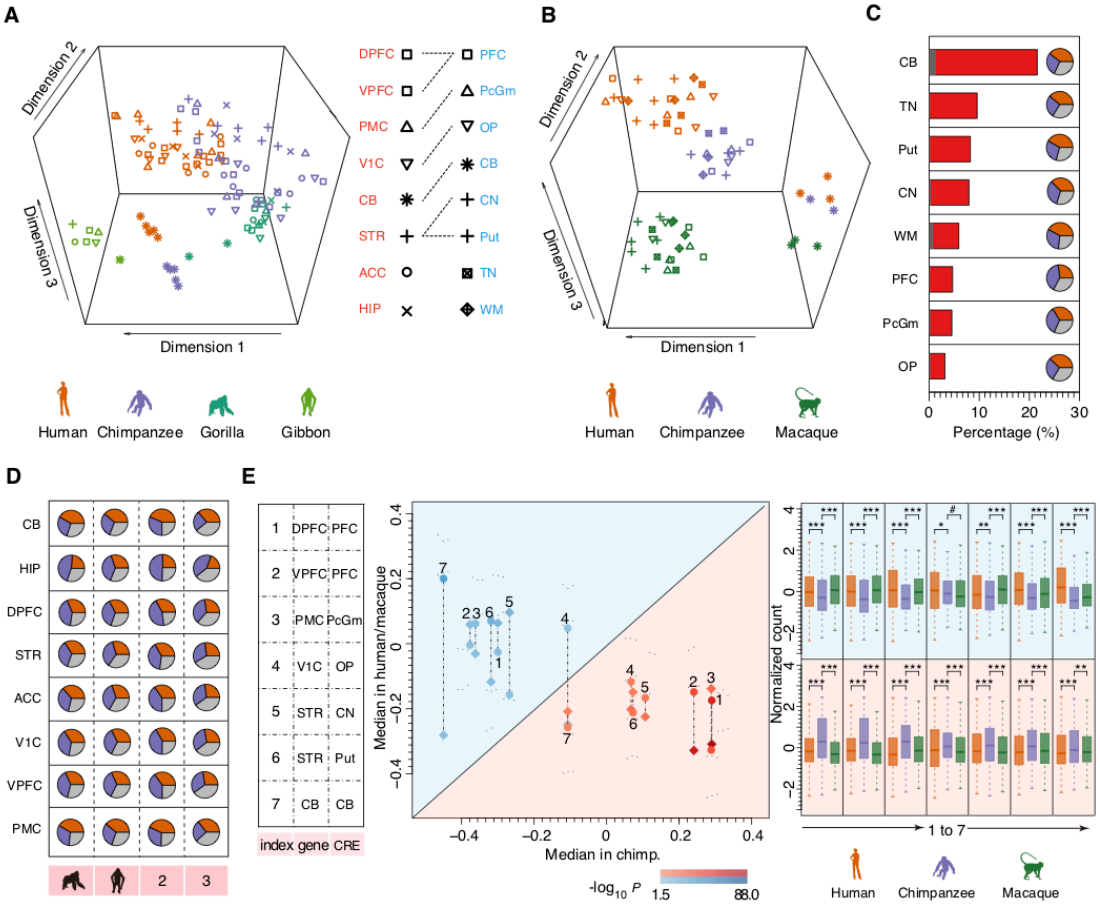
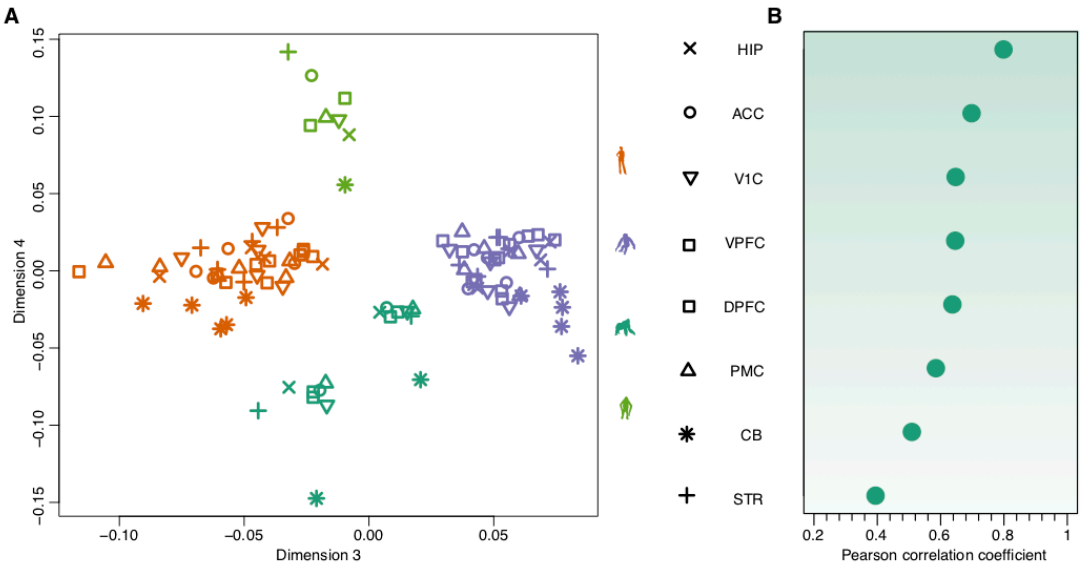
Supplemental Figures



