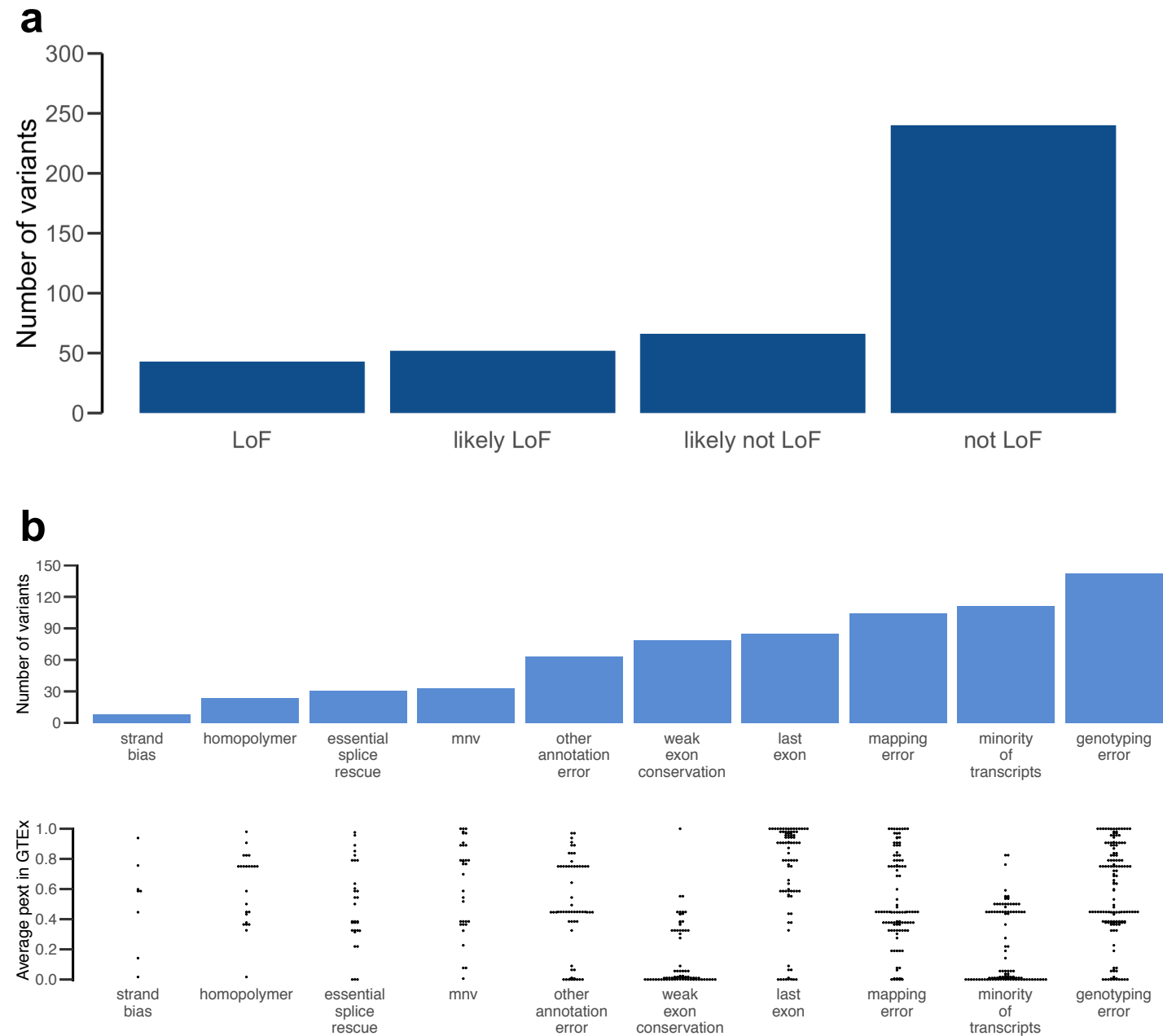
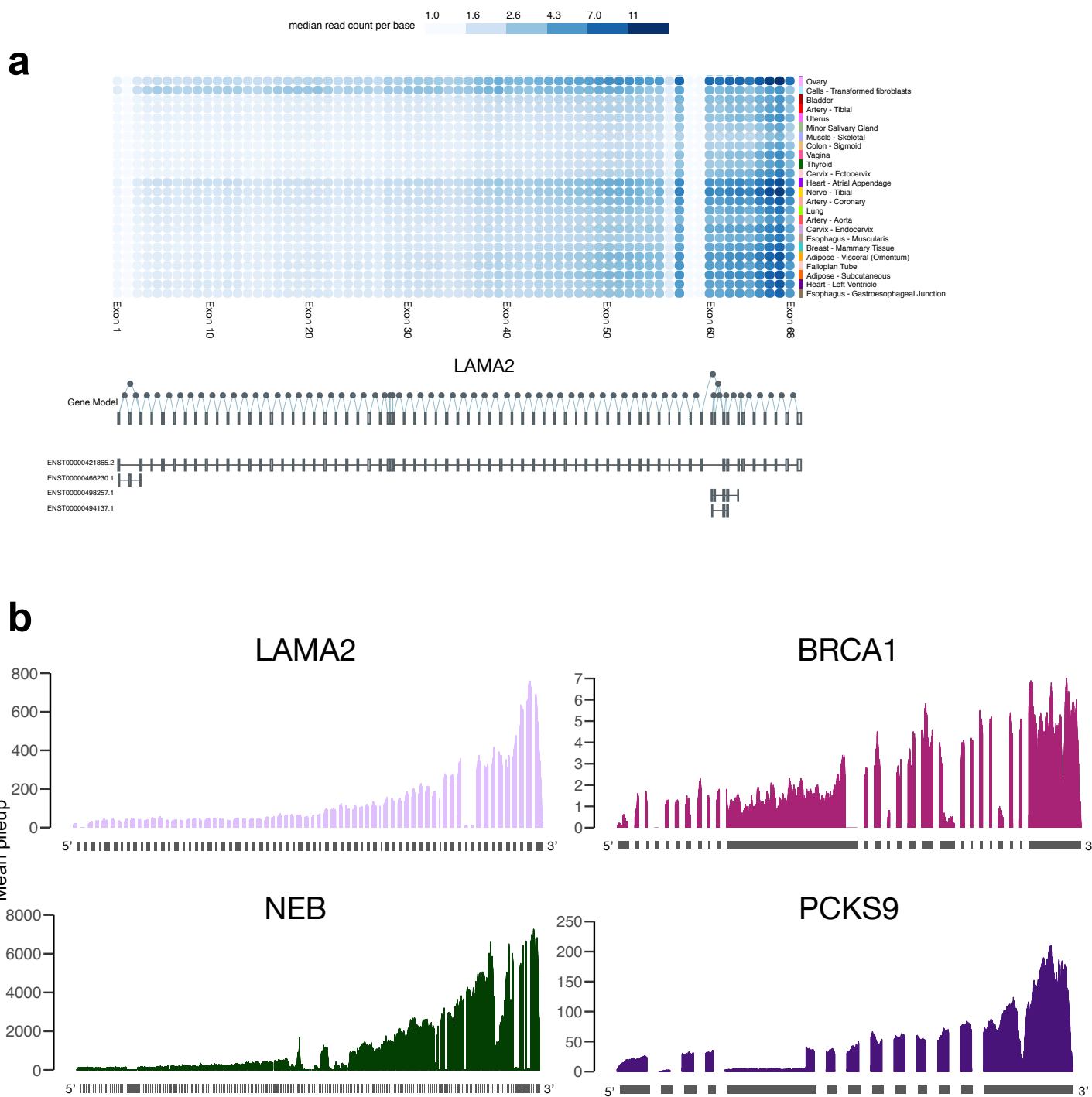

Supplementary information

Transcript expression-aware annotation improves rare variant interpretation

In the format provided by the
authors and unedited



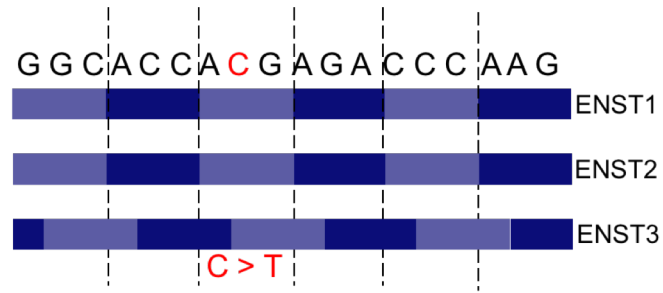
Supplementary Figure 1: Details of manual curation of 401 pLoF variants in 61 haploinsufficient developmental disease genes **a.** Distribution of curation verdicts for the 401 pLoF variants. We categorized 240 variants (76%) as not being LoF, 66 as likely not LoF, 52 as likely LoF and 43 as LoF. **b.** Full distribution of the flags refuting true LoF status for 306 not LoF and likely not LoF variants (top) and their corresponding pext score in GTEx (bottom). A variant with multiple flags is assigned to each flag as in Figure 1 (ie. double counted). Minority of transcripts and weak exon conservation were grouped as transcript errors, genotyping errors and homopolymers grouped as sequencing errors, essential splice rescue and MNVs grouped as rescue and strand bias was included in other annotation errors. While the pext values are randomly distributed for other error modes, they are enriched for lower values in transcript errors. Criteria for curation for each flag and verdict, and full curation results are available in Supplementary Tables 1, 2, and 3, respectively.



