

Supplementary Figures and Table

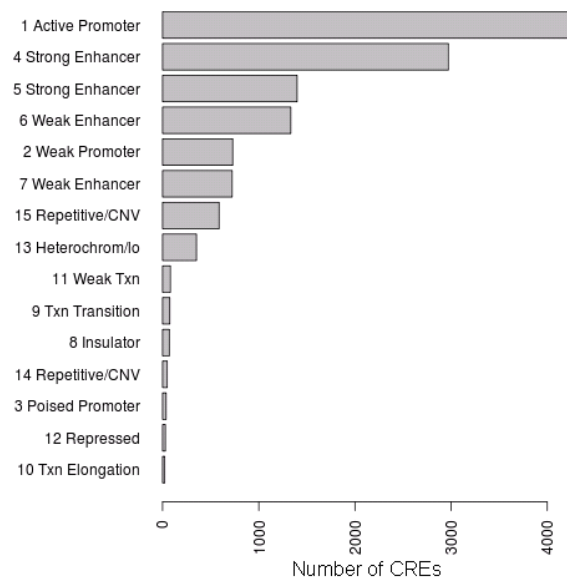


Figure 1: Frequency of each chromatin state at detected CREs. In comparison to the portion of the genome cover (<1% for promoter and enhancer states versus >70% for inactive state 13), promoter and enhancer states are enriched at CREs.

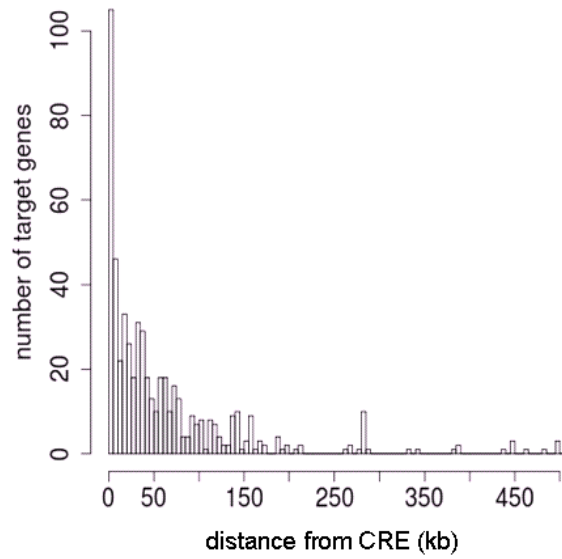


Figure 2: Distribution of target genes relative to distance from CREs. Target genes considered distal to a CRE were identified as being further than