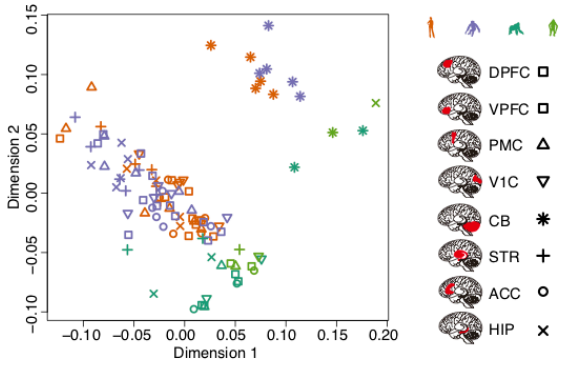
Supplemental Figure S1. Differences between humans and chimpanzees in transcriptome and epigenome. (*A*, *B*) Plots showing three dimensions resulting from a *t*-distributed stochastic neighbor embedding (t-SNE) analysis of transcriptome (*A*) and epigenome (*B*) data. Middle: symbols and labels of brain regions included in the transcriptome (red) and epigenome (blue) datasets. The dashed lines show brain regions shared between the datasets. Bottom: silhouette figures with colors marking species identity. The colors and shapes of the symbols on the t-SNE plots indicate species and brain region identities of the samples. (*C*) Left: the percentage of CREs showing significant coverage differences between humans and chimpanzees (red bars) and expected by chance (gray bars). Right: the pie charts showing the percentages of human-specific (orange), chimpanzee-specific (purple) and unassigned (gray) CRE

coverage differences sorted using macaque outgroup information. (*D*) Replicable lineage assignment under different outgroups and thresholds in transcriptome. The pie charts showing the percentages of human-specific (orange), chimpanzee-specific (purple) and unassigned (gray) gene expression differences sorted using only gorilla (first column) or gibbon (second column) outgroup information under the threshold of 2.5 fold-changes, or using combined gorilla and gibbon outgroup information under threshold of 2 fold-changes (third column) or 3 fold-changes (fourth column). (*E*) Normalized H3K27ac read counts for CREs located next to genes showing chimpanzee-specific gene expression changes. Left: the color of the dots shows the direction of the expression difference: red – higher in chimpanzee, blue – lower in chimpanzee. Each dot's shape represents non-chimpanzee species: circle – humans, diamond – macaques. Right: the box plots show the distributions of CRE coverage in humans (orange), chimpanzees (purple), and macaques (green) for the corresponding dots. The significance of the coverage differences between chimpanzees and non-chimpanzee primates in a one-sided Wilcoxon test was shown by each dot's color gradient (left) and the asterisks (right): *** – $p < 0.0005$, ** – $p < 0.005$, * – $p < 0.05$, $^{\#}$ – $p > 0.05$ as the gray dot shows. The numbers indicate brain regions in accordance with the index shown on left side.



Supplemental Figure S2. Assessment of the gibbon hippocampus sample identity based on its expression. (*A*) Species' identity assessment. The plot shows the third and the fourth dimensions of the multidimensional scaling (MDS) analysis of the transcriptome data. The colors and shapes of the symbols indicate the species and brain region identities of the samples, as illustrated by the legend on the right. DPFC: dorsolateral prefrontal cortex; VPFC: ventrolateral prefrontal cortex; PMC: premotor cortex; V1C: primary visual cortex; CB: cerebellum; STR: striatum; ACC: anterior cingulate cortex; HIP: hippocampus. (*B*) Brain region identity assessment. The mean of the Pearson correlation coefficients calculated between the gibbon hippocampus sample and the gorilla samples from each brain region based on the expression of the 561 hippocampal marker genes.

