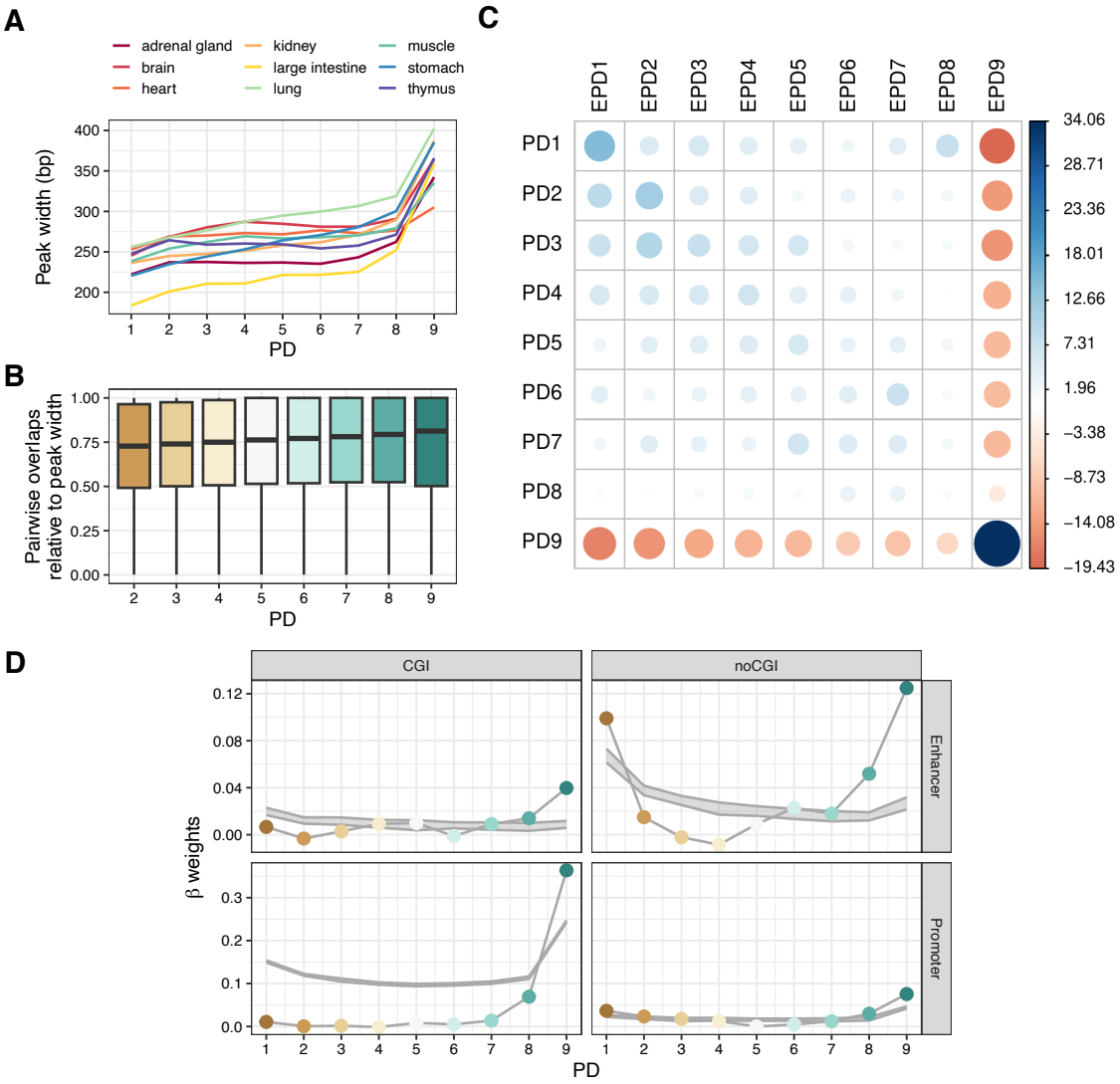
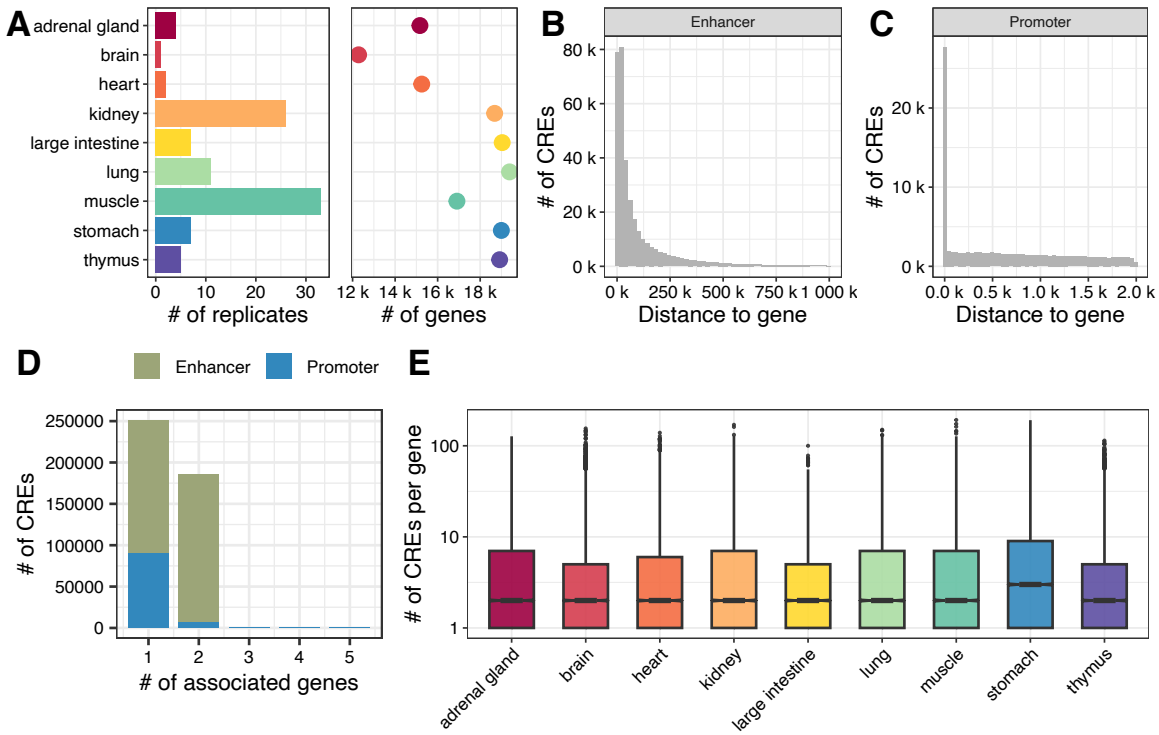
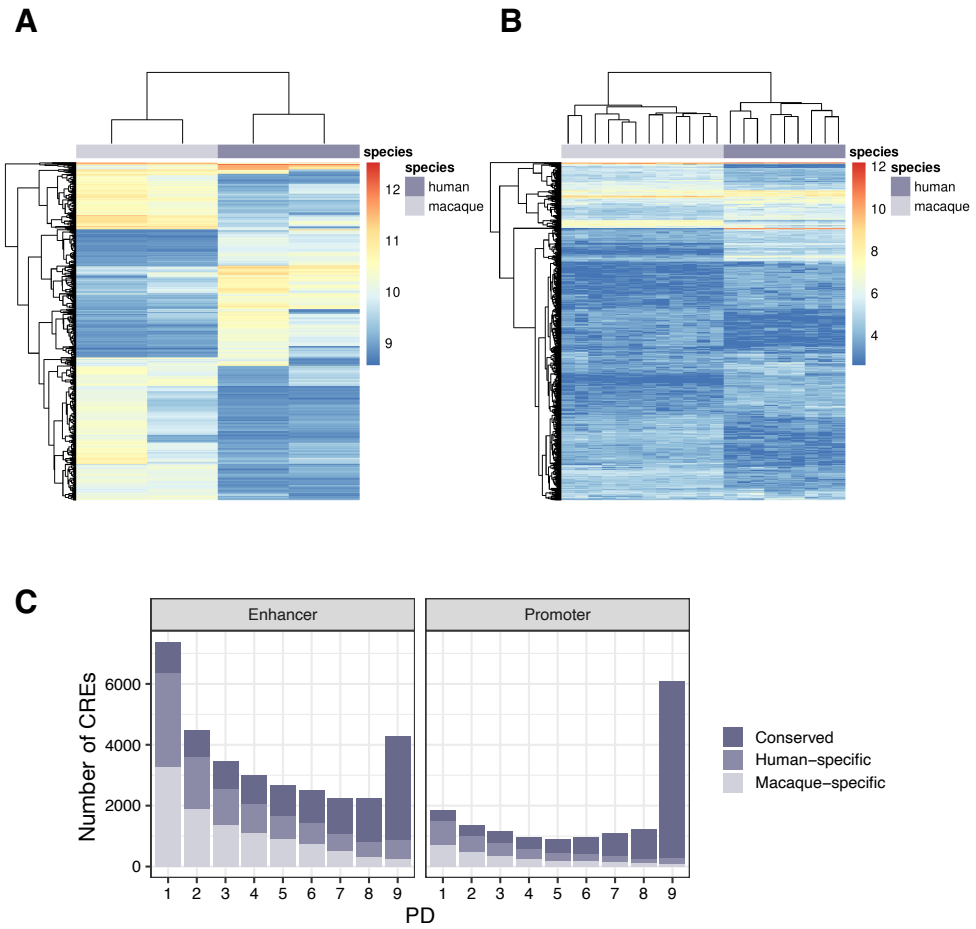
Supplemental Figures



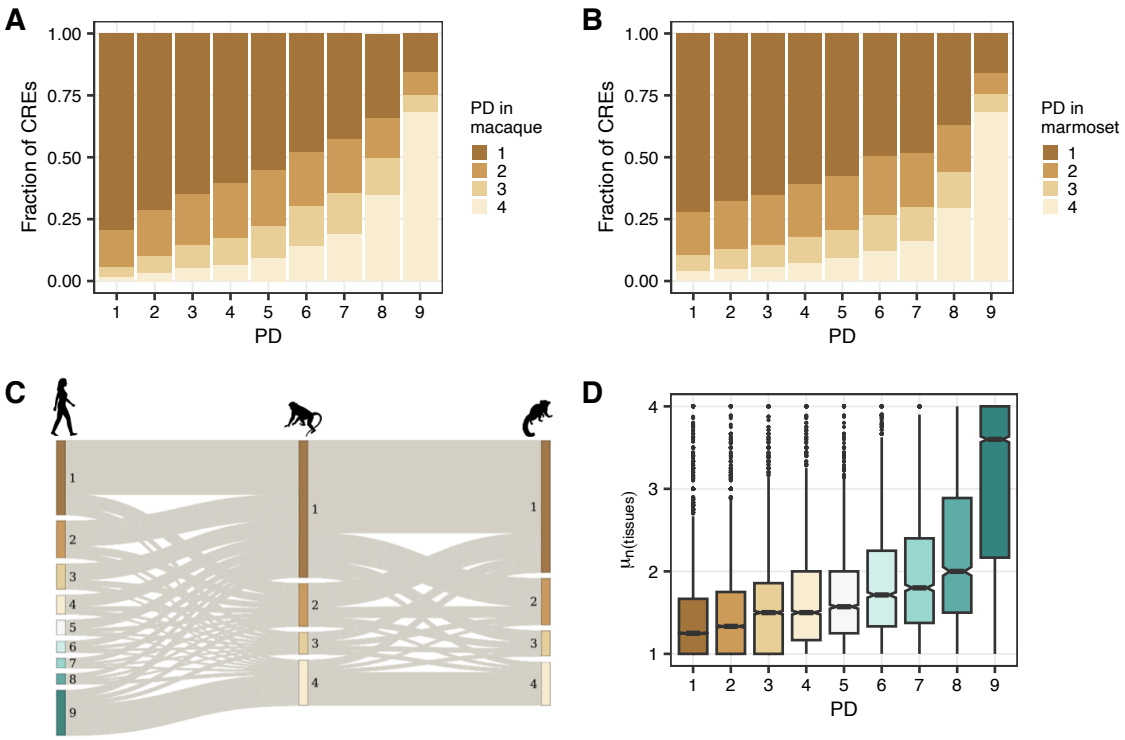
Supplemental Figure S1. Peak widths per tissue and overlaps across tissues. **A.** Average peak width per tissue across specificity groups prior to cross-tissue merging. **B.** Pairwise overlap fraction between overlapping peaks that were later merged into the same CRE. **C.** Pearson residuals between promoter PD and the associated gene PD (χ^2-test p-value<2.2e-16, χ^2=5286, df=64). **D.** Observed coefficients for the CREs from different PDs (thin line with colored points) are different from the control estimates (90% CI, in gray) where PD labels were shuffled 30× across the CREs.