Supplementary Figure 2: Technical artifacts in transcriptome sequencing experiments prevent the use of read pileup at exons as an unbiased proxy for expression **a.** Example of exon expression information on the GTEx web browser (gtexportal.org) for *LAMA2*, which has 3 annotated transcripts. Blue-gray gradient represents median read count per base. While exons 5 and 55 are annotated on a single transcript, the mean read count for the exons are 0.8 and 3.25 in GTEx ovary, respectively, reflecting the confounding effect of 3' bias **b.** Examples of 3' bias in genes of varying lengths and expression levels shows 3' bias is pervasive. Base-level coverage of uniquely mapped reads were calculated in 10 random GTEx samples per tissue using samtools depth in tissues where the genes are highly expressed. Plots show (1) *LAMA2* in tibial nerve (2) *BRCA1* in mammary tissue (breast) (3) *NEB* in skeletal muscle and (4) *PCKS9* in liver, all of which display 3' bias.

a**1 – Variant table or VCF**

CHROM	POS	REF	ALT	CONSEQUENCES
X	34242	C	T	ENST1: missense, ENST2: missense, ENST3: stop_gained

**2 – Isoform expression matrix**

	Heart	Liver	Lung
ENST1	10	0.5	6
ENST2	5	0.5	18
ENST3	0	0.5	3

Base level expression for chrX:34242

→ sum expression(ENST1, ENST2, ENST3)

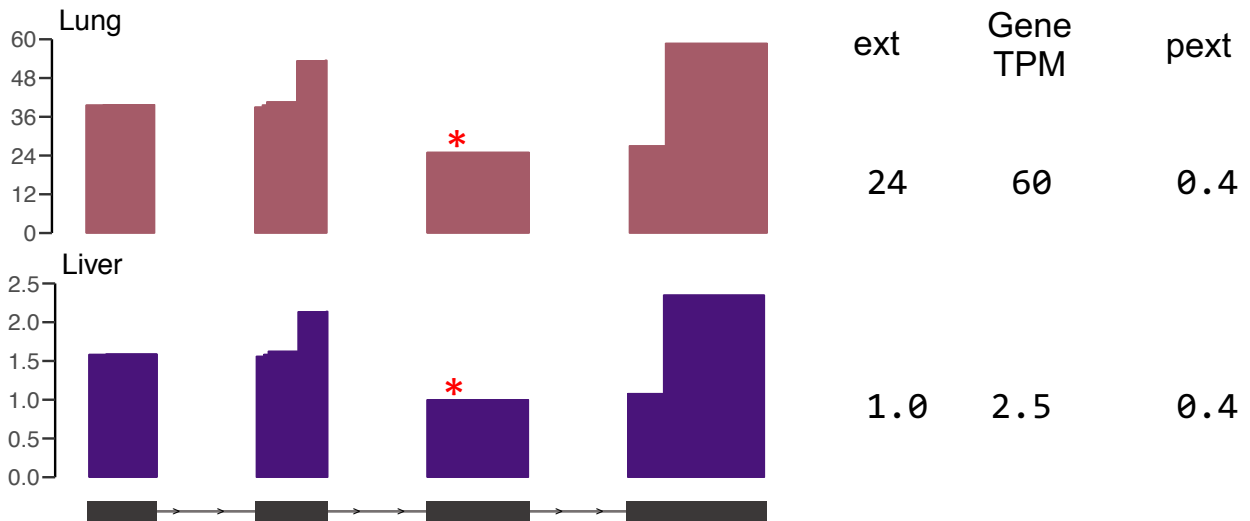
Annotation-level expression across transcripts for chrX:34242-C-T

→ missense : sum expression(ENST1 & ENST2)

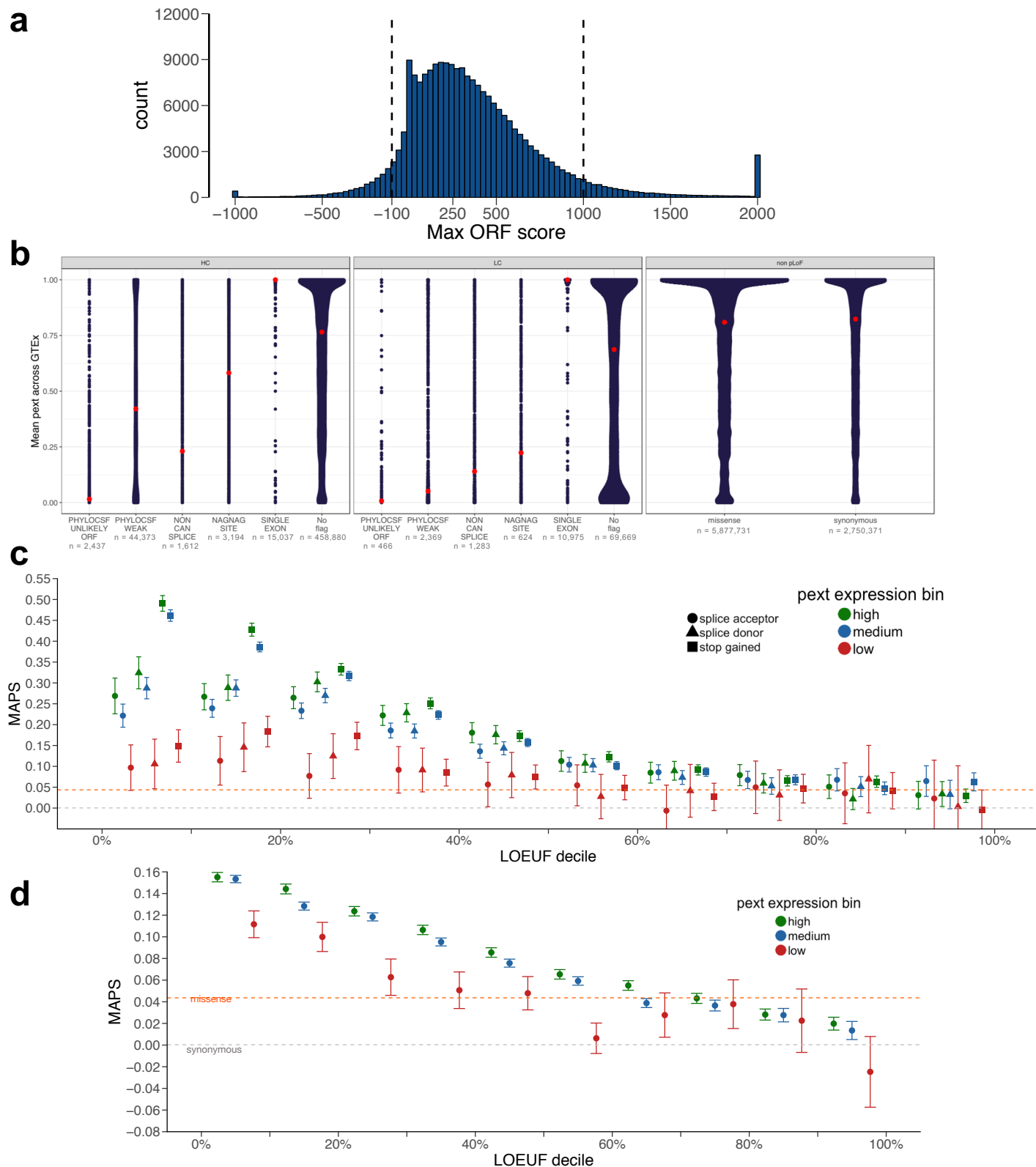
→ stop gained : expression(ENST3)