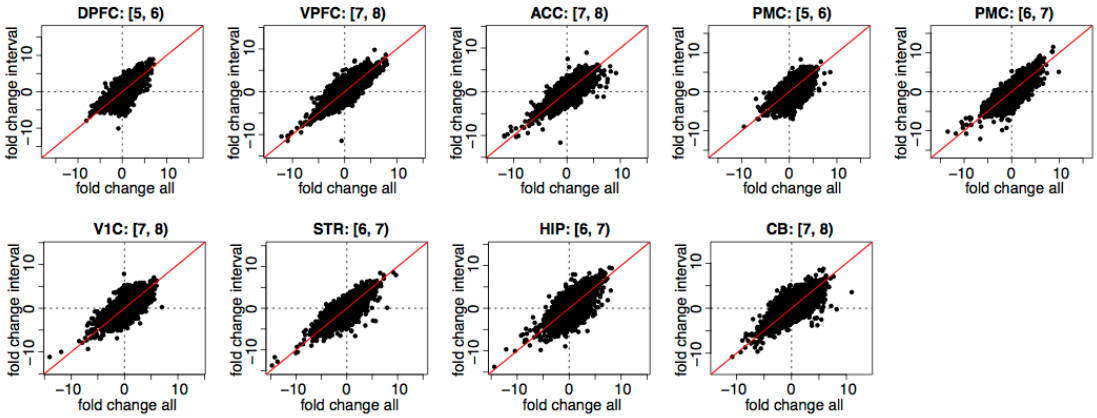
A



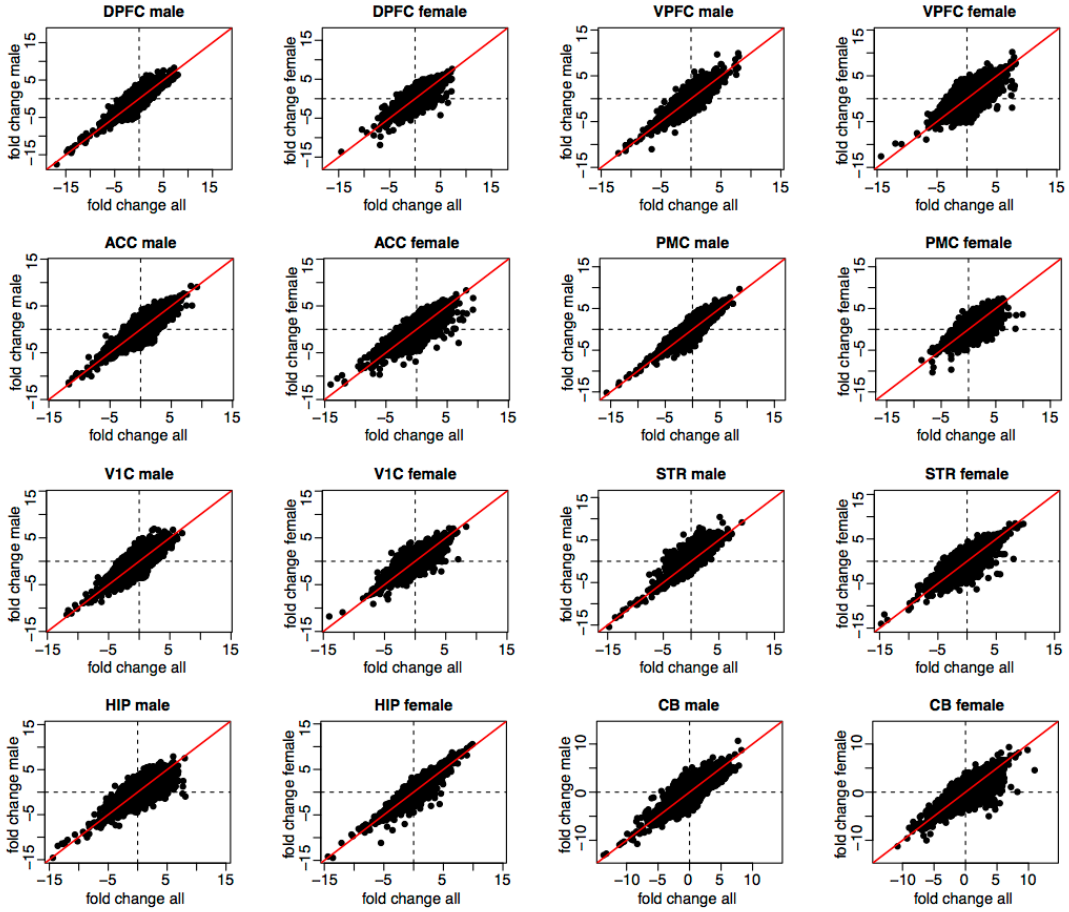
B

# DE genes without duplicate removal	# DE genes with duplicate removal	# overlap	% overlap
3,300	3,283	3,129	95.06
2,652	2,656	2,596	97.81
2,563	2,546	2,467	96.57
2,825	2,785	2,693	96.01
4,863	4,821	4,692	96.90
3,317	3,268	3,186	96.77
3,247	3,186	3,103	96.47
4,234	4,197	4,097	97.19

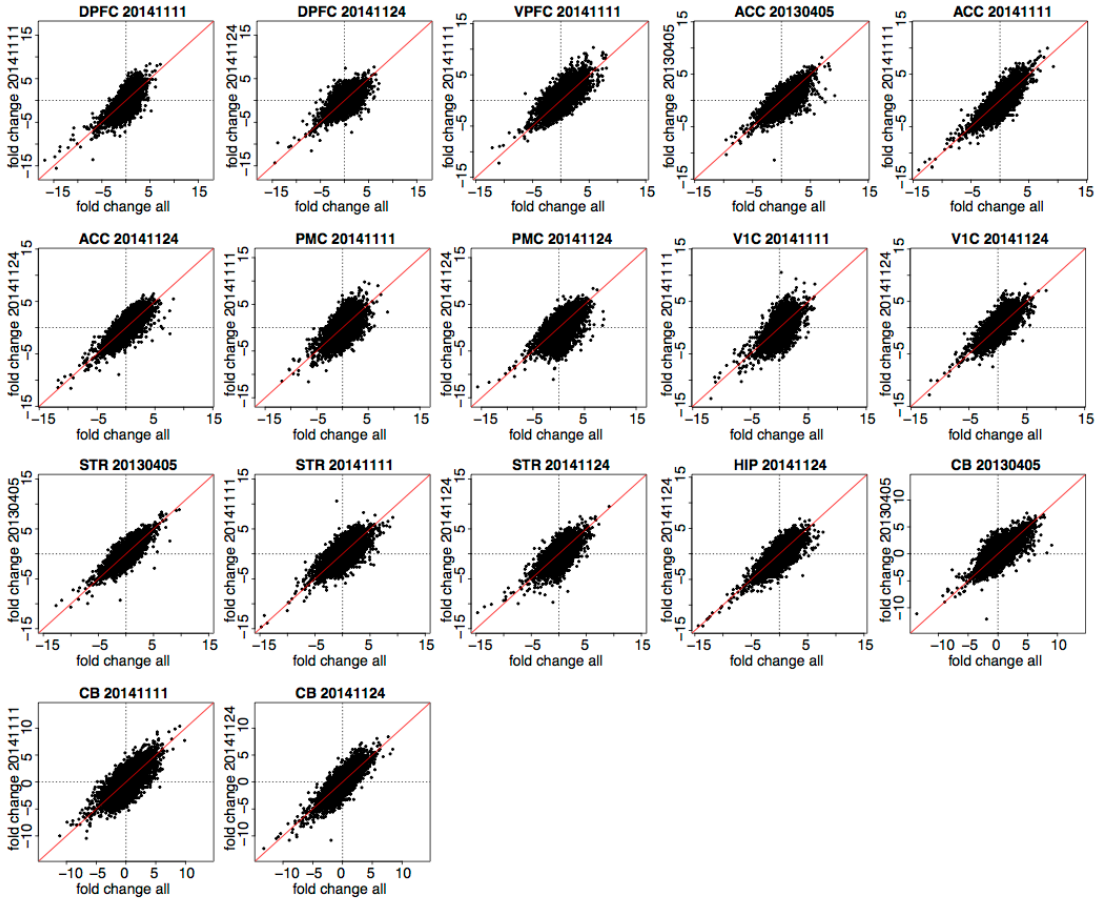
Supplemental Figure S3. PCR duplicates had no effect on the transcriptome comparison between humans and chimpanzees. (*A*) The plot shows two dimensions resulting from the multidimensional scaling (MDS) analysis of the transcriptome data without PCR duplicate removal step. The colors and shapes of the symbols indicate the species and brain region identities of the samples, as illustrated by the legend on the right. (*B*) The numbers and the overlap of genes differentially expressed between humans and chimpanzees in each brain region detected with and without PCR duplicate removal step.



