Supplemental Figures

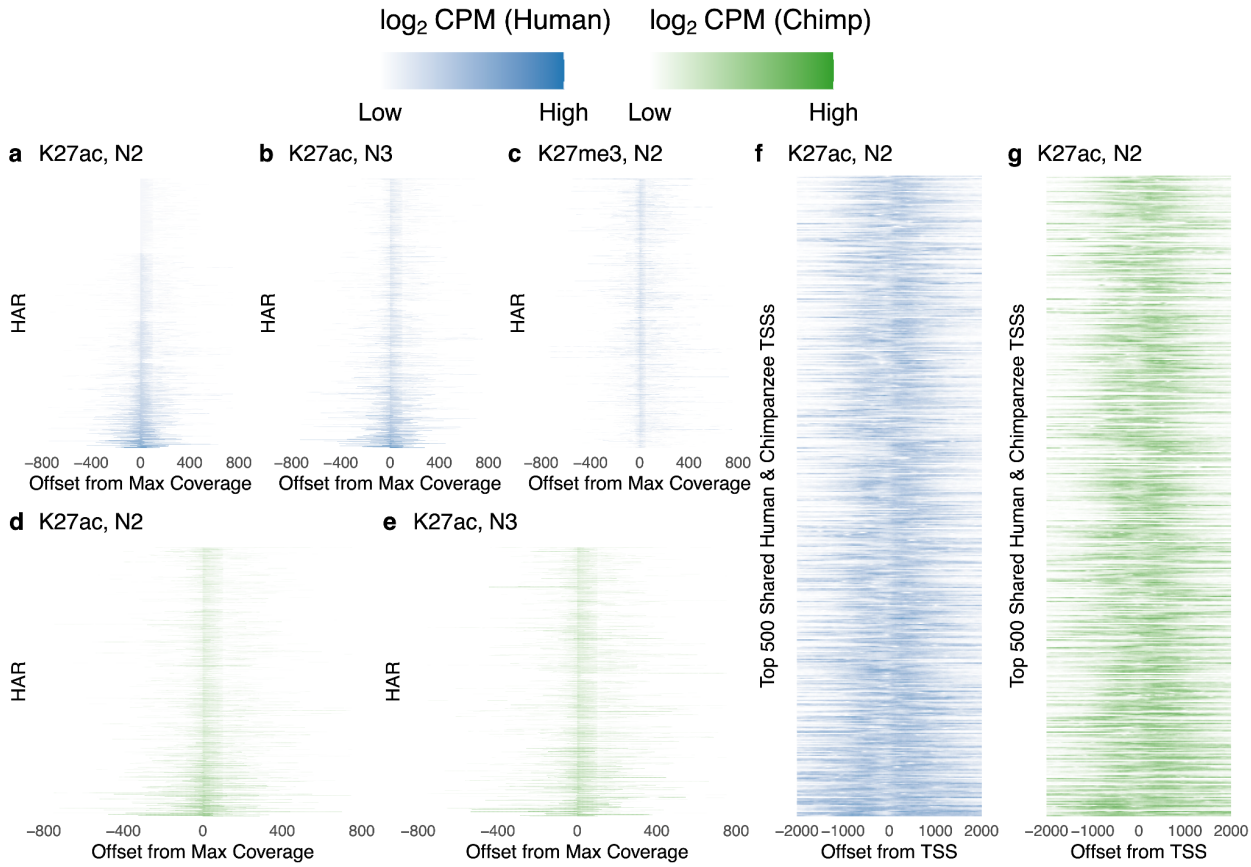

Figure S1. ChIP-seq signal across cell lines at HARs and genome-wide, related to Figure 1. (a-e) Normalized ChIP-seq read coverage (CPM) at HARs. For each HAR, reads are centered on the position with maximum coverage. Rows in all panels are ordered by coverage in panel a. One representative replicate is shown for each condition. (a) H3K27ac in human N2, (b) H3K27ac in human N3, (c) H3K27me3 in human N2, (d) H3K27ac in chimpanzee N2, (e) H3K27ac in chimpanzee N3. (f-g) Normalized read coverage (CPM) genome-wide at transcription start sites (TSS). Rows in both panels are ordered by coverage at the TSS (offset 0) in panel f. All TSS shared between the human and chimpanzee genomes are analyzed. (f) H3K27ac in human N2, (g) H3K27ac in chimpanzee N2.