3 – Resulting annotated variant table or VCF

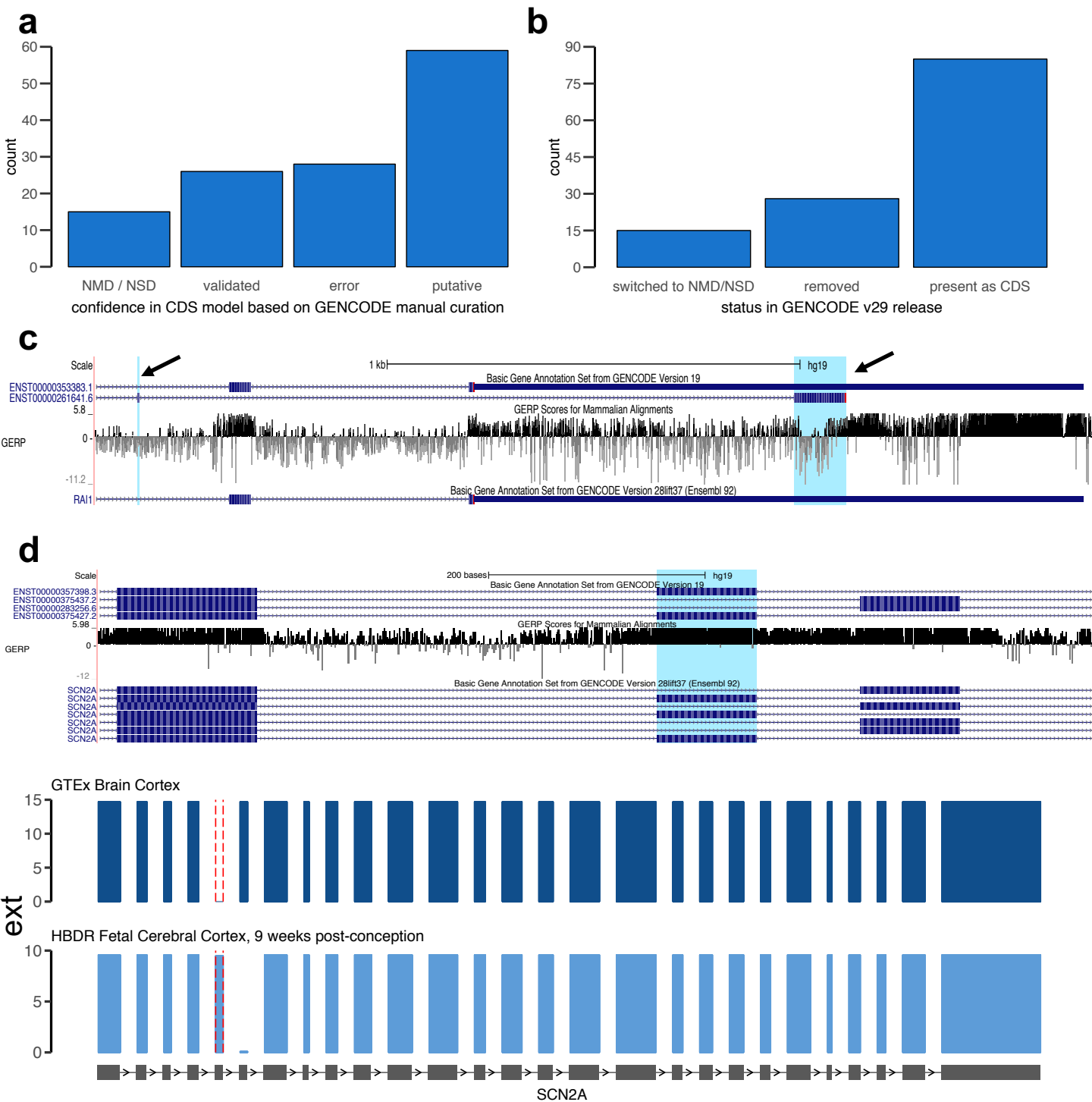
CHROM	POS	REF	ALT	CONSEQUENCES	BASE_LEVEL	ANNOTATION_LEVEL (ext)
X	34242	C	T	ENST1: missense; ENST2: missense; ENST3: stop_gained	Heart: 15, Liver: 1.5, Lung : 27	missense: Heart: 15, Liver: 1, Lung : 24 stop_gained: Heart: 0, Liver: 0.5, Lung : 3

b

Supplementary Figure 3: Details of calculating transcript-expression annotation. a. A SNV can have different consequences across transcripts. For example, an SNV on a region with three annotated transcripts can have a missense effect on two transcripts and a nonsense effect on one transcript. The base-level expression, mainly to be utilized for quick visualization of variant expression in genes, is calculated as the sum of the three transcripts. The annotation level expression across transcript (ext) metric defines the expression of a variant as the sum of the expression of transcripts on which an annotation exists. In this example, the expression value for the missense variant will be the sum of the expression on transcripts where the variant is a missense (ENST1 and ENST2) and the value for the nonsense will be the sum of the expression of transcripts where the variant is a nonsense (ENST3). **b.** To account for gene expression differences between tissues, we normalize the ext value by the sum of expression of all transcript (representing gene expression) in the tissue to calculate the proportion expressed across transcript (pext) value used in the manuscript. This allows for combining pext values across tissues to for example, get the mean pext value across GTEx.

