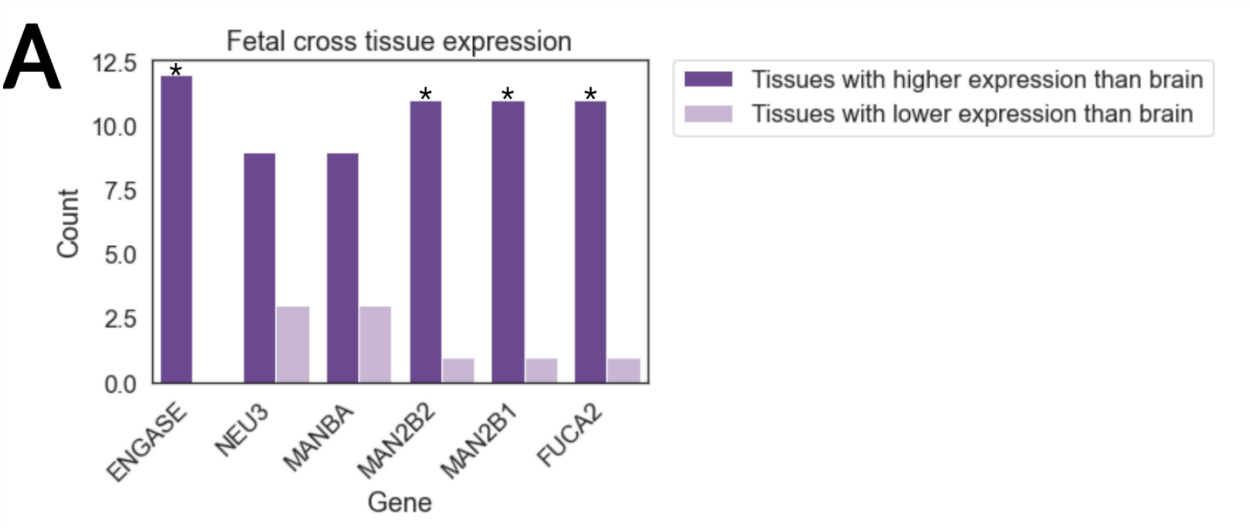
Supplementary Figure 5: A) Plot showing mean expression of EDNRB across chimpanzee and human cell clusters (Kanton et al. 2019). Expression in the "RGCs early 2" cluster is significantly higher than in all chimpanzee clusters by Mann-Whitney U Test ($p < 10^{-16}$ for all comparisons).