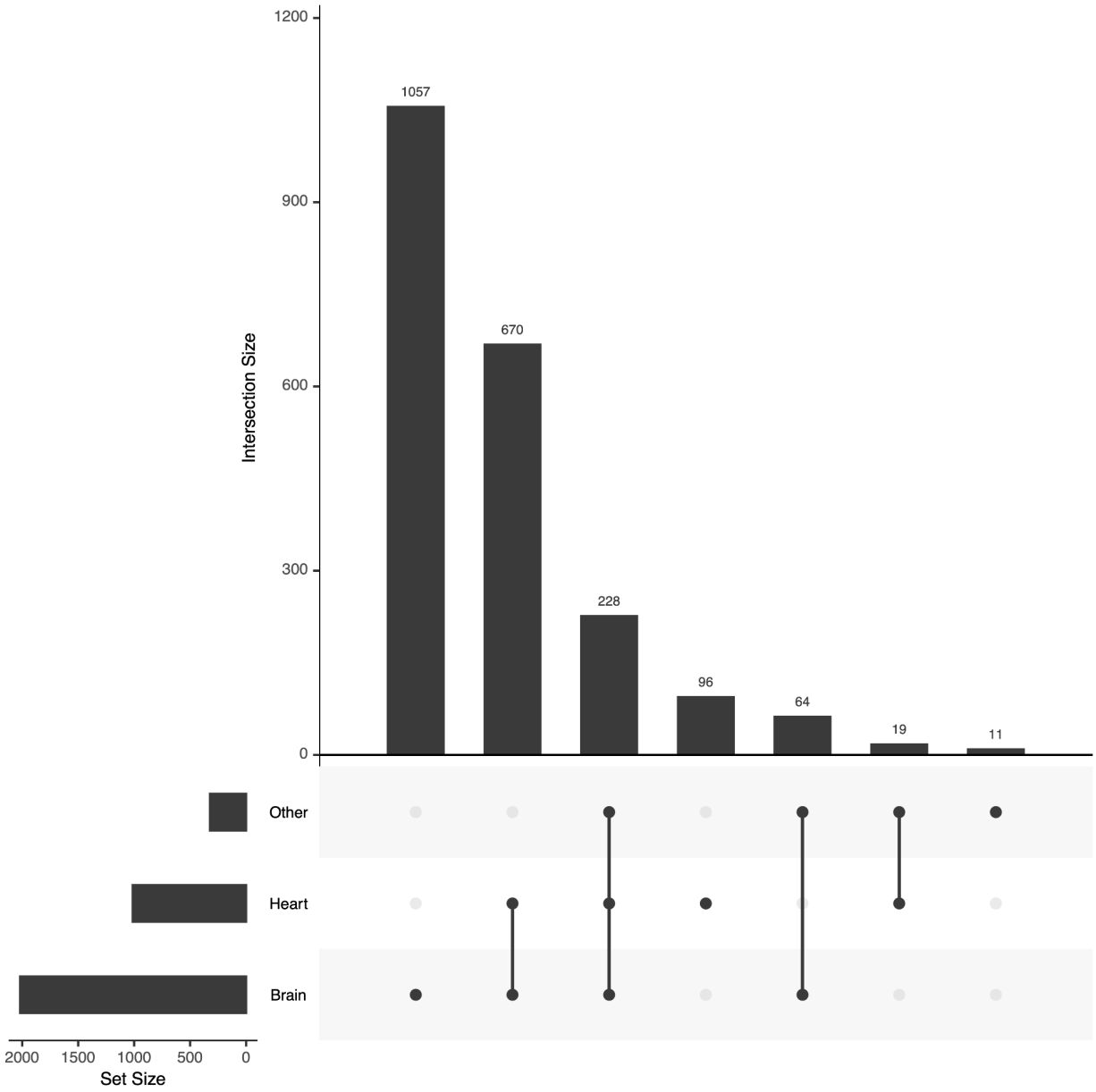

Figure S2. Overlap of HARs with epigenetic datasets for heart, brain, and other primary tissues, related to Figure 2.