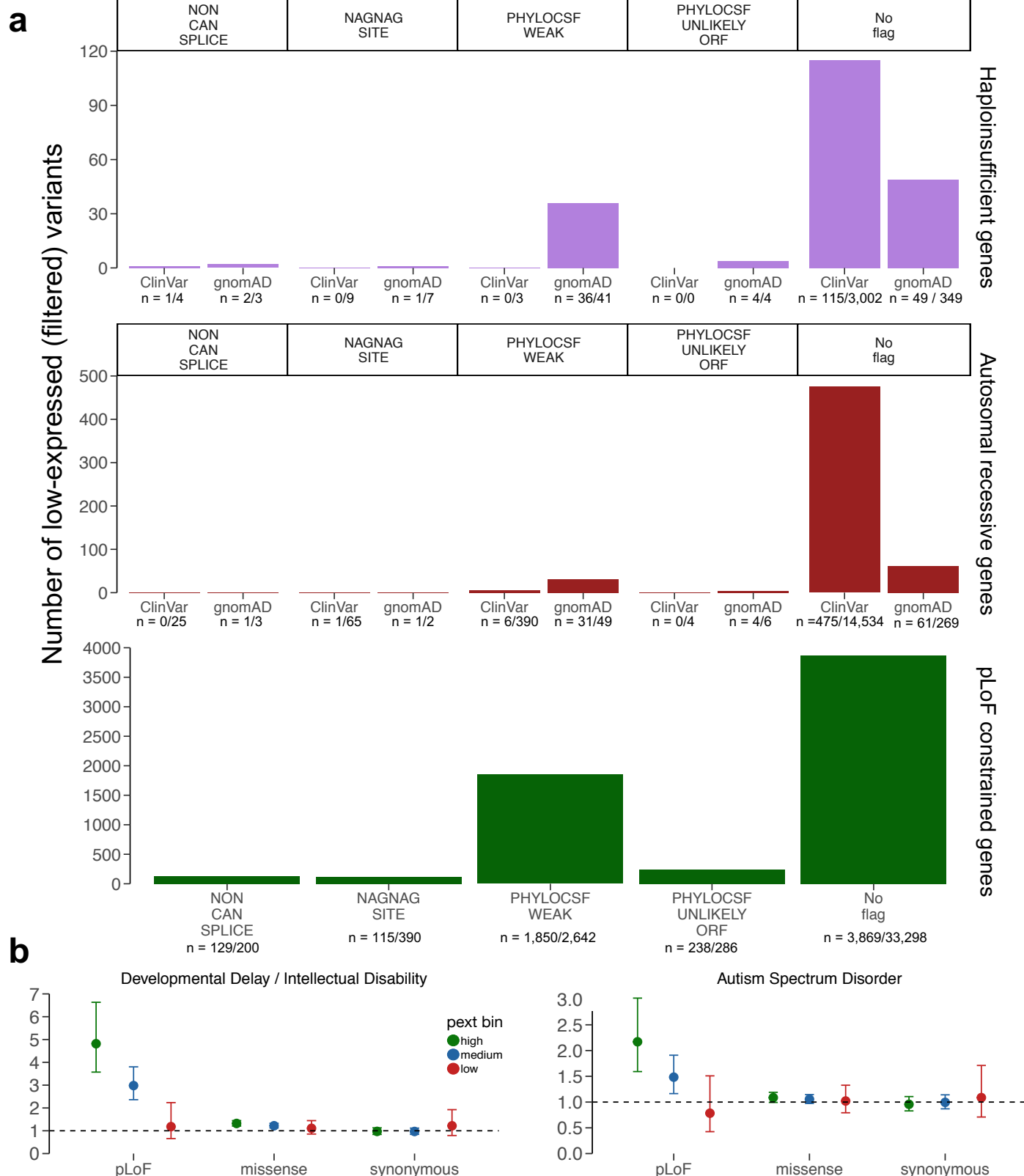


Supplementary Figure 5: Functional validation of pext **a.** Distribution of max open reading frame (ORF) scores from phyloCSF across the genome. We denoted exons with a maximum phyloCSF ORF score > 1000 as highly conserved and those with maximum phyloCSF ORF score < -100 as unconserved. Max ORF scores were capped at -1000 and 2000 for plotting. **b.** Sina plots of pext distribution in all gnomAD exome variants, partitioned on LOFTEE flags and filters (filters denoted in gray bars above plots). Red dots denote median average pext value per category and the number of variants per category listed below the plots. **d.** MAPS score for pLoF variants broken down by specific pLoF consequence shows consistent differences in MAPS for each pLoF category between high, medium and low pext expression bins. **d.** MAPS score of only missense variants denoted probably damaging by Polyphen-2 shows consistent gradient based on high (>0.9), medium ($0.1 \leq x \leq 0.9$) and low (<0.1) average GTEx pext expression bins showing that pext scores offer supplemental information to existing variant prioritization tools. For C and D, points represent MAPS values and error bars represent the 95% confidence interval, and the number of variants evaluated per category and unadjusted proportion singleton values can be found in Supplementary Table 5B and C, respectively.

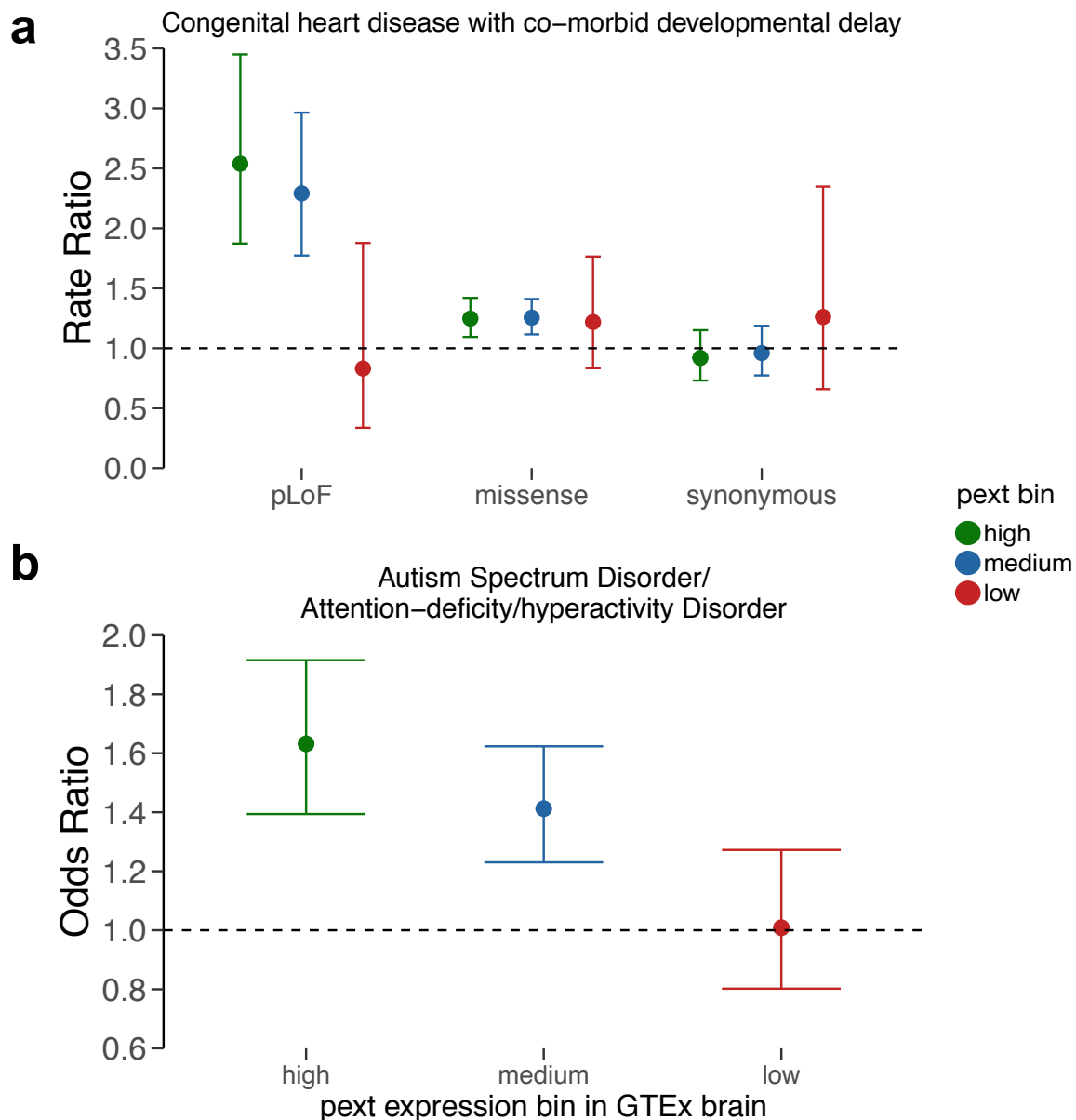


Supplementary Figure 6: Results of GENCODE of 128 unexpressed regions in haploinsufficient disease genes