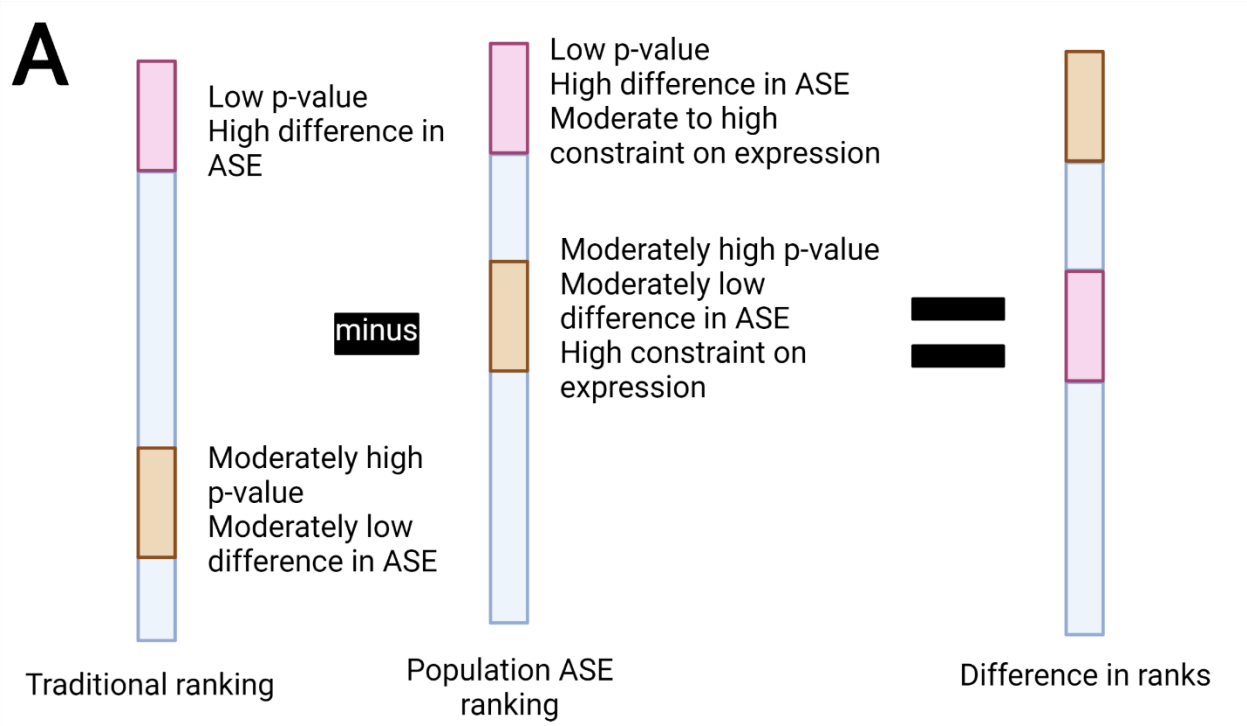
Supplementary Figure 6: The distribution of EDNRB expression in cells with non-zero counts (Kanton et al. 2019). As there were many more chimpanzee cells in the Chimp NSC/radial glia category, we down-sampled the number of cells so that there were an equal number of EDNRB-expressing cells. The distribution in chimpanzee NSC/radial glia is much more shifted left than that of the human “RGCs Early 2” cluster indicating higher expression in the human cluster.



Supplementary Figure 7: A) Barplot showing that many changes in the expression of gluconeogenesis genes are human-derived and occur in parental organoids. Data are from Week 5 cerebral organoids (Field et al. 2019). Asterisks indicate genes for which the human-orangutan difference is significant at an FDR cutoff of 0.05 and with lower expression in human. Genes whose down-regulation is not human-derived generally show insignificant differences in expression between humans and chimpanzees potentially indicating compensatory trans-acting genetic changes.



Supplementary Figure 8: A) Barplot showing that expression of the genes driving the "Other glycan degradation" enrichment is generally lower in fetal brain than in other tissues (Cao et al. 2020). Asterisks indicate significance (p-value < 0.05) by binomial test with expected p = 0.5.



Supplementary Figure 9: A) Explanation of computation of difference in ranks. Population ASE Ranking was subtracted from the traditional ranking. Due to this, genes that are highly ranked in both the traditional and population ASE rankings (likely those with low p-values and high difference in ASE) are near the middle of the list in the difference of ranks. On the other hand, genes with moderately high p-values and high constraint on expression are ranked lowly in the traditional ranking, but near the middle in the population ASE ranking, and so are very high in the difference in ranks.

Adams, Katrina L., Giulia Riparini, Payal Banerjee, Marjolein Breur, Marianna Bugiani, and Vittorio Gallo. 2020. “Endothelin-1 Signaling Maintains Glial Progenitor Proliferation in the Postnatal Subventricular Zone.” *Nature Communications* 11 (1): 2138. https://doi.org/10.1038/s41467-020-16028-8.

Agoglia, Rachel M., Danqiong Sun, Fikri Birey, Se-Jin Yoon, Yuki Miura, Karen Sabatini, Sergiu P. Pașca, and Hunter B. Fraser. 2021. “Primate Cell Fusion Disentangles Gene Regulatory Divergence in Neurodevelopment.” *Nature* 592 (7854): 421–27. https://doi.org/10.1038/s41586-021-03343-3.