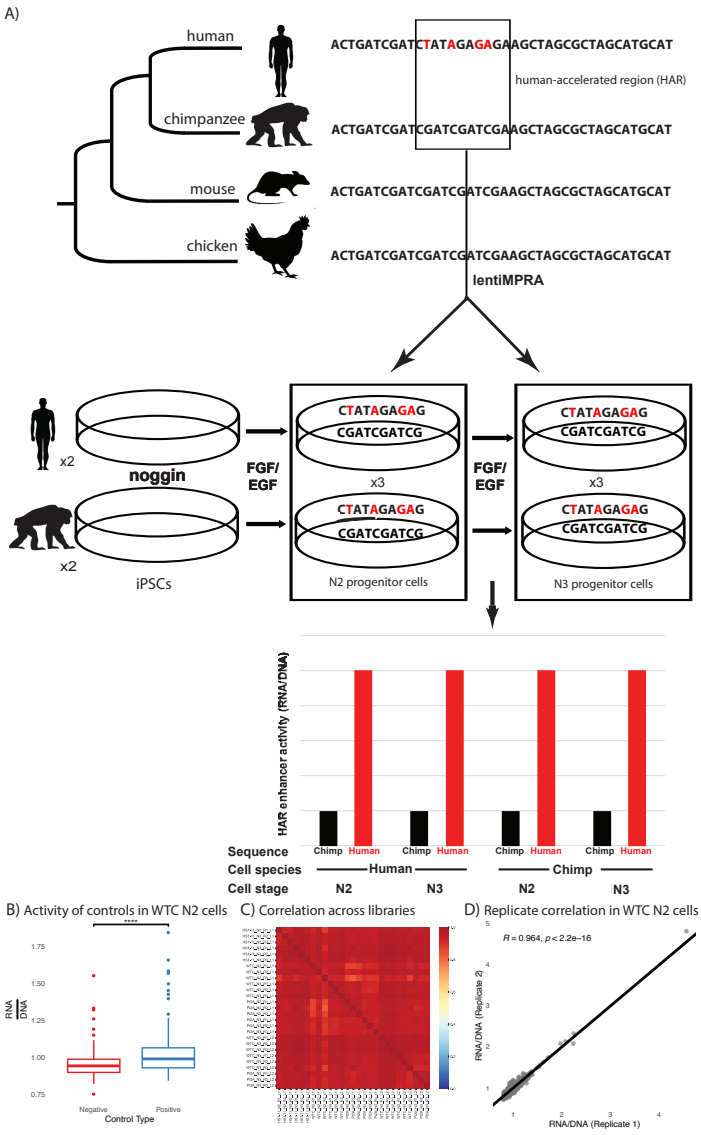

Figure S3. LentiMPRA study design and validation, related to Figure 5. (a) Study design for lentiMPRA experiments. (b) LentiMPRA activity of control sequences in WTC N2 cells. The distribution of RNA/DNA ratios is lower for negative controls (H3K27me3, red) compared to positive controls (H3K27ac, blue). A similar difference was observed in all cell lines and stages. (c) Permutation lentiMPRA measurements are highly concordant within and between two independent libraries. We observed high correlation between lentiMPRA activity levels of permutation oligos in library 1 and library 2 after batch correction. Correlation is high across all cell lines, cell stages, and replicates. (d) LentiMPRA activity measurements are highly reproducible between two technical replicates of WTC N2 cells. We observed high correlations of RNA/DNA ratios for all combinations of species, cell line, cell stage, and replicate, with the highest correlation for technical replicates (range 0.90-0.97), followed by biological replicates (different cell line) of the same species at the same stage (range 0.86-0.97), biological replicates of the same species at N2 versus N3 (range 0.80-0.97), and replicates from different species (range 0.54-0.97).