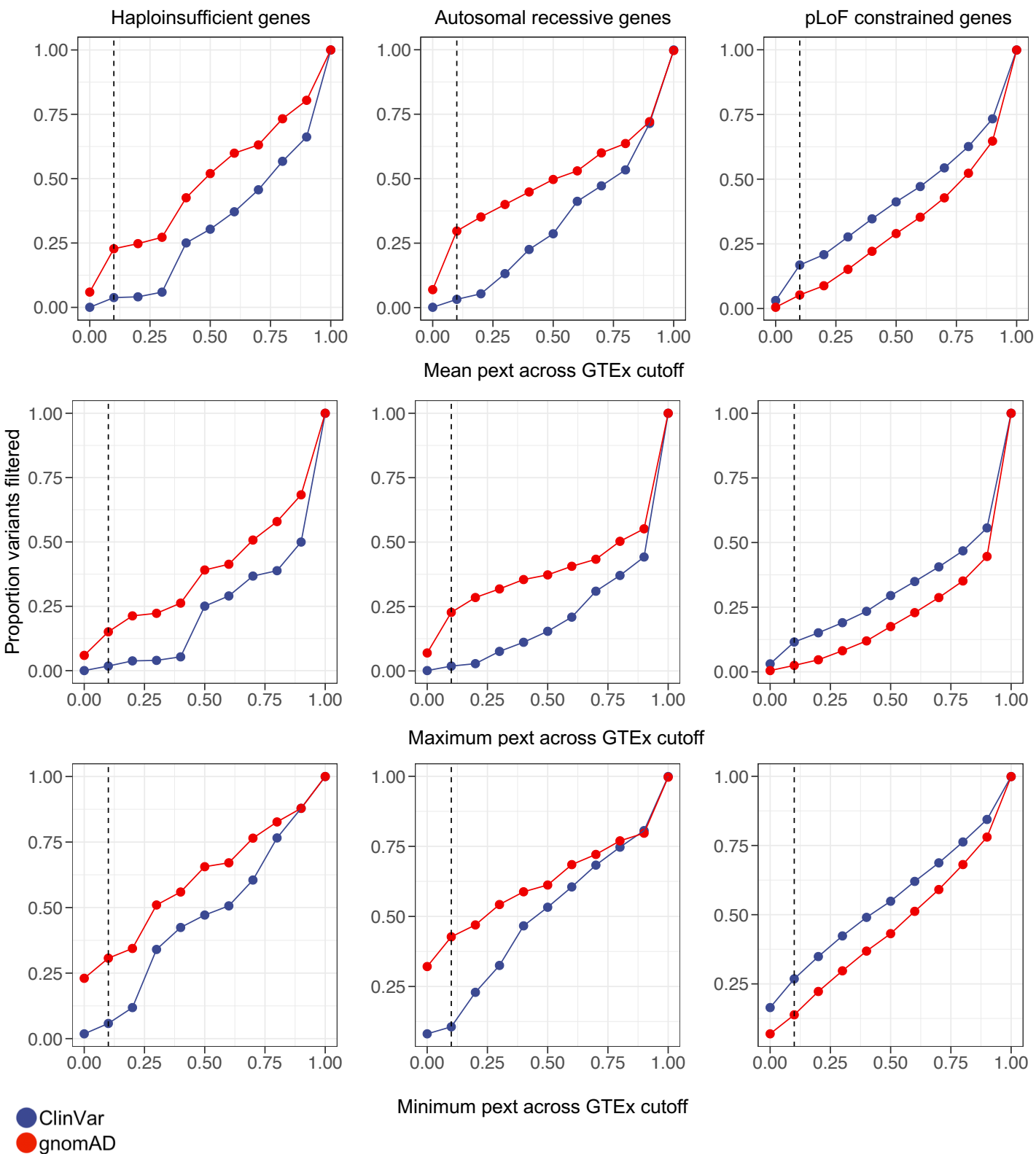
a. Summary of confidence in the CDS models tagged as unexpressed in GTEx based on expert manual evaluation. The major curation mode was putative annotation, where regions meet minimal annotation criteria but the coding potential of the region remains speculative. This was followed by regions that were marked as errors and non-coding regions, and have since been removed or are marked for removal based on this analysis. **b.** Summary of current annotation status of the regions in GENCODE v29. While some regions have been removed, or have been switched to a noncoding biotype, a majority remain in subsequent annotation sets **c.** An example of two erroneous gencode v19 CDS' in *RAI1* (chr17:17712481-17712483 and chr17:17714069-17714194, highlighted in blue) flagged by pext as unexpressed. The regions exhibit poor conservation and represents an incorrectly computationally predicted microexon and its downstream CDS, likely due to a poor quality cDNA alignment **d.** An example of a likely-coding CDS in *SCN2A* (chr2:166165675-166165766) which is well-conserved (gene model on top). While the region is unexpressed in GTEx, it exhibits considerable expression in fetal tissues from the Human Brain Developmental Resource (shown on bottom), highlighting the importance of incorporating multiple isoform datasets for accurate interpretation.



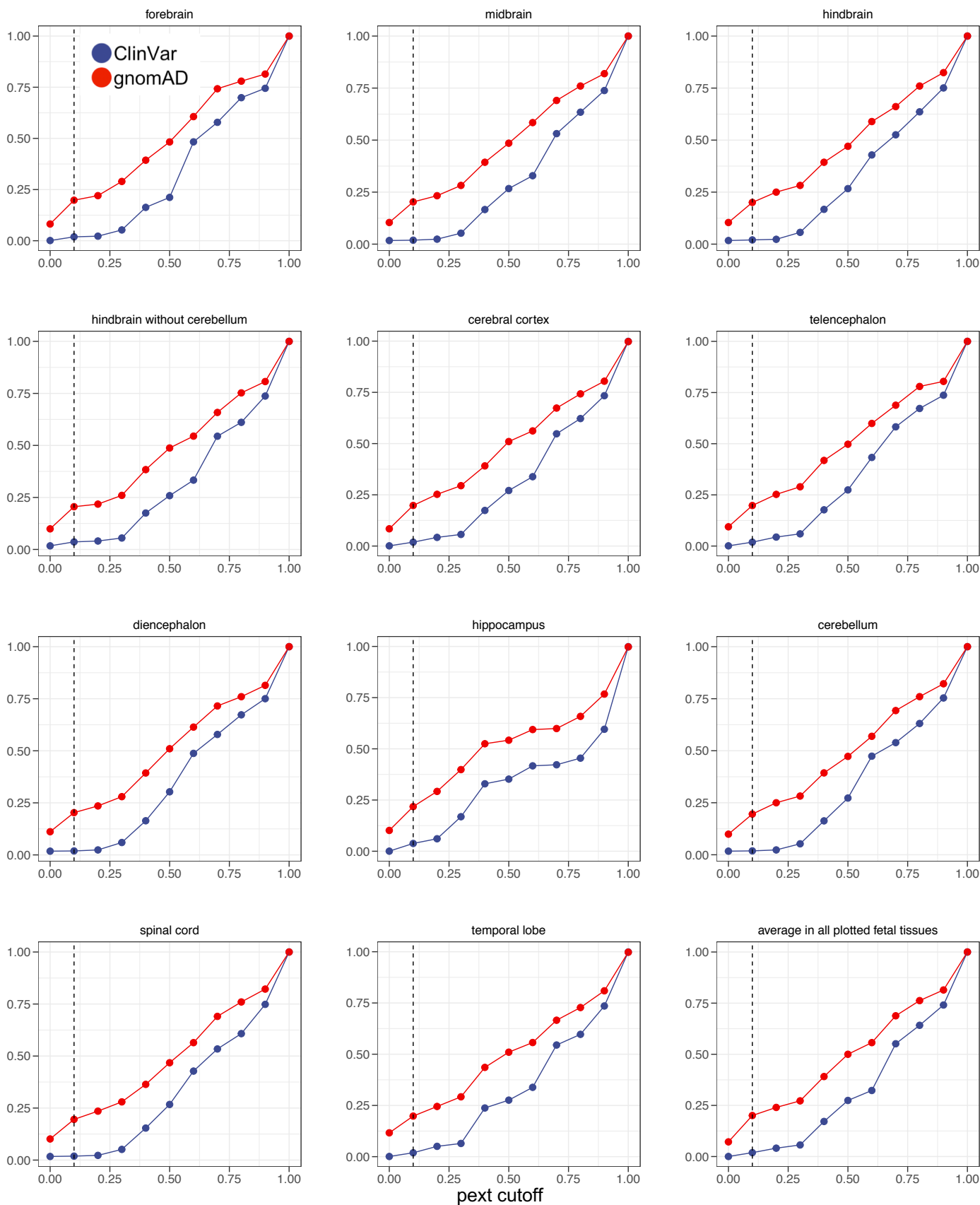
Supplementary Figure 7: Evaluation of pext after accounting for LOFTEE flags in ClinVar and *de novo* analyses **a.** Majority of pLoF variants filtered by pext are LOFTEE-HC with no flags. Comparison of proportion of high quality pLoF variants filtered in a curated list of 61 haploinsufficient developmental delays genes in gnomAD vs ClinVar (top), pLoF variants found in gnomAD in a homozygous state in at least one individual (middle) and pLoF and variants in gnomAD pLoF-intolerant genes (defined by LOEF < 0.35). Numbers below each x axis label represent the number of filtered variants over the number of total variants in the category. Synonymous variants are not shown for the bottom graph, given that synonymous variants do not receive LOFTEE designations. **b.** To show that results for *de novo* variant analysis do not change after controlling from LOFTEE flags, we regenerated Figure 4 restricted to only LOFTEE HC variants with no flags, removing 43, 43, and 13 *de novo* pLoF variants from DD/ID, ASD, and control groups, respectively. This remove 99 of all *de novo* pLoF variants from a total of 2,161 (representing 4.6% of the analyzed pLoF variants). Points represent rate ratio and error bars represent 95% confidence interval from the Poisson exact test.