Alkhayat, A. H., S. A. Kraemer, J. R. Leipprandt, M. Macek, W. J. Kleijer, and K. H. Friderici. 1998. “Human -Mannosidase CDNA Characterization and First Identification of a Mutation Associated with Human -Mannosidosis.” *Human Molecular Genetics* 7 (1): 75–83. https://doi.org/10.1093/hmg/7.1.75.

Aygün, Nil, Angela L. Elwell, Dan Liang, Michael J. Lafferty, Kerry E. Cheek, Kenan P. Courtney, Jessica Mory, et al. 2021. “Brain-Trait-Associated Variants Impact Cell-Type-Specific Gene Regulation during Neurogenesis.” *The American Journal of Human Genetics* 108 (9): 1647–68. https://doi.org/10.1016/j.ajhg.2021.07.011.

Benito-Kwiecinski, Silvia, Stefano L. Giandomenico, Magdalena Sutcliffe, Erlend S. Riis, Paula Freire-Pritchett, Iva Kelava, Stephanie Wunderlich, et al. 2021. “An Early Cell Shape Transition Drives Evolutionary Expansion of the Human Forebrain.” *Cell* 184 (8): 2084-2102.e19. https://doi.org/10.1016/j.cell.2021.02.050.

Benjamini, Yoav, and Yosef Hochberg. 1995. “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.” *Journal of the Royal Statistical Society: Series B (Methodological)* 57 (1): 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Blekhman, Ran, Alicia Oshlack, Adrien E. Chabot, Gordon K. Smyth, and Yoav Gilad. 2008. “Gene Regulation in Primates Evolves under Tissue-Specific Selection Pressures.” *PLoS Genetics* 4 (11): e1000271. https://doi.org/10.1371/journal.pgen.1000271.

Blomqvist, Maria, Marie Falkenberg Smeland, Julia Lindgren, Per Sikora, Hilde Monica Frostad Riise Stensland, and Jorge Asin-Cayuela. 2019. “β-Mannosidosis Caused by a Novel Homozygous Intragenic Inverted Duplication in *MANBA*.” *Molecular Case Studies* 5 (3): a003954. https://doi.org/10.1101/mcs.a003954.

Broad Institute. 2019. “Picard Toolkit.”

Cao, Junyue, Diana R. O’Day, Hannah A. Pliner, Paul D. Kingsley, Mei Deng, Riza M. Daza, Michael A. Zager, et al. 2020. “A Human Cell Atlas of Fetal Gene Expression.” *Science* 370 (6518). https://doi.org/10.1126/science.aba7721.

Carbon, Seth, Amelia Ireland, Christopher J. Mungall, ShengQiang Shu, Brad Marshall, and Suzanna Lewis. 2009. “AmiGO: Online Access to Ontology and Annotation Data.” *Bioinformatics* 25 (2): 288–89. https://doi.org/10.1093/bioinformatics/btn615.

Castel, Stephane E., François Aguet, Pejman Mohammadi, François Aguet, Shankara Anand, Kristin G. Ardlie, Stacey Gabriel, et al. 2020. “A Vast Resource of Allelic Expression Data Spanning Human Tissues.” *Genome Biology* 21 (1): 234. https://doi.org/10.1186/s13059-020-02122-z.

Cathey, Sara S., Sara M. Sarasua, Richard Simensen, Katie Pietris, Gordon Kimbrell, David Sillence, Callum Wilson, and Lucia Horowitz. 2019. “Intellectual Functioning in Alpha-mannosidosis.” *JIMD Reports* 50 (1): 44–49. https://doi.org/10.1002/jmd2.12073.

Clark, H. M., J. R. Duffy, J. L. Whitwell, J. E. Ahlskog, E. J. Sorenson, and K. A. Josephs. 2014. “Clinical and Imaging Characterization of Progressive Spastic Dysarthria.” *European Journal of Neurology* 21 (3): 368–76. https://doi.org/10.1111/ene.12271.