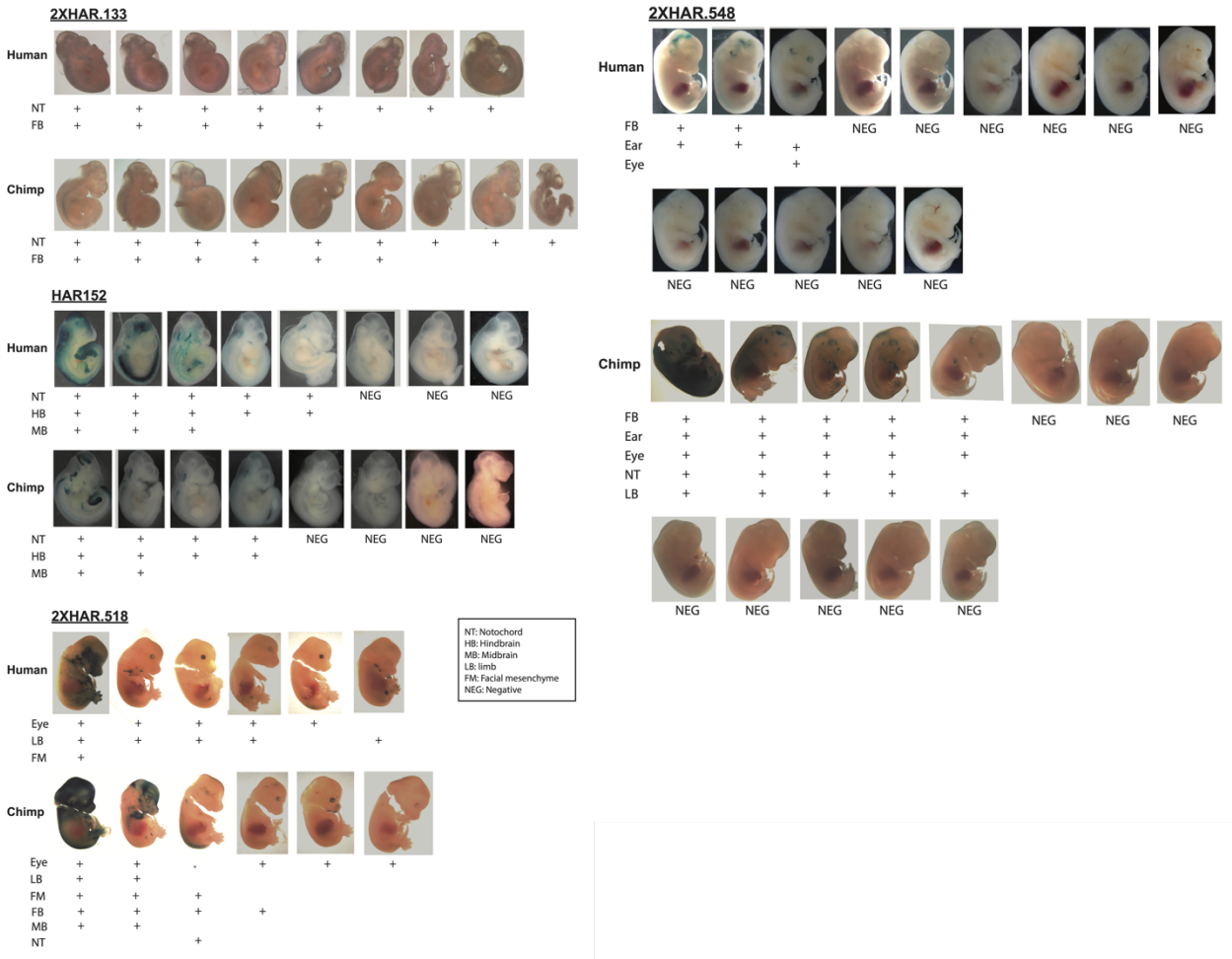

Figure S4. Transgenic mouse embryos for 2xHAR.133, HAR152, 2xHAR.518 and 2xHAR.548, related to Figures 3 and 5. Pictures of all PCR positive embryos from mouse transgenic enhancer assays. The name of the tested HAR is noted at the top left of each set of images, and the sequence origin (human or chimpanzee) is given in each row. Tissue names are abbreviated as notochord (NT), forebrain (FB), midbrain (MB), limb (LB), and facial mesenchyme (FM). A plus sign (+) in the row of the tissue name means LacZ expression was observed for that tissue. Embryos that did not show LacZ expression but were LacZ positive by PCR, are marked as negative (NEG).