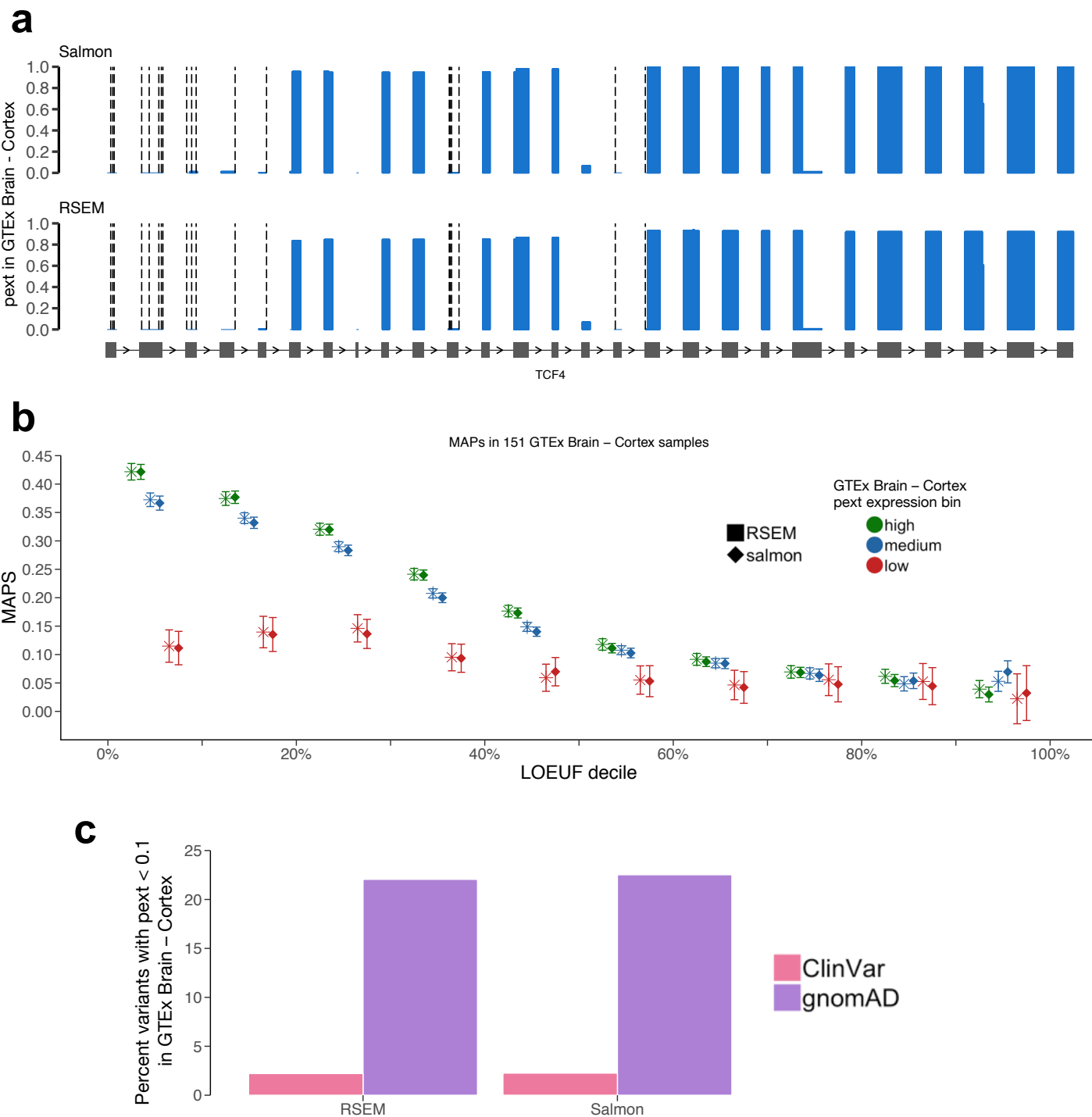
Supplementary Figure 8: Application of transcript-expression based annotation to *de novo* and rare variant analysis in additional datasets **a.** Using *de novo* variants identified in probands with congenital heart disease and co-morbid developmental delay we find a consistent effect of *de novo* pLoF variants found on high expressed regions in GTEx brain having larger effect sizes than *de novo* pLoF variants in weakly expressed regions. Once again, *de novo* pLoF variants found on regions with little evidence for expression are similarly distributed in cases and controls as *de novo* synonymous variants, suggesting such variants can be removed from gene burden testing analyses to boost discovery power. The high pext expression bin contains 41%, 32.7% and the low expression bin contains 3.7%, 9.3%, of 244, 327 *de novo* pLoF variants found in CHD/DD patients and controls, respectively. Points represent rate ratio estimate and error bars represent 95% confidence interval from the poisson exact test. **b.** Rare pLoF variants (combined AC in cases and controls ≤ 10) identified in highly constrained genes (first decile in LOEUF) grouped by pext expression bins show that those with high expression in GTEx brain have higher effect sizes than those identified in low-expressed regions, which are equally distributed in cases and controls. Points represent odds ratio estimate and error bars represent 95% confidence interval from two sided fisher's exact test on counts.



Supplementary Figure 9: Evaluation of using differing cutoffs for average, minimum or maximum pext across tissues While in this manuscript we have focused on binning average pext across GTEx tissues into low, medium and high-expressing categories, we also evaluated other summarizations of the GTEx pext scores in terms of the proportion of variants pLoF variants filtered from ClinVar and gnomAD in haploinsufficient severe disease genes, autosomal recessive disease genes, and pLoF constrained genes. All gene lists and variants represented are the same as used in Figure 4. We find that general trends of selective filtering of variants from gnomAD over ClinVar with pext is consistent across the varying methods for summarizing pext (average, minimum or maximum). Increasing the pext cut-off results in filtering more variants, and the given threshold used will depend on analysis needs. The dashed vertical line represents the 0.1 cutoff representing the low expression bin used for filtering in Figure 4.



Supplementary Figure 10: Using pext based on fetal RNA-seq isoform expression data for filtering variants in haploinsufficient genes Given that the 61 haploinsufficient genes are likely expressed during development, we evaluated the proportion of variants filtered from gnomAD and ClinVar in these genes using pext scores derived from the Human Brain Developmental Resource (see methods). We observe that fetal-expression based pext does not provide more selective filtering than using GTEx, highlighting that pext-based filtering captures mostly annotation errors.