Collins, Ryan L, Joseph T Glessner, Eleonora Porcu, Lisa-Marie Niestroj, Jacob Ulirsch, Georgios Kellaris, Daniel P Howrigan, et al. medRxiv “A Cross-Disorder Dosage Sensitivity Map of the Human Genome.” *Genomic Medicine Institute* 15 (2): 23–25. https://doi.org/10.1101/2021.01.26.21250098.

Combs, Peter A., Joshua J. Krupp, Neil M. Khosla, Dennis Bua, Dmitri A. Petrov, Joel D. Levine, and Hunter B. Fraser. 2018. “Tissue-Specific Cis-Regulatory Divergence Implicates EloF in Inhibiting Interspecies Mating in Drosophila.” *Current Biology* 28 (24): 3969-3975.e3. https://doi.org/10.1016/j.cub.2018.10.036.

Doan, Ryan N., Byoung-Il Bae, Beatriz Cubelos, Cindy Chang, Amer A. Hossain, Samira Al-Saad, Nahit M. Mukaddes, et al. 2016. “Mutations in Human Accelerated Regions Disrupt Cognition and Social Behavior.” *Cell* 167 (2): 341-354.e12. https://doi.org/10.1016/j.cell.2016.08.071.

Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. “STAR: Ultrafast Universal RNA-Seq Aligner.” *Bioinformatics* 29 (1): 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Eichmüller, Oliver L., Nina S. Corsini, Ábel Vértesy, Ilaria Morassut, Theresa Scholl, Victoria-Elisabeth Gruber, Angela M. Peer, et al. 2022. “Amplification of Human Interneuron Progenitors Promotes Brain Tumors and Neurological Defects.” *Science* 375 (6579). https://doi.org/10.1126/science.abf5546.

Ferraro, Nicole M., Benjamin J. Strober, Jonah Einson, Nathan S. Abell, Francois Aguet, Alvaro N. Barbeira, Margot Brandt, et al. 2020. “Transcriptomic Signatures across Human Tissues Identify Functional Rare Genetic Variation.” *Science* 369 (6509): eaaz5900. https://doi.org/10.1126/science.aaz5900.

Field, Andrew R., Frank M.J. Jacobs, Ian T. Fiddes, Alex P.R. Phillips, Andrea M. Reyes-Ortiz, Erin LaMontagne, Lila Whitehead, et al. 2019. “Structurally Conserved Primate LncRNAs Are Transiently Expressed during Human Cortical Differentiation and Influence Cell-Type-Specific Genes.” *Stem Cell Reports* 12 (2): 245–57. https://doi.org/10.1016/j.stemcr.2018.12.006.

Fraser, Hunter B. 2011. “Genome-Wide Approaches to the Study of Adaptive Gene Expression Evolution.” *BioEssays* 33 (6): 469–77. https://doi.org/10.1002/bies.201000094.

Fraser, Hunter B. 2013. “Gene Expression Drives Local Adaptation in Humans.” *Genome Research* 23 (7): 1089–96. https://doi.org/10.1101/gr.152710.112.

Gilad, Yoav, Alicia Oshlack, Gordon K. Smyth, Terence P. Speed, and Kevin P. White. 2006. "Expression Profiling in Primates Reveals a Rapid Evolution of Human Transcription Factors." *Nature* 440 (7081): 242–45. https://doi.org/10.1038/nature04559.

Girskis, Kelly M., Andrew B. Stergachis, Ellen M. DeGennaro, Ryan N. Doan, Xuyu Qian, Matthew B. Johnson, Peter P. Wang, et al. 2021. "Rewiring of Human Neurodevelopmental Gene Regulatory Programs by Human Accelerated Regions." *Neuron* 109 (20): 3239-3251.e7. https://doi.org/10.1016/j.neuron.2021.08.005.

Givanna H Putri, Simon Anders, Paul Theodor Pyl, John E Pimanda, Fabio Zanini. 2021. "Analysing High-Throughput Sequencing Data in Python with HTSeq 2.0." *ArXiv*.