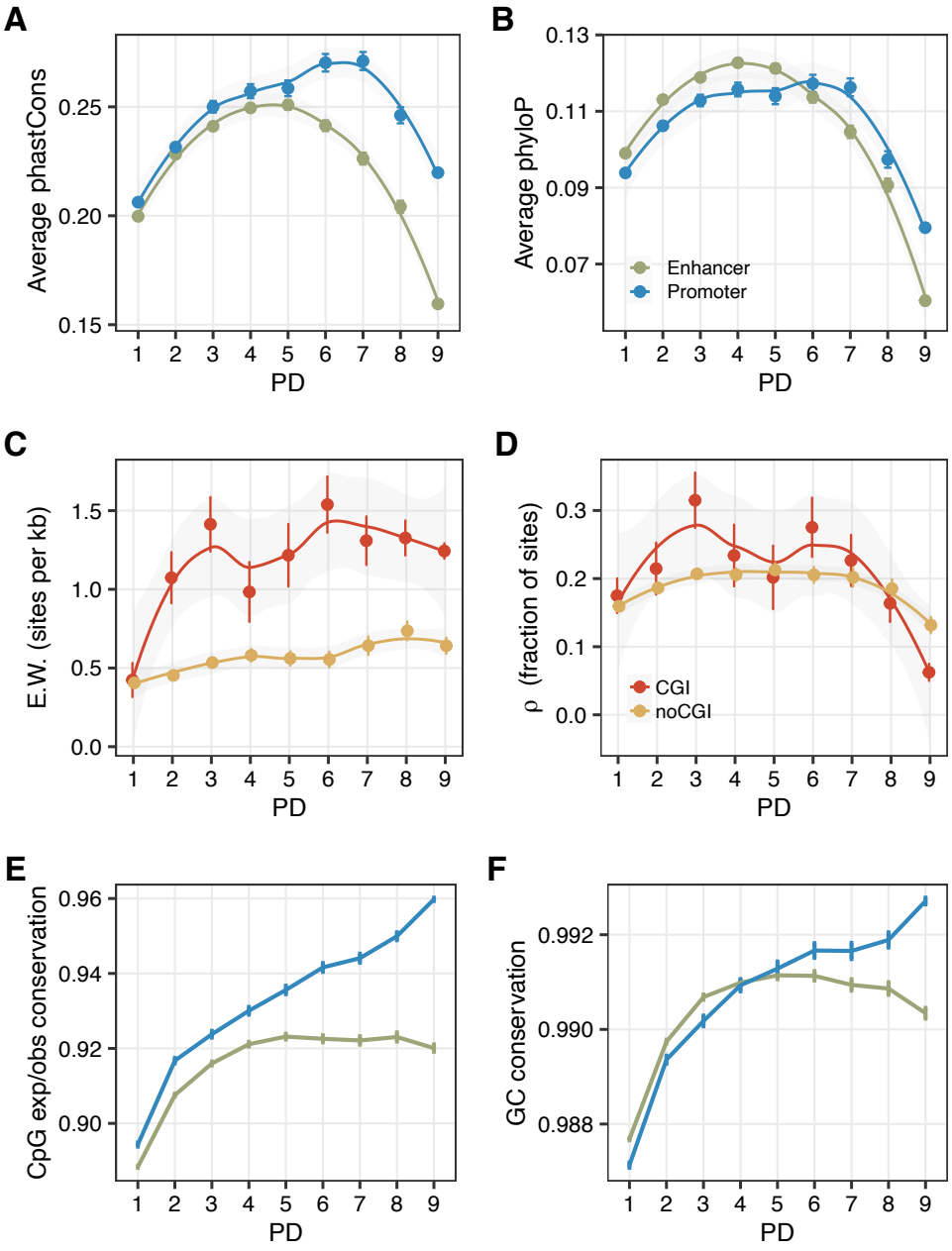
Supplemental Figure S2. CRE to gene association across tissues. **A.** Expression data overview: Number of replicates and number of expressed genes for each tissue. **B-C.** Distance distribution of CREs annotated as enhancers (**B**) and promoters (**C**) to their closest gene. **D.** Number of associated genes per CRE. Enhancers were associated with up to 2 genes, promoters up to 5. **E.** The associated number of CREs per gene is comparable across tissues.