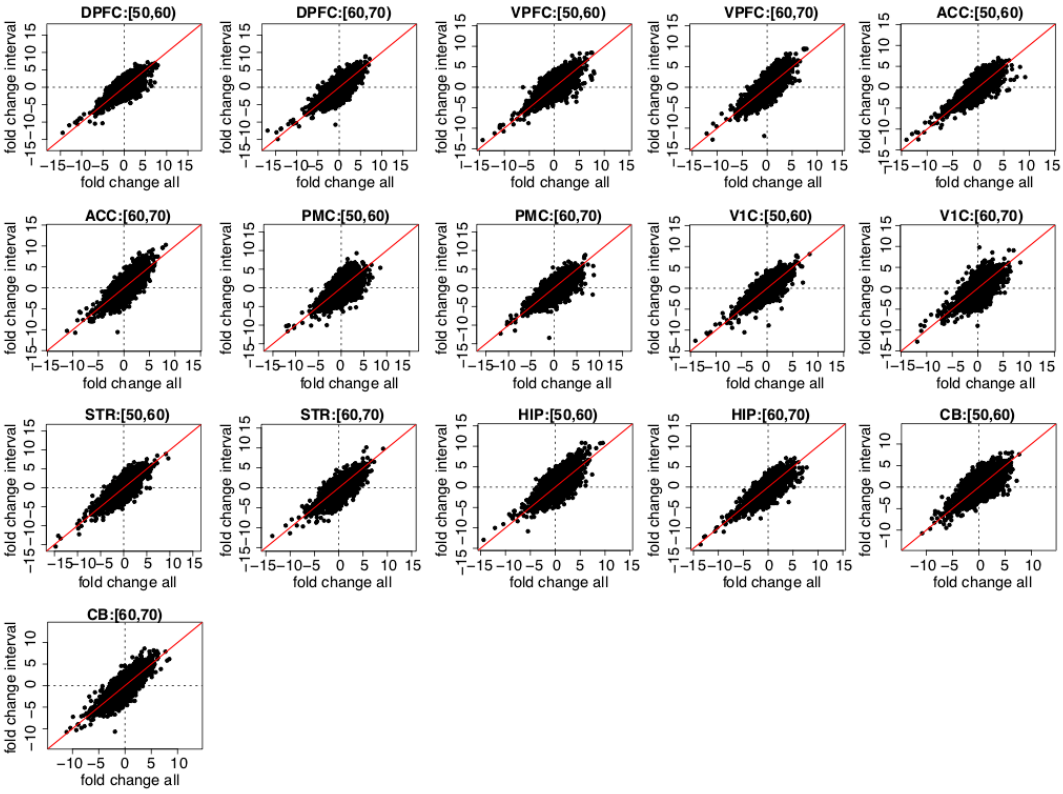
Supplemental Figure S4. RIN caused no bias in the transcriptome comparison between humans and chimpanzees. Each dot represents the fold change calculated between humans and chimpanzees based on all human and chimpanzee samples (x axis) as compared with the one based on samples within the same interval of RIN (y axis) in each brain region. Brain regions and RIN intervals are illustrated above the panel. The red line shows the reference line y = x.



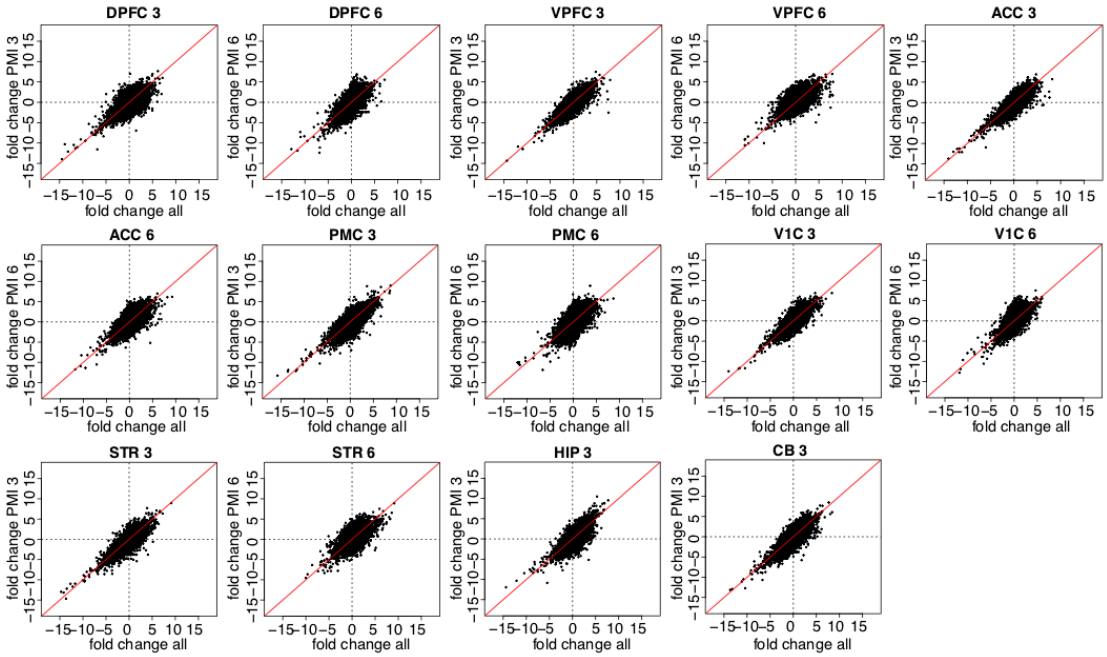
Supplemental Figure S5. Individual sexes caused no bias in the transcriptome comparison between humans and chimpanzees. Each dot represents the fold change calculated between humans and chimpanzees based on all human and chimpanzee samples (x axis) as compared with the one based on samples with the same sex (y axis) in each brain region. Brain regions and individual sexes are illustrated above the panel. The red line shows the reference line y = x.