Gkini, Vasiliki, and Takashi Namba. 2022. "Glutaminolysis and the Control of Neural Progenitors in Neocortical Development and Evolution." *The Neuroscientist*, January, 107385842110690. https://doi.org/10.1177/10738584211069060.

Gokhman, David, Rachel M. Agoglia, Maia Kinnebrew, Wei Gordon, Danqiong Sun, Vivek K. Bajpai, Sahin Naqvi, et al. 2021. "Human–Chimpanzee Fused Cells Reveal Cis-Regulatory Divergence Underlying Skeletal Evolution." *Nature Genetics* 53 (4): 467–76. https://doi.org/10.1038/s41588-021-00804-3.

Housman, Genevieve, and Yoav Gilad. 2020. "Prime Time for Primate Functional Genomics." *Current Opinion in Genetics & Development* 62 (June): 1–7. https://doi.org/10.1016/j.gde.2020.04.007.

Hu, Caroline K., Ryan A. York, Hillery C. Metz, Nicole L. Bedford, Hunter B. Fraser, and Hopi E. Hoekstra. 2022. "Cis-Regulatory Changes in Locomotor Genes Are Associated with the Evolution of Burrowing Behavior." *Cell Reports* 38 (7): 110360. https://doi.org/10.1016/j.celrep.2022.110360.

Jassal, Bijay, Lisa Matthews, Guilherme Viteri, Chuqiao Gong, Pascual Lorente, Antonio Fabregat, Konstantinos Sidiropoulos, et al. 2019. "The Reactome Pathway Knowledgebase." *Nucleic Acids Research*, November. https://doi.org/10.1093/nar/gkz1031.

Kanehisa, Minoru, Yoko Sato, Masayuki Kawashima, Miho Furumichi, and Mao Tanabe. 2016. "KEGG as a Reference Resource for Gene and Protein Annotation." *Nucleic Acids Research* 44 (D1): D457–62. https://doi.org/10.1093/nar/gkv1070.

Kanton, Sabina, Michael James Boyle, Zhisong He, Malgorzata Santel, Anne Weigert, Fátima Sanchís-Calleja, Patricia Guijarro, et al. 2019. "Organoid Single-Cell Genomic Atlas Uncovers Human-Specific Features of Brain Development." *Nature* 574 (7778): 418–22. https://doi.org/10.1038/s41586-019-1654-9.

Kelley, Joanna L., and Yoav Gilad. 2020. "Effective Study Design for Comparative Functional Genomics." *Nature Reviews Genetics* 21 (7): 385–86. https://doi.org/10.1038/s41576-020-0242-z.

Khrameeva, Ekaterina, Ilia Kurochkin, Dingding Han, Patricia Guijarro, Sabina Kanton, Malgorzata Santel, Zhengzong Qian, et al. 2020. "Single-Cell-Resolution Transcriptome Map of Human, Chimpanzee, Bonobo, and Macaque Brains." *Genome Research* 30 (5): 776–89. https://doi.org/10.1101/gr.256958.119.

King, Mary-Claire, and A. C. Wilson. 1975. “Evolution at Two Levels in Humans and Chimpanzees.” *Science* 188 (4184): 107–16. https://doi.org/10.1126/science.1090005.

Kita, Ryosuke, Sandeep Venkataram, Yiqi Zhou, and Hunter B. Fraser. 2017. “High-Resolution Mapping of *Cis* -Regulatory Variation in Budding Yeast.” *Proceedings of the National Academy of Sciences* 114 (50): E10736–44. https://doi.org/10.1073/pnas.1717421114.

Köhler, Sebastian, Sandra C. Doelken, Christopher J. Mungall, Sebastian Bauer, Helen v. Firth, Isabelle Bailleul-Forestier, Graeme C. M. Black, et al. 2014. “The Human Phenotype Ontology Project: Linking Molecular Biology and Disease through Phenotype Data.” *Nucleic Acids Research* 42 (D1): D966–74. https://doi.org/10.1093/nar/gkt1026.