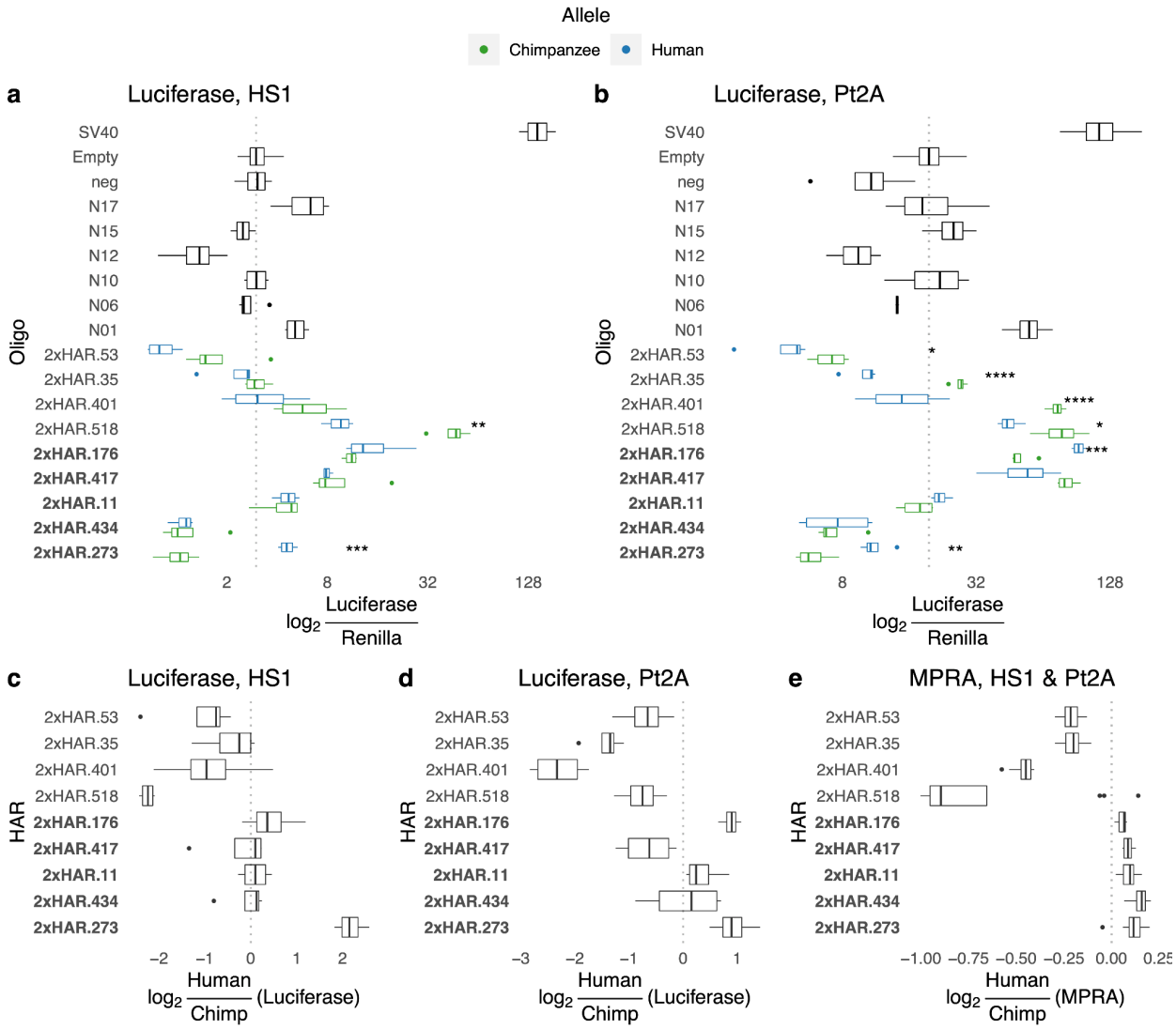

Figure S5. Luciferase assays for validation of lentiMPRA, related to Figure 5. (a-b) Relative luciferase activity in human (HS1, panel a) and chimpanzee (Pt2a, panel b) N3 cells for human (blue) or chimpanzee (green) HAR sequences. We selected HARs across a range of RNA/DNA levels (mean of human and chimpanzee alleles) and tested whether relative activity of the human and chimpanzee alleles was similar in luciferase assays compared to lentiMPRA. 2xHAR.273, 2xHAR.434, 2xHAR.11, 2xHAR.417, and 2xHAR.176 (bold font) had higher activity of the human allele in lentiMPRA. 2xHAR.518, 2xHAR.401, 2xHAR.35, 2xHAR.53, and 2xHAR.364 (plain font) had higher activity of the chimpanzee allele in lentiMPRA. 2xHAR.401, 2xHAR.518, 2xHAR.176, 2xHAR.417, 2xHAR.11, and 2xHAR.173 each have at least one allele (human and/or chimpanzee sequence) significantly higher than the empty vector. We also tested seven “inactive” sequences with low activity in lentiMPRA (neg, N01, N06, N10, N12, N15, N17), empty pLS-mP-luc vector (Empty), and pLS-SV40-mP-luc (SV40). Asterisks indicate significant differences between the human and chimpanzee alleles by Student's t-test. (c-e) Relative activity of human versus chimpanzee alleles for HARs from (a-b) that showed species-biased activity in lentiMPRA. (c) HS1 luciferase, (d) Pt2a luciferase, and (e) lentiMPRA.