Supplemental Figure S3. Cross-species NPC expression and accessibility. **A.** Heatmap of the top 1000 most variable CREs across the ATAC-seq data in terms of their openness (Euclidean distance, method: Complete). **B.** Heatmap of the top 1000 most variable genes across the RNA-seq data (Euclidean distance, method: Complete). **C.** Number of NPC ATAC-seq peaks that overlap a CRE from a certain PD category. If a peak is present in NPCs from both species, it is labeled as conserved, while species-specific labels mean openness in the NPCs from only one species.

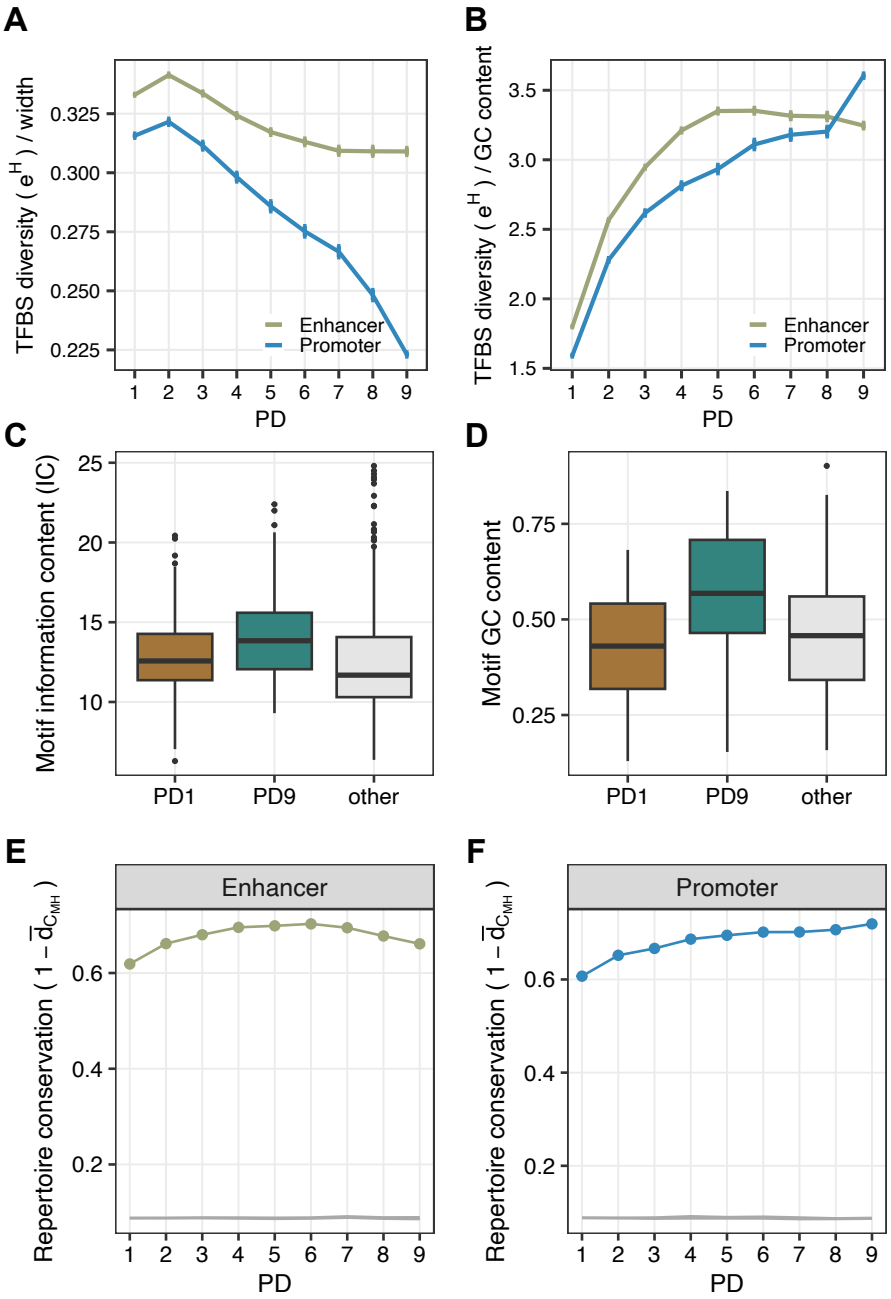


Supplemental Figure S4. Agreement in PD between human CREs (PD 1-9) in our study and the CREs in other mammals (PD 1-4) from Roller et al., 2021 where regulatory activity was annotated using histone modification ChIP-seq across 4 tissues: Brain, liver, muscle, testis. Our PD annotations were reciprocally liftOvered to the genome of rhesus macaque genome. Only 1-to-1 corresponding CREs were used for all analyses (47,377 CREs). Pairwise CRE orthology annotations from macaque to all other species from Roller et al., 2021 were used to investigate PD in other mammals and estimate CRE evolutionary age. **A.** PD across 9 human tissues compared to the PD of the overlapping orthologous CREs in macaque. The color indicates the number of tissues in which the CRE was detected as active in macaque. **B.** PD across nine human tissues compared to the PD of the overlapping orthologous CREs in marmoset. The color indicates the number of tissues in which the CRE was detected as active in marmoset. **C.** Sankey plot showing PD agreement across the three primates, human, rhesus macaque and marmoset. There is an overall agreement in PD between species, particularly at the PD1 and PD9 (i.e. PD4 in macaque and marmoset) categories. **D.