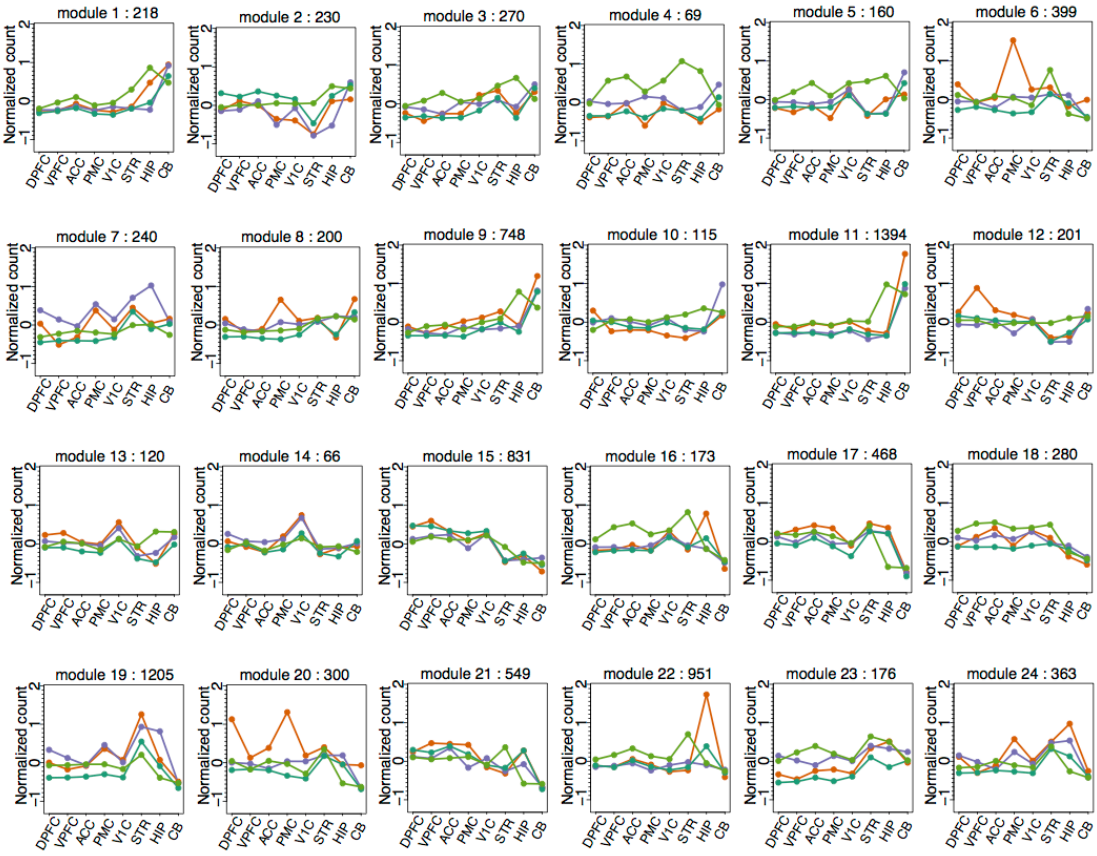
Supplemental Figure S6. The experimental batch caused no bias in the transcriptome comparison between humans and chimpanzees. Each dot represents the fold change calculated between humans and chimpanzees based on all human and chimpanzee samples (x axis) as compared with the one based on samples processed on the same experimental batch (y axis) in each brain region. Brain regions and the experimental batch are illustrated above the panel. The red line shows the reference line y = x.



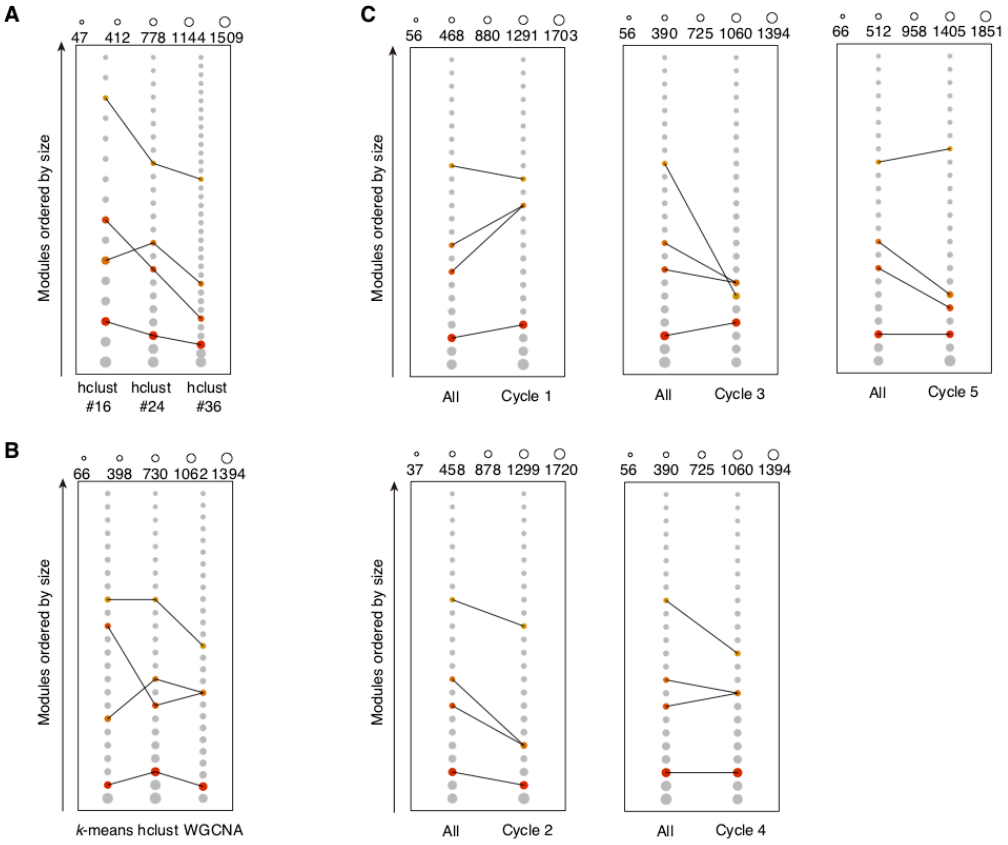
Supplemental Figure S7. Individuals' ages caused no bias in the transcriptome comparison between humans and chimpanzees. Each dot represents the fold change calculated between humans and chimpanzees based on all human and chimpanzee samples (x axis) as compared with the one based on samples within the same age interval (y axis) in each brain region. Brain regions and age intervals (in year unit) are illustrated above the panel. The red line shows the reference line y = x.