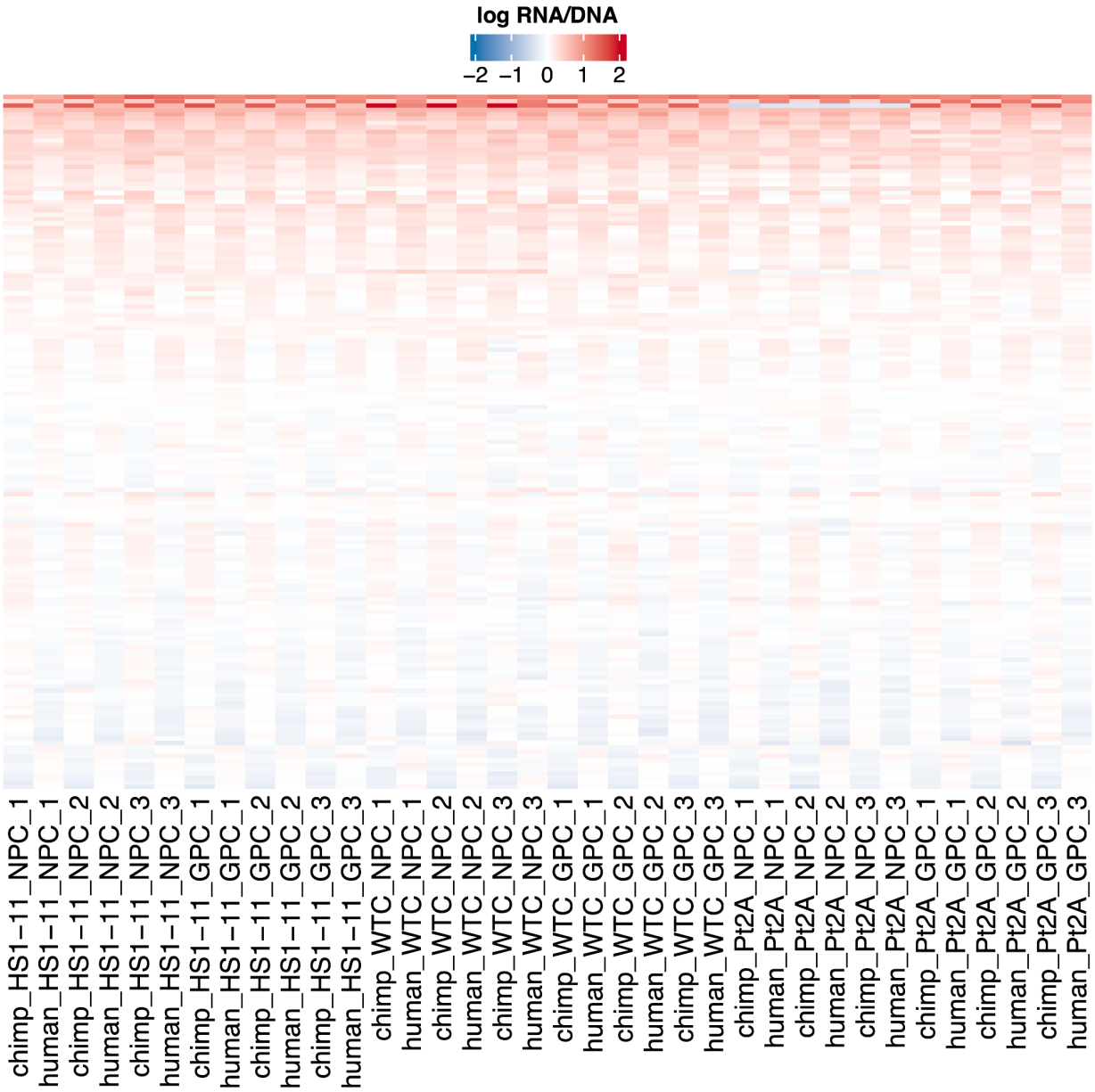

Figure S6. Heatmap of lentiMPRA activity for all species-biased HARs, related to Figure 5. Chimpanzee and human sequences of the same HAR are plotted next to each other. The stripes present across all cell lines, cell stages, and replicates show consistent differences in activity between the alleles from the two species.

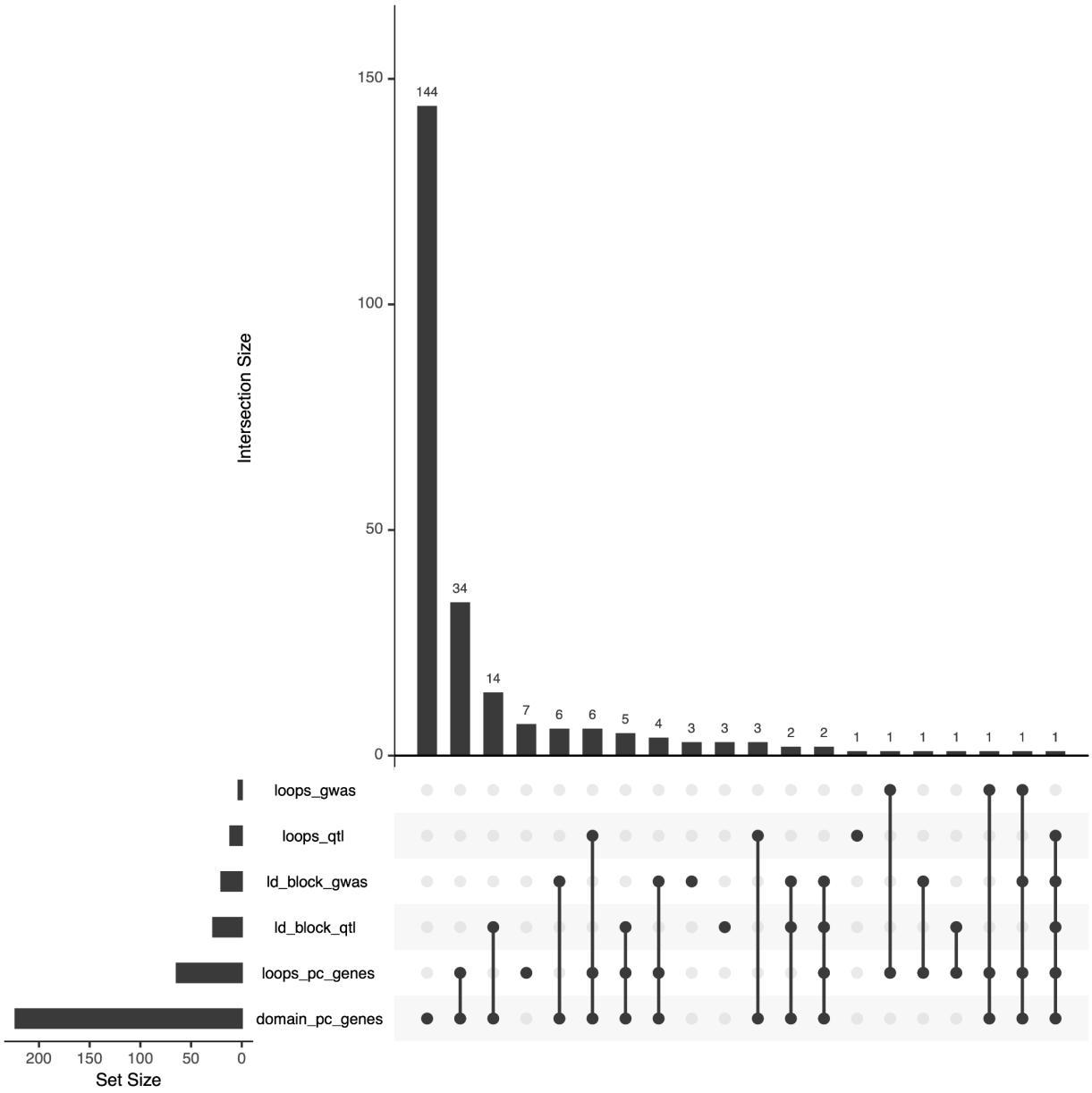

Figure S7. Associating active HARs with neurodevelopmental genes, related to Figures 4-5. Number of active HARs (lentiMPRA) interacting with GWAS variants via chromatin loops, interacting with QTLs via chromatin loops, sharing an LD block with GWAS variants, sharing an LD block with QTLs, interacting with the promoter of protein coding genes via chromatin loops, and sharing a contact domain (sub-TAD) with promoters of protein-coding genes.