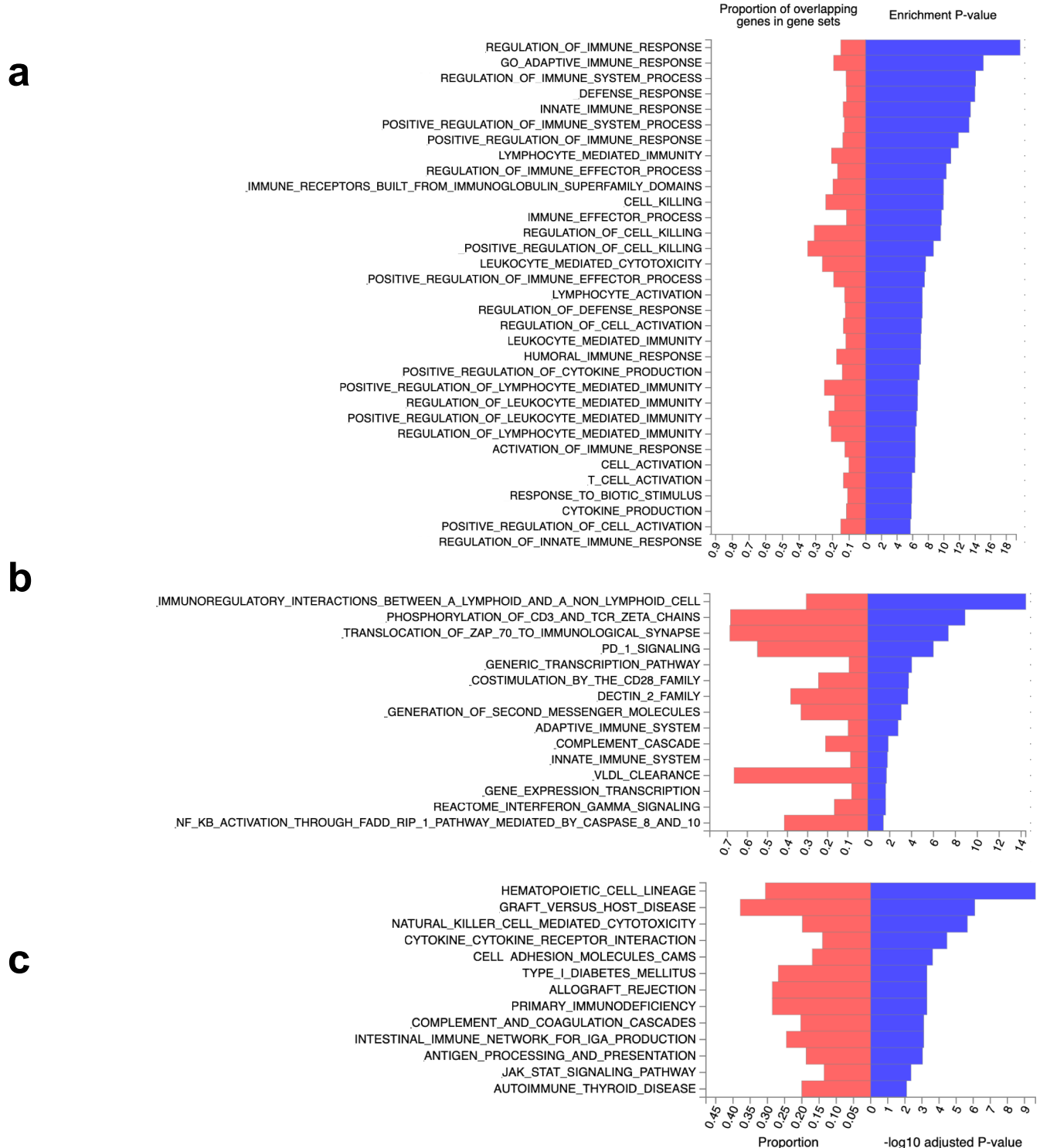


Supplementary Figure 11 : Comparison of key results using Salmon vs RSEM. Using isoform quantifications on 151 GTEx Brain - Cortex samples, we compared the results from analyses in the manuscript using quantifications from salmon and RSEM. Results from the analyses were consistent, suggesting that the pext calculation is robust to isoform quantification tool used. To keep analysis consistent, genes where the maximum pext was below 0.2 were removed, resulting in the filtering of 691 and 325 genes from the RSEM and salmon quantifications, respectively (see Methods)

a. The baselevel pext metric in *TCF4* using the two quantification tools. The 20 pLoF variants identified in gnomAD in *TCF4*, denoted by dashed lines, lie on unexpressed regions in Brain - Cortex samples using salmon and RSEM

b. No difference in the MAPs score is seen for pLoF variants in pext expression bins with RSEM and salmon quantifications. Points represent MAPs values and error bars represent the 95% confidence interval. The number of variants evaluated per category and unadjusted proportion singleton values can be found in Supplementary Table 6.

d. The number of pLoF variants filtered with a Brain - Cortex pext cutoff of 0.1 in gnomAD vs ClinVar was similar.



Supplementary Figure 12: Results from FUMA GENE2FUNC analysis in unconserved regions with high expression values We ran pathway analysis on 1,308 genes harboring 2,411 regions with low conservation (phyloCSF < -100) but high expression (pext > 0.9) shown in Figure 3A using the FUMA GENE2FUNC web browser. Results from **a**. Gene Ontology Biological Processes **b**. Reactome pathways and **c**. KEGG pathways show that these regions are enriched for immune pathways, which are selected for diversity but represent true coding regions, emphasizing the orthogonal information provided by pext over conservation alone. All plots and statistics generated were generated by FUMA, described in Ref 40. Full results from FUMA analysis are available in Supplementary Table 6.

Supplementary Table 1: Summary of manual curation flags for 401 pLoF in 61 haploinsufficient disease genes identified in gnomAD

Mapping error	Strand Bias	Reference error	Genotyping error	Homopolymer	MNV	Essential splice site rescue	Minority of transcripts	Weak exon conservation	Last exon	Other annotation error
Human self-chained repeats (UCSC)	Variant present on >90% of either forward or reverse strand	Reference contains small (<5bp) deletion within exon	Allele balance ≤ 35%	Repeat of ≥ 6 base pairs	In phase MNV that abolishes stop codon	In frame splice site rescue within 36 bp and validated by Alamut	Variant falls on <50% of coding NCBI RefSeq transcripts	Flagged by phyloCSF as weak	Variant falls within the terminal coding exon	Variant falls on exactly 50% of NCBI coding RefSeq transcripts
Tandem repeats (UCSC)			GC rich region	Repeat of ≥ 6 trinucleotide repeats	Frame restoring indel			Entire exon is weakly conserved upon visual inspection in the UCSC browser	Variant falls within 50 base pairs of penultimate coding exon	Exon is partially conserved, specifically lacking conservation of variant/transcript
Segmental dups (UCSC)			GQ ≤ 20						Variants affects >25% coding sequence	Re-initiation by downstream methionine in first coding exon
Complex variation in region e.g. multiple indels, SNVs			Low complexity sequence							Stop codon occurs upstream of variant
			Low read depth <15							

Supplementary Table 2: Summary of criteria for LoF verdicts of 401 pLoF in 61 haploinsufficient disease genes identified in gnomAD

LoF	Likely LoF	Likely not LoF	Not LoF
GQ > 99, absence of any evidence to refute a LOF consequence and one of two criteria met: AB>35 and read depth >15	Low complexity region	Re-initiation by downstream methionine in first coding exon	Minority of coding RefSeq transcripts (except when exon well conserved)
	Allele balance ≤ 35% but >25%	Mapping ambiguity (UCSC)	Variant falls in last coding exon
	GQ ≤ 20	phyloCSF weak	Weak conservation of exon
	GC rich region	In frame splice site rescue between 6 and 21 base pairs from the intron/exon boundary and validated by Alamut	Frame restoring indel
	Strand bias in regions where coverage is skewed towards a strand	Homopolymer repeat	In phase multi-nucleotide variant abolishing stop codon
	Read depth <15	Falls within terminal exon although will disrupt >25% but <50% of coding sequence	Complete splice site rescue within 6 base pairs and validated in Alamut
	Any single other transcript error	Minority of transcripts but exon well conserved	Complex mapping and assembly error
	Potential splice site rescue >21 BP away and weakly supported by Alamut	Combination of multiple flags (e.g. On 50% of RefSeq transcripts and partial loss of exon conservation; homopolymer repeat and mapping error etc.)	Reference error
	Splice rescue within 21 BP but very weak signal as per Alamut	Strand bias despite equal coverage of forward and reverse strand across region	Combination of multiple flags (e.g. in frame splice rescue within 21 bp and minority of transcripts with exon well conserved; mapping ambiguity and phyloCSF weak etc.)
	Falls within terminal exon although will disrupt >50% of coding sequence		

Supplementary Table 5: Total number of variants and unadjusted proportion singleton per category in MAPS analyses. The number of variants that went into analyses and the proportion singleton in parenthesis is presented for figures **A**. Figure 3C **B**. Supplementary Figure 5C **C**. Supplementary Figure 5D