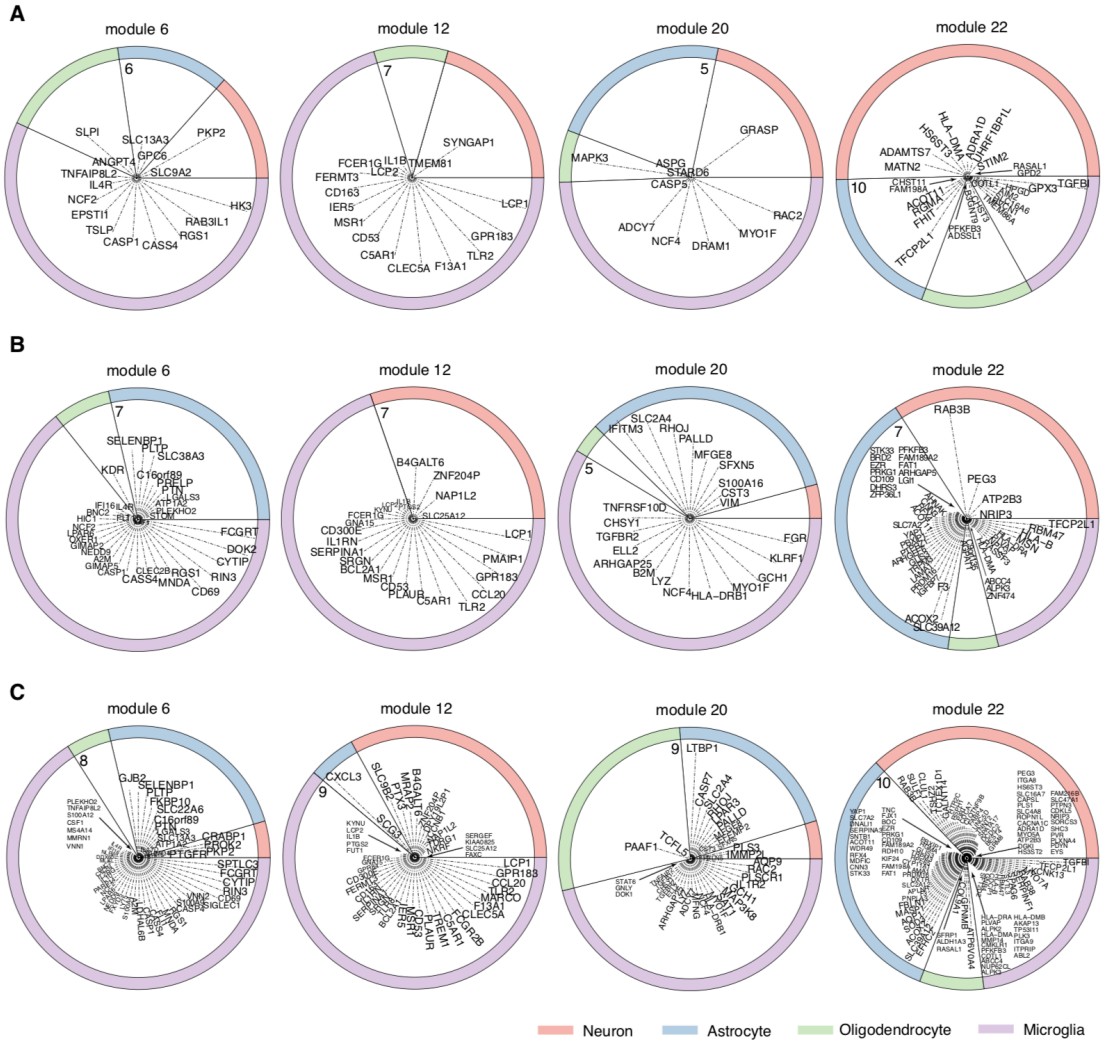
Supplemental Figure S8. Postmortem interval caused no bias in the transcriptome comparison between humans and chimpanzees. Each dot represents the fold change calculated between humans and chimpanzees based on all human and chimpanzee samples (x axis) as compared with the one based on samples within the same postmortem interval (y axis) in each brain region. Brain regions and postmortem intervals (in hour unit) are illustrated above the panel. The red line shows the reference line y = x.



Supplemental Figure S9. Gene expression profiles of the 24 modules derived from unsupervised hierarchical clustering of gene expression in humans. The dots show normalized read count averaged across genes within each module for respective species (orange – humans, purple – chimpanzees, dark green – gorillas, light green – gibbon). The number of genes within each module was shown above each panel.

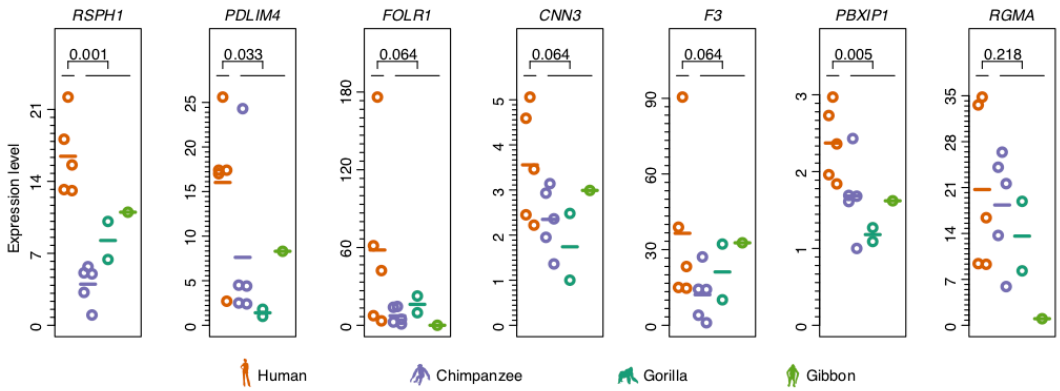


Supplemental Figure S10. Reproducibility of human-specific modules. (*A*, *B*, *C*) The four human-specific modules persisted despite module definition cutoffs under unsupervised hierarchical clustering (hclust) (*A*), clustering procedures like *k*-means clustering or signed WGCNA (*B*) or jackknife resampling (*C*). Each dot represents a module ordered by number of genes within the module under different clustering criteria. A bigger dot indicates a higher number of genes within the module, as illustrated above the panel. Red dots indicate the four human-specific modules. Edges denote significant gene set overlap between modules under different clustering criteria assessed by a hypergeometric test with Bonferroni-corrected *p*-values of less than 1E-10 (i.e. heavy overlap) shown for the four human-specific modules. The five cycles in panel *C* were obtained by a five-fold jackknife resampling: randomly classifying all samples into 5 equal-sized groups and removing one group in each cycle.