Köhler, Sebastian, Michael Gargano, Nicolas Matentzoglu, Leigh C Carmody, David Lewis-Smith, Nicole A Vasilevsky, Daniel Danis, et al. 2021. “The Human Phenotype Ontology in 2021.” *Nucleic Acids Research* 49 (D1): D1207–17. https://doi.org/10.1093/nar/gkaa1043.

Li, Bo, Victor Ruotti, Ron M. Stewart, James A. Thomson, and Colin N. Dewey. 2010. “RNA-Seq Gene Expression Estimation with Read Mapping Uncertainty.” *Bioinformatics* 26 (4): 493–500. https://doi.org/10.1093/bioinformatics/btp692.

Love, Michael I, Wolfgang Huber, and Simon Anders. 2014. “Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2.” *Genome Biology* 15 (12): 550. https://doi.org/10.1186/s13059-014-0550-8.

Mack, Katya L., Polly Campbell, and Michael W. Nachman. 2016. “Gene Regulation and Speciation in House Mice.” *Genome Research* 26 (4): 451–61. https://doi.org/10.1101/gr.195743.115.

Paciorkowski, Alex R., Ryan N. Traylor, Jill A. Rosenfeld, Jacqueline M. Hoover, Catharine J. Harris, Susan Winter, Yves Lacassie, et al. 2013. “MEF2C Haploinsufficiency Features Consistent Hyperkinesis, Variable Epilepsy, and Has a Role in Dorsal and Ventral Neuronal Developmental Pathways.” *Neurogenetics* 14 (2): 99–111. https://doi.org/10.1007/s10048-013-0356-y.

Paşca, Anca M, Steven A Sloan, Laura E Clarke, Yuan Tian, Christopher D Makinson, Nina Huber, Chul Hoon Kim, et al. 2015. “Functional Cortical Neurons and Astrocytes from Human Pluripotent Stem Cells in 3D Culture.” *Nature Methods* 12 (7): 671–78. https://doi.org/10.1038/nmeth.3415.

Platzer, Konrad, Benjamin Cogné, Jennifer Hague, Carlo L. Marcelis, Diana Mitter, Katrin Oberndorff, Soo-Mi Park, et al. 2018. “Haploinsufficiency of *CUX1* Causes Nonsyndromic Global Developmental Delay With Possible Catch-up Development.” *Annals of Neurology* 84 (2): 200–207. https://doi.org/10.1002/ana.25278.

Prud’homme, B., N. Gompel, and S. B. Carroll. 2007. “Emerging Principles of Regulatory Evolution.” *Proceedings of the National Academy of Sciences* 104 (Supplement 1): 8605–12. https://doi.org/10.1073/pnas.0700488104.

Reilly, Steven K., and James P. Noonan. 2016. “Evolution of Gene Regulation in Humans.” *Annual Review of Genomics and Human Genetics* 17 (1): 45–67. https://doi.org/10.1146/annurev-genom-090314-045935.

Romero, Irene Gallego, Ilya Ruvinsky, and Yoav Gilad. 2012. “Comparative Studies of Gene Expression and the Evolution of Gene Regulation.” *Nature Reviews Genetics* 13 (7): 505–16. https://doi.org/10.1038/nrg3229.

Roy, Sushmita, Ilan Wapinski, Jenna Pfiffner, Courtney French, Amanda Socha, Jay Konieczka, Naomi Habib, Manolis Kellis, Dawn Thompson, and Aviv Regev. 2013. “Arboretum: Reconstruction and Analysis of the Evolutionary History of Condition-Specific Transcriptional Modules.” *Genome Research* 23 (6): 1039–50. https://doi.org/10.1101/gr.146233.112.

Runge, Karen, Rémi Mathieu, Stéphane Bugeon, Sahra Lafi, Corinne Beurrier, Surajit Sahu, Fabienne Schaller, et al. 2021. “Disruption of NEUROD2 Causes a Neurodevelopmental Syndrome with Autistic Features via Cell-Autonomous Defects in Forebrain Glutamatergic Neurons.” *Molecular Psychiatry* 26 (11): 6125–48. https://doi.org/10.1038/s41380-021-01179-x.