Supplementary Figure 5C C. Supplementary Figure 5D

LOEUF decile												
A.	pext bin	0-10%	10-20%	20-30%	30-40%	40-50%	50-60%	60-70%	70-80%	80-90%	90-100%	
	synonymous	low	14,493 (0.50)	9,267 (0.50)	9,309 (0.49)	9,402 (0.48)	11,934 (0.49)	10,199 (0.45)	5,306 (0.50)	4,834 (0.48)	2,803 (0.49)	2,176 (0.46)
		medium	240,168 (0.48)	173,420 (0.48)	150,198 (0.48)	137,859 (0.48)	125,339 (0.48)	107,059 (0.48)	90,436 (0.47)	61,464 (0.47)	38,439 (0.48)	21,086 (0.47)
		high	138,152 (0.48)	113,316 (0.48)	105,853 (0.48)	97,491 (0.48)	84,640 (0.48)	80,150 (0.48)	72,823 (0.48)	68,441 (0.48)	56,124 (0.48)	45,881 (0.48)
missense	low	34,294 (0.55)	25,084 (0.53)	25,150 (0.52)	26,031 (0.52)	32,509 (0.51)	27,507 (0.49)	16,803 (0.50)	14,865 (0.49)	9,077 (0.50)	6,453 (0.50)	
	medium	420,567 (0.57)	340,944 (0.55)	305,989 (0.54)	301,723 (0.53)	289,188 (0.52)	256,704 (0.52)	221,591 (0.51)	150,958 (0.51)	98,752 (0.51)	52,775 (0.50)	
	high	235,748 (0.57)	202,928 (0.56)	205,232 (0.55)	201,511 (0.55)	187,259 (0.53)	184,222 (0.53)	169,631 (0.53)	162,410 (0.52)	134,430 (0.51)	108,390 (0.50)	
pLoF	low	1,154 (0.62)	1,172 (0.65)	1,431 (0.63)	1,544 (0.58)	1,772 (0.58)	1,828 (0.55)	1,387 (0.52)	1,297 (0.55)	824 (0.56)	629 (0.52)	
	medium	4,289 (0.83)	7,661 (0.79)	10,146 (0.75)	1,337 (0.68)	16,134 (0.64)	16,436 (0.59)	15,352 (0.58)	11,074 (0.57)	7,006 (0.57)	3,467 (0.59)	
	high	1,843 (0.86)	3,576 (0.82)	5,554 (0.78)	7,489 (0.71)	8,423 (0.66)	9,438 (0.61)	9,468 (0.59)	9,278 (0.57)	7,693 (0.57)	5,558 (0.56)	

B.	pext bin	0-10%	10-20%	20-30%	30-40%	40-50%	50-60%	60-70%	70-80%	80-90%	90-100%	
	splice acceptor	low	294 (0.65)	253 (0.66)	312 (0.63)	284 (0.61)	323 (0.61)	380 (0.61)	254 (0.55)	232 (0.60)	175 (0.59)	111 (0.58)
		medium	850 (0.78)	1,330 (0.80)	1,803 (0.80)	2,287 (0.75)	2,842 (0.70)	2,790 (0.67)	2,710 (0.65)	2,069 (0.63)	1,264 (0.63)	667 (0.63)
		high	297 (0.83)	561 (0.83)	808 (0.82)	11,45 (0.78)	1,283 (0.74)	1,404 (0.67)	1411 (0.64)	1,391 (0.64)	11,45 (0.61)	850 (0.59)
splice donor	low	254 (0.63)	250 (0.67)	306 (0.64)	329 (0.62)	314 (0.60)	336 (0.55)	236 (0.56)	258 (0.55)	140 (0.61)	100 (0.53)	
	medium	930 (0.80)	1,606 (0.80)	2,151 (0.79)	2,832 (0.70)	3,476 (0.67)	3,565 (0.63)	3,358 (0.60)	2,356 (0.59)	1,619 (0.60)	806 (0.58)	
	high	387 (0.82)	681 (0.80)	1,046 (0.81)	1,517 (0.75)	1,681 (0.70)	1,963 (0.63)	1,799 (0.62)	1,727 (0.59)	1,481 (0.55)	1,058 (0.57)	
stop gained	low	606 (0.61)	665 (0.63)	821 (0.63)	924 (0.55)	1139 (0.56)	1115 (0.53)	893 (0.50)	808 (0.54)	504 (0.53)	418 (0.50)	
	medium	2,507 (0.86)	4,727 (0.79)	6,193 (0.73)	8,258 (0.65)	9,815 (0.61)	9,990 (0.56)	9,297 (0.56)	6,636 (0.55)	4,123 (0.55)	1,995 (0.59)	
	high	1159 (0.88)	2,334 (0.82)	3,678 (0.75)	4844 (0.69)	5474 (0.62)	6064 (0.59)	6259 (0.57)	6162 (0.55)	5073 (0.56)	3637 (0.55)	

C.	pext bin	0-10%	10-20%	20-30%	30-40%	40-50%	50-60%	60-70%	70-80%	80-90%	90-100%	
	Polyphen-2 probably damaging missense	low	6,121 (0.58)	5,170 (0.58)	3,381 (0.54)	3,351 (0.53)	4,050 (0.52)	4,855 (0.51)	2,283 (0.53)	1,893 (0.53)	1,115 (0.53)	900 (0.51)
		medium	79,410 (0.61)	68,398 (0.59)	68,473 (0.58)	70,816 (0.56)	70,083 (0.55)	63,015 (0.54)	56,645 (0.53)	38,342 (0.53)	24,853 (0.53)	13,629 (0.53)
		high	48,478 (0.61)	44,437 (0.60)	49,038 (0.59)	50,499 (0.58)	49,448 (0.56)	48,989 (0.55)	48,435 (0.55)	44,595 (0.54)	36,757 (0.53)	27,229 (0.53)

Supplementary Table 6: Total number of variants and unadjusted proportion singleton per category in MAPS analyses. The number of variants evaluated for MAPS comparison between RSEM and salmon in Supplementary Figure 11B and the proportion singleton in parenthesis.

		LOEUF decile										
A.		pext bin	0-10%	10-20%	20-30%	30-40%	40-50%	50-60%	60-70%	70-80%	80-90%	90-100%
synonymous	RSEM	low	12,128 (0.50)	9,239 (0.50)	10,277 (0.49)	10,209 (0.48)	9,454 (0.47)	8,288 (0.44)	5,830 (0.44)	4,758 (0.47)	3,496 (0.49)	1,849 (0.47)
		medium	215,842 (0.48)	156,432 (0.48)	134,961 (0.47)	128,388 (0.48)	113,157 (0.48)	97,731 (0.48)	79,710 (0.47)	52,873 (0.47)	32,130 (0.47)	18,624 (0.47)
		high	158,044 (0.48)	123,040 (0.48)	113,344 (0.48)	97,120 (0.48)	84,008 (0.48)	74,947 (0.49)	65,301 (0.48)	57,738 (0.48)	45,067 (0.48)	34,563 (0.48)
	salmon	low	11,501 (0.50)	7,901 (0.49)	8,698 (0.49)	9,268 (0.48)	8,635 (0.47)	7,555 (0.43)	4,952 (0.49)	3,594 (0.48)	3,638 (0.48)	1,586 (0.48)
		medium	197,101 (0.49)	141,550 (0.47)	12,1424 (0.47)	11,4973 (0.48)	104,930 (0.48)	83,645 (0.48)	65,777 (0.47)	43,395 (0.47)	26,561 (0.48)	15,809 (0.47)
		high	188,240 (0.48)	146,243 (0.48)	137,011 (0.48)	124,447 (0.48)	109,508 (0.48)	102,721 (0.49)	91,849 (0.48)	78,706 (0.48)	57,388 (0.48)	42,695 (0.47)
pLoF	RSEM	low	1,127 (0.61)	1,172 (0.63)	1,550 (0.63)	1,633 (0.59)	1,660 (0.56)	1,529 (0.55)	1,419 (0.54)	1,211 (0.56)	947 (0.56)	494 (0.54)
		medium	3,828 (0.83)	6,983 (0.79)	9,128 (0.75)	12,396 (0.68)	14,446 (0.64)	14,797 (0.60)	13,578 (0.59)	9,597 (0.57)	5,907 (0.57)	2,982 (0.59)
		high	2,179 (0.86)	3,955 (0.82)	6,020 (0.78)	7,597 (0.71)	8,543 (0.66)	9,002 (0.61)	8,614 (0.59)	7,824 (0.57)	6,157 (0.58)	4,073 (0.57)
	salmon	low	1,061 (0.61)	1,010 (0.62)	1,382 (0.62)	1,523 (0.59)	1,518 (0.57)	1,278 (0.55)	1,227 (0.54)	993 (0.55)	892 (0.55)	409 (0.55)
		medium	3,672 (0.82)	6,565 (0.79)	8,463 (0.75)	11,506 (0.67)	13,785 (0.63)	12,832 (0.60)	11,300 (0.59)	7,866 (0.57)	5,022 (0.57)	2,427 (0.60)
		high	2,616 (0.86)	4,819 (0.82)	7,452 (0.77)	10,001 (0.71)	11,334 (0.65)	13,250 (0.60)	12,924 (0.59)	11,537 (0.57)	8,232 (0.57)	5,543 (0.56)