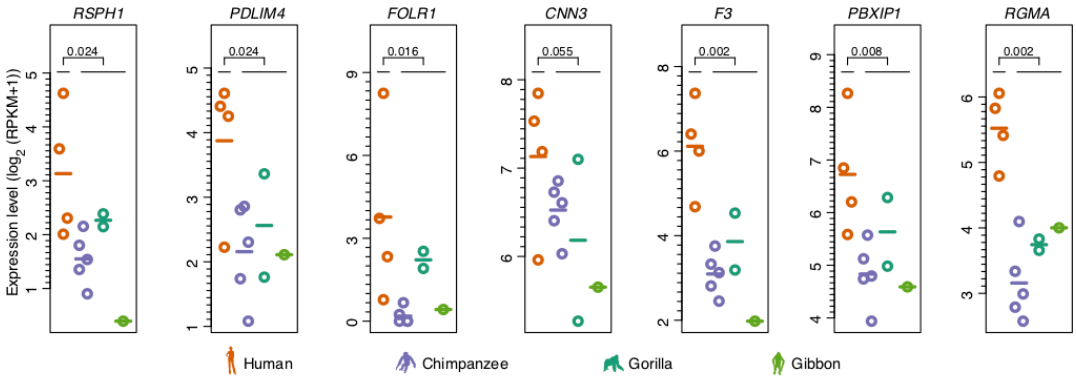


Supplemental Figure S11. The distribution of cell type marker genes within the four human-specific modules based on marker sets from Zeisel et al. 2015 (*A*), Darmanis et al. 2015 (*B*), and Zhang et al. 2016 (*C*). Each ring represents one human-specific module, showing the relative proportions of neuronal, astrocytic, oligodendrocytic, and microglial marker genes within the module with expression human-specificity ratio greater than one, normalized to the overall number of marker genes for each cell type. The corresponding marker gene names are shown within the rings, with the distance between each gene and the center representing the log2-transformed human-specificity ratio. The numbers show the maximum of the log2-transformed human-specificity ratio range within the modules. The human-specificity ratio was calculated as the ratio of the absolute expression difference between humans and the

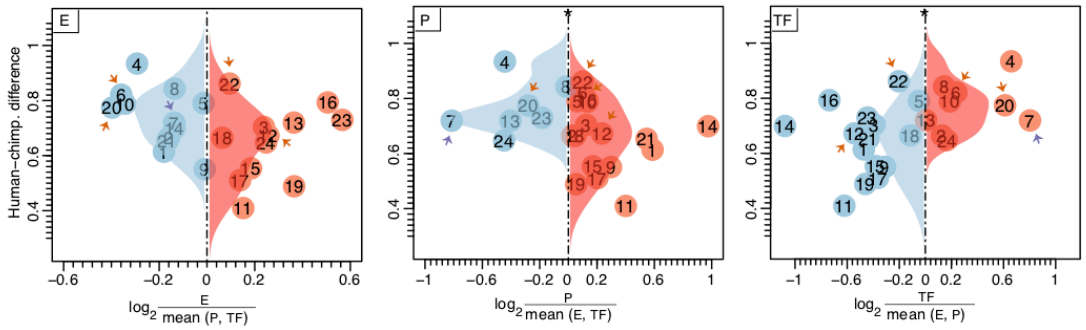
outgroup species to the absolute difference between chimpanzees and the outgroup species.



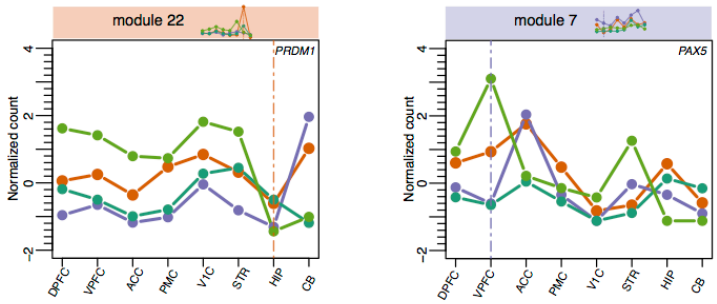
Supplemental Figure S12. Expression levels of the seven selected genes measured using RT-qPCR. The human sample with the highest expression level for a given gene was removed from the plot and statistical analysis. The colors represent species, as shown by the legend below. Each measurement represents a biological replicate (n = 5 for humans, n = 5 for chimpanzees, n = 2 for gorillas, and n = 1 for gibbon). The horizontal lines represent the mean of expression across biological replicates. The significance of expression differences between humans and the other three primate species is indicated by the *p*-value of a one-sided Wilcoxon test.



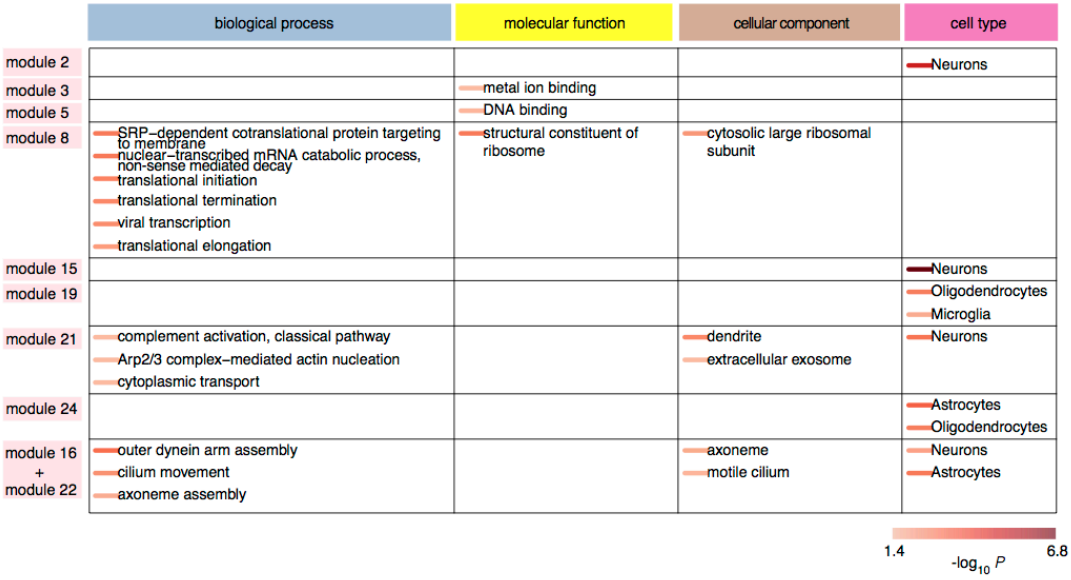
Supplemental Figure S13. The log2-transformed expression levels of the seven selected genes measured using RNA-seq data. The colors represent species, as shown by the legend below. Each measurement represents a biological replicate (n = 4 for humans, n = 5 for chimpanzees, n = 2 for gorillas, and n = 1 for gibbon). The horizontal lines represent the mean of expression across biological replicates. The significance of expression differences between humans and the other three primate species is indicated by the *p*-value of a one-sided Wilcoxon test.