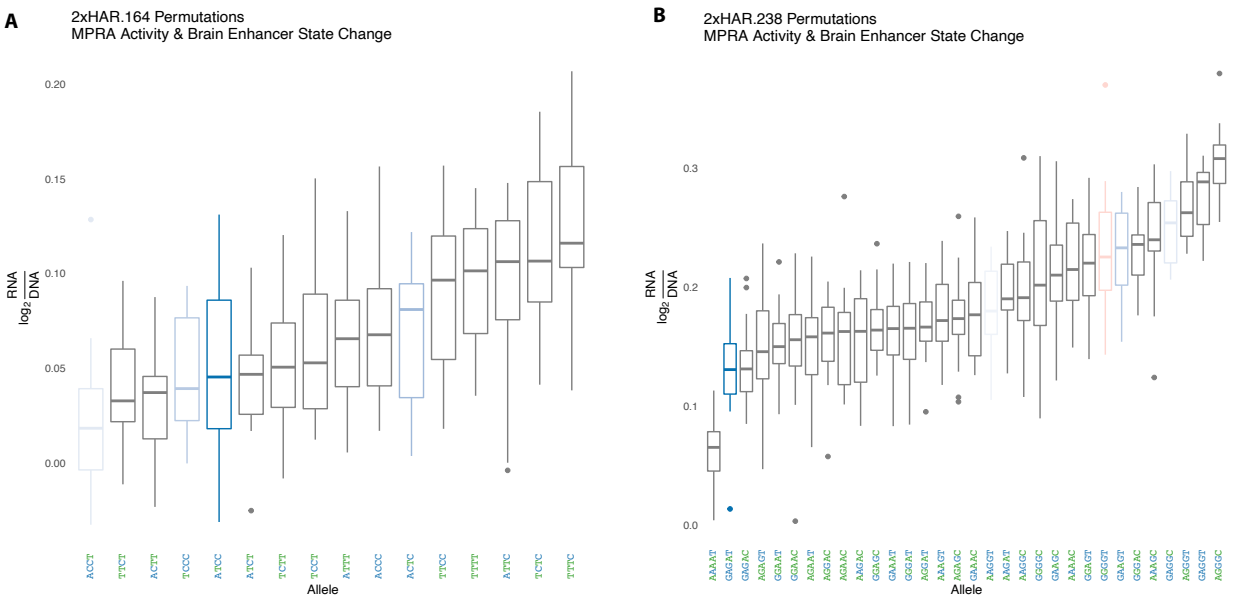

Figure S8. LentiMPRA activity of HAR permutations, related to Figure 7. Activity of all evolutionary intermediates for (a) 2xHAR.164 and (b) 2xHAR.238. Human bases of the allele are shown in blue and chimpanzee bases in green. Single nucleotide changes are colored with the predicted brain enhancer state change from Sei, showing moderate concordance with lentiMPRA activity.

Supplemental Tables

Species	Stage	Donor	Capture Plate ID	# Cells
Human	N2	WTC	1771064068	49
Chimp	N2	Pt2	1771064069	27
Chimp	N2	Pt5	1784059104	20
Human	N2	Hs1	1784060057	50
Chimp	N3	Pt5	1784060058	84
Chimp	N3	Pt2	1784073115	69
Human	N3	WTC	1784073116	57

Table S1. Single-cell RNA-seq cell counts, related to Figure 1.

		Any Tissue		Brain		Telencephalon		NPC MPRA
HAR	**Refs**	**Human**	**Chimp**	**Human**	**Chimp**	**Human**	**Chimp**	**(this study)**
HAR2/2xHAR.3/ HACNS1	6,7	Yes	Yes	No	No	No	No	Inactive
HAR5/2xHAR.557	7	No	Not tested	No	Not tested	No	Not tested	Chimp biased
2xHAR.5	2	No	Not tested	No	Not tested	No	Not tested	Inactive
2xHAR.20	1,2	Yes	Yes	Yes	Yes	Yes	Yes	Human biased
HAR25	2	Yes	Yes	Yes	Yes	No	No	Human biased
HAR31	5,7	Yes	Not tested	No	Not tested	No	Not tested	Inactive
HAR34	2,7	Yes	Yes	Yes	Yes	Yes	Yes	Inactive
2xHAR.46	2	Yes	Yes	No	No	No	No	Chimp biased
2xHAR.59	7	No	Not tested	No	Not tested	No	Not tested	Inactive
HAR69	2	No	Not tested	No	Not tested	No	Not tested	Inactive
2xHAR.82	7	No	Not tested	No	Not tested	No	Not tested	Inactive
2xHAR.90	2	No	No	No	No	No	No	Active
2xHAR.92	7	No	Not tested	No	Not tested	No	Not tested	Active
2xHAR.93	2	Yes	Yes	No	No	No	No	Inactive
2xHAR.97	7	Yes	Not tested	No	Not tested	No	Not tested	Active
2xHAR.99	2	Yes	Yes	No	No	No	No	Inactive
HAR104	7	Yes	Not tested	Yes	Not tested	Yes	Not tested	Inactive
2xHAR.114	2	Yes	Yes	Yes	Yes	No	No	Human biased
2xHAR.116	2	Yes	Yes	Yes	Yes	Yes	Yes	Inactive
HAR118	7	Yes	Not tested	Yes	Not tested	Yes	Not tested	Chimp biased
HAR119/ 2xHAR.18	7	No	Not tested	No	Not tested	No	Not tested	Chimp biased
HAR122	7	Yes	Not tested	No	Not tested	No	Not tested	Active
2xHAR.122	2	No	No	No	No	No	No	Inactive
2xHAR.128	2	Yes	Yes	No	No	No	No	Inactive
2xHAR.138	2	Yes	Yes	Yes	Yes	Yes	Yes	Chimp biased
2xHAR.142	3	Yes	Yes	Yes	Yes	Yes	Yes	Inactive
HAR143	7	Yes	Not tested	Yes	Not tested	Yes	Not tested	Inactive
HAR157	7	No	Not tested	No	Not tested	No	Not tested	Inactive
HAR164	7	Yes	Not tested	No	Not tested	No	Not tested	Inactive
2xHAR.164	2	Yes	Yes	Yes	Yes	Yes	Yes	Active
2xHAR.170	2	Yes	Yes	Yes	Yes	Yes	Yes	Human biased
HAR180/ 2xHAR.427	7	No	Not tested	No	Not tested	No	Not tested	Inactive
HAR196	7	Yes	Not tested	Yes	Not tested	Yes	Not tested	Chimp biased
2xHAR.222	2	Yes	Yes	Yes	Yes	No	No	Active
2xHAR.238	2,4, 7	Yes	Yes	Yes	Yes	Yes	Yes	Human biased
2xHAR.240	2	Yes	Yes	No	No	No	No	Inactive
2xHAR.243	2	No	No	No	No	No	No	Human biased
2xHAR.247	7	No	Not tested	No	Not tested	No	Not tested	Inactive
2xHAR.274	2	Yes	Yes	Yes	Yes	No	No	Human biased
2xHAR.283	2	No	No	No	No	No	No	Inactive
2xHAR.287	2	No	No	No	No	No	No	Inactive
2xHAR.332	2	No	No	No	No	No	No	Chimp biased
2xHAR.349	2	Yes	Yes	No	No	No	No	Active
2xHAR.374	2	Yes	Yes	No	No	No	No	Inactive
2xHAR.377	2	No	No	No	No	No	No	Inactive
2xHAR.384	2	No	No	No	No	No	No	Active
2xHAR.408	2	Yes	Yes	Yes	Yes	Yes	Yes	Human biased
2xHAR.427	7	No	Not tested	No	Not tested	No	Not tested	Inactive
2xHAR.447	7	Yes	Not tested	Yes	Not tested	Yes	Not tested	Inactive

2xHAR.482	2	No	No	No	No	No	No	Inactive
2xHAR.499	7	No	Not tested	No	Not tested	No	Not tested	Active
2xHAR.514	7	Yes	Not tested	Yes	Not tested	No	Not tested	Active
Total		31/52	20/30	19/52	13/30	14/52	9/30	
Active %		59.61%	66.67%	36.54%	43.33%	26.92%	30.00%	
Telecephalon: Brain %						73.68%	69.23%	

References

[1] Aldea et al 2021: https://doi.org/10.1073/pnas.2021722118

[2] Capra et al 2013: https://royalsocietypublishing.org/doi/10.1098/rstb.2013.0025

[3] Kamm et al 2013b: https://royalsocietypublishing.org/doi/10.1098/rstb.2013.0019

[4] Norman et al 2021: https://www.biorxiv.org/content/10.1101/2021.01.27.428524v1

[5] Oskenberg et al 2013: https://doi.org/10.1371/journal.pgen.1003221

[6] Prabhakar et al 2008: https://pubmed.ncbi.nlm.nih.gov/18772437/

[7] VISTA: http://enhancer.lbl.gov

Table S4. Telencephalon expression in HARs characterized in published mouse enhancer assays, related to Figure 5.

Supplemental Table Captions

Table S2. TF footprints called in each HAR using the human genome and human H3K27ac or the chimpanzee genome and chimpanzee H3K27ac, as well as footprints disrupted by human:chimpanzee variants, related to Figure 6.

Table S3. In vivo epigenetic datasets used in this study, related to Figure 2.

Table S5. Sei predicted epigenetic state changes for each HAR variant, related to Figure 4.

Table S6. Gene set enrichment analysis results, related to Figure 5.

Table S7. Annotations of HARs active in lentiMPRA, related to Figure 5.

Table S8. Primers used in this study, related to Methods.