Shinohara, Haruo, Jun Udagawa, Rika Morishita, Hiroshi Ueda, Hiroki Otani, Reiji Semba, Kanefusa Kato, and Tomiko Asano. 2004. “Gi2 Signaling Enhances Proliferation of Neural Progenitor Cells in the Developing Brain.” *Journal of Biological Chemistry* 279 (39): 41141–48. https://doi.org/10.1074/jbc.M406721200.

Subramanian, A., P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, et al. 2005. “Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles.” *Proceedings of the National Academy of Sciences* 102 (43): 15545–50. https://doi.org/10.1073/pnas.0506580102.

Supek, Fran, Matko Bošnjak, Nives Škunca, and Tomislav Šmuc. 2011. “REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms.” *PLoS ONE* 6 (7): e21800. https://doi.org/10.1371/journal.pone.0021800.

Vidovic, Maria, Ming-Ming Chen, Qun-Ying Lu, Katherine F. Kalloniatis, Ben M. Martin, Abel H. Y. Tan, Celina Lynch, G. David H. Croaker, Daniel T. Cass, and Zan-Min Song. 2008. “Deficiency in Endothelin Receptor B Reduces Proliferation of Neuronal Progenitors and Increases Apoptosis in Postnatal Rat Cerebellum.” *Cellular and Molecular Neurobiology* 28 (8): 1129–38. https://doi.org/10.1007/s10571-008-9292-z.

Williams, Sarah E., Maxence Noel, Sylvain Lehoux, Murat Cetinbas, Ramnik J. Xavier, Ruslan I. Sadreyev, Edward M. Scolnick, Jordan W. Smoller, Richard D. Cummings, and Robert G. Mealer. 2022. “Mammalian Brain Glycoproteins Exhibit Diminished Glycan Complexity Compared to Other Tissues.” *Nature Communications* 13 (1): 275. https://doi.org/10.1038/s41467-021-27781-9.

Wittkopp, Patricia J., and Gizem Kalay. 2012. “Cis-Regulatory Elements: Molecular Mechanisms and Evolutionary Processes Underlying Divergence.” *Nature Reviews Genetics* 13 (1): 59–69. https://doi.org/10.1038/nrg3095.

Wolf, F. Alexander, Philipp Angerer, and Fabian J. Theis. 2018. “SCANPY: Large-Scale Single-Cell Gene Expression Data Analysis.” *Genome Biology* 19 (1): 15. https://doi.org/10.1186/s13059-017-1382-0.

Zhang, Hong, Fei Zhang, Yiming Yu, Li Feng, Jinbu Jia, Bo Liu, Bosheng Li, Hongwei Guo, and Jixian Zhai. 2020. “A Comprehensive Online Database for Exploring ~20,000 Public Arabidopsis RNA-Seq Libraries.” *Molecular Plant* 13 (9): 1231–33. https://doi.org/10.1016/j.molp.2020.08.001.

Zhang, Xu, and Justin O Borevitz. 2009. “Global Analysis of Allele-Specific Expression in *Arabidopsis Thaliana*.” *Genetics* 182 (4): 943–54. https://doi.org/10.1534/genetics.109.103499.

Zhu, Anqi, Joseph G Ibrahim, and Michael I Love. 2019. “Heavy-Tailed Prior Distributions for Sequence Count Data: Removing the Noise and Preserving Large Differences.” *Bioinformatics* 35 (12): 2084–92. https://doi.org/10.1093/bioinformatics/bty895.

Zhu, Ying, André M. M. Sousa, Tianliuyun Gao, Mario Skarica, Mingfeng Li, Gabriel Santpere, Paula Esteller-Cucala, et al. 2018a. “Spatiotemporal Transcriptomic Divergence across Human and Macaque Brain Development.” *Science* 362 (6420). https://doi.org/10.1126/science.aat8077.