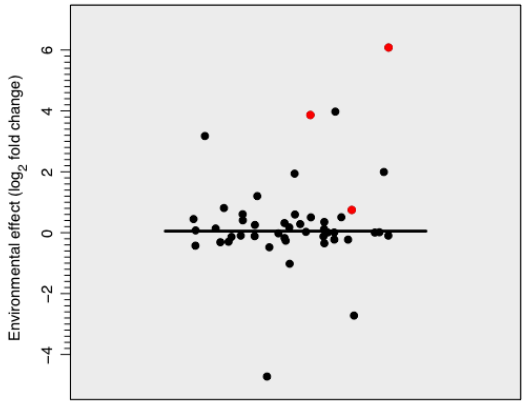
Supplemental Figure S14. Influence of the three regulatory mechanisms on spatial gene expression differences between humans and chimpanzees: enhancer CREs (E), promoter CREs (P), TF expression levels for TFs identified using binding sites present in distal enhancer CREs (TF). The dots represent modules. The horizontal axis shows the relative influence of a regulatory mechanism in each module. The vertical axis and the distributions show human-chimpanzee expression differences for modules with the major (red) and minor (blue) contribution of a given regulatory mechanism. The asterisks indicate the significance of the difference between two distributions (one-sided Wilcoxon test, * – $p < 0.05$). The arrows mark species-specific modules (orange – human, purple – chimpanzee).



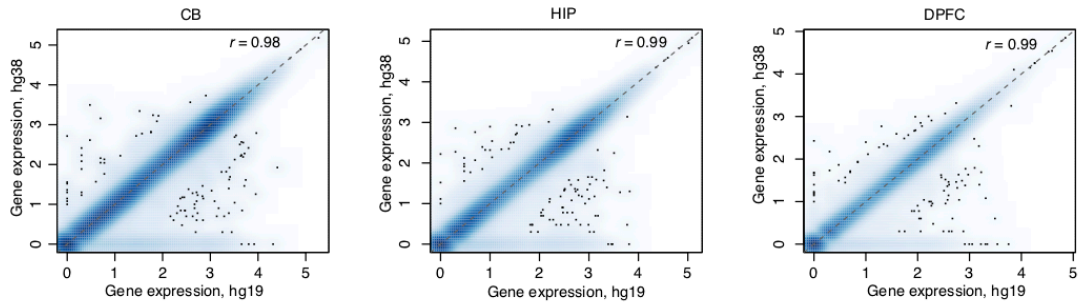
Supplemental Figure S15. Spatial expression patterns of TFs *PRDM1* and *PAX5*. Spatial expression patterns of *PRDM1* and *PAX5* predicted to negatively regulate target genes in modules 22 and 7 respectively. The colors represent species: orange – human, purple – chimpanzee, dark green – gorilla, light green – gibbon. The module expression patterns are shown above the panels. The dotted vertical lines indicate the brain regions responsible for species-specificity of the modules.



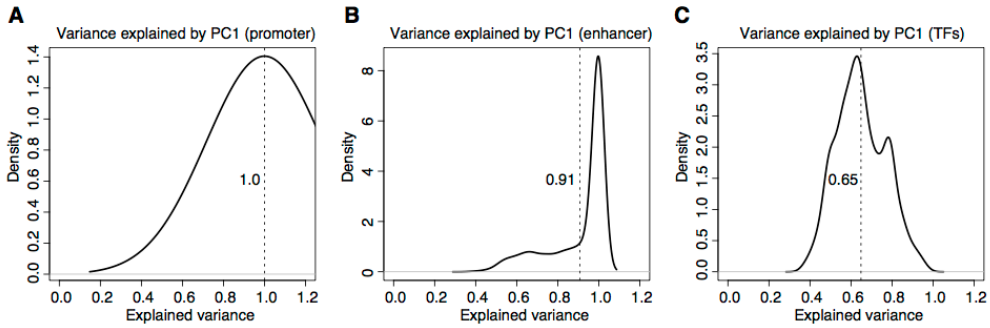
Supplemental Figure S16. GO terms and cell type markers significantly enriched in genes within non-species-specific co-expression modules. The shade of the color bars next to term and cell type names indicates *p*-values of hypergeometric tests after BH correction, as shown by the bar below.



Supplemental Figure S17. The environmental transition effect on the expression of the 46 microglial marker genes used to specify the three human-specific modules. Each dot represents a gene. The vertical axis shows the log2-transformed values of the expression level fold change after the environmental transition according to Gosselin et al. 2017. The black line indicates the median value. The horizontal axis is added to improve data representation. The red dots denote the three genes with expression significantly affected by the environmental transition according to Gosselin et al. 2017.



Supplemental Figure S18. Scatter plots showing the log10-transformed sequence read counts mapped to the human genes based on the hg19 and hg38 genome versions. The color gradient represents the data point density calculated as the bivariate normal transformation kernel of the data points. The top 100 points within the lowest density areas are shown by the dots. The dotted line shows the y = x diagonal. The Pearson correlation coefficients (*r*) are shown within each panel.



Supplemental Figure S19. The adequacy of the first principal component (PC1) in representing the spatial profile of each kind of regulator coupled to the gene. (*A*-*C*) The distributions show the proportion of variance for PC1's spatial profile explaining the whole profiles of promoter CREs (*A*), enhancer CREs (*B*) or TFs (*C*) coupled to the gene in a principal component analysis (PCA). Dotted lines represent the mean values of the distributions. Note that the number of promoter CREs for every gene is one as shown in *A*.