** The average PD (i.e., number of tissues with activity) for each detected orthologous CRE across all 10 mammalian species from Roller et al., 2021, stratified by our human PD assignment. A value close to 4 means that in most species the orthologous CRE was active in all tissues. Thus, pleiotropic CREs from humans also tend to be more pleiotropic in other mammals.

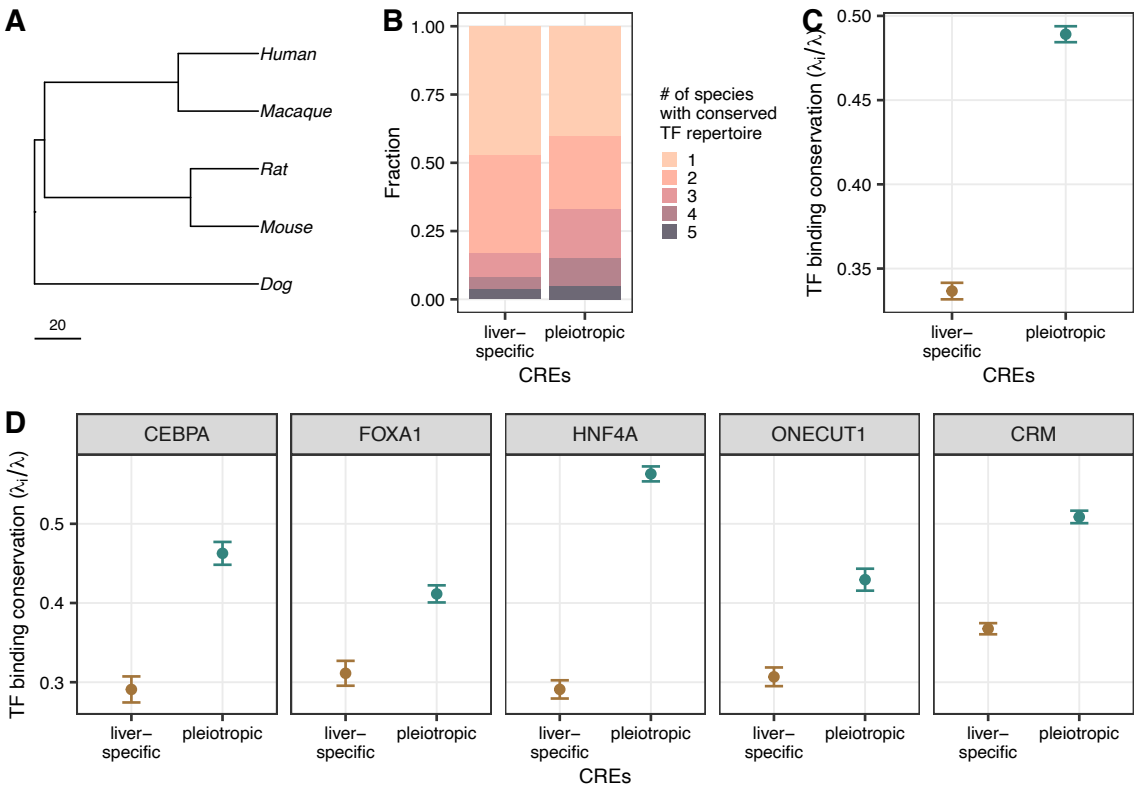


Supplemental Figure S5. Evolutionary sequence analysis of CREs across tissue specificity groups. **A.** phastCons conservation scores based on a 10-species primate tree. **B.** PhyloP conservation scores based on a 10-species primate tree. **C.** Weak negative selection patterns between CGI and non-CGI CREs suggest an increasing weak selection with higher PD in both groups. **D.** (Strong) negative selection patterns between CGI and non-CGI CREs show similar trend across PDs. **E.** Absolute pairwise similarity in CpG expected/observed ratio between human and cynomolgus macaque orthologous CREs. **F.** Absolute pairwise similarity in GC content between human and cynomolgus macaque orthologous CREs. **A-F.** Depicted are mean estimates +/- SEM.

.



Supplemental Figure S6. TFBS repertoire diversity and conservation. **A.** TFBS diversity normalized by CRE width. Depicted are mean values +/- SEM. **B.** TFBS diversity normalized by GC content. Depicted are mean values +/- SEM. **C.** Motif information content of the PD1-enriched, PD9-enriched and other motifs, classified as in Fig. 2C. **D.** Motif GC content of the PD1-enriched, PD9-enriched and other motifs, classified as in Fig. 2C. **E,F.** 95% CIs of the average TFBS repertoire conservation when shuffling macaque CRE identifiers within the respective PD class 10× (gray line). Random CRE similarity is below 10% and does not increase with PD. In comparison, the real observed enhancer and promoter repertoire conservation is depicted in green (**E.**) and blue (**F.**), respectively.



Supplemental Figure S7. TF binding conservation based on ChIP-seq data across five mammalian liver tissues for four TFs from Ballester et al., 2014. Liver-specific (PD1) and pleiotropic (PD9) elements were identified by intersecting our data and PD annotations for liver CREs from Roller et al., 2021 (see Methods). **A.** The five mammalian species that were included in Ballester et al., 2014 are human, rhesus macaque, rat, mouse and dog. The phylogeny and the total branch length λ was calculated by pruning the mammalian tree from Bininda-Emonds et al., 2007. Scale: Million years ago. **B.** Liver-specific and pleiotropic CRE positions were intersected with recorded four TF binding events in at least one species. The provided cross-species binding conservation annotations from Ballester et al., 2014 were used to quantify the number of species in which a binding event is conserved across the orthologous sequences. PD1 CRE sequences show higher proportions of species-specific and two-species binding events than PD9 CREs, suggesting lower transcription factor repertoire conservation. n(PD1)=6,300, n(PD9)=14,442. **C.** TF binding conservation across the phylogeny is higher in PD9 than PD1 elements (Wilcoxon rank sum test p-value$< 2.2^{-16}$). For each binding event i, branch length λ_i among the species with binding over the total branch length of the phylogeny λ in A was calculated. Depicted are mean values +/- SEM. **D.** TF binding conservation split into individual TF and *cis*-regulatory module (CRM) annotations. Conservation is consistently higher in PD9 than PD1 (all Wilcoxon rank sum test BH-adjusted p-values< 0.001). Depicted are mean values +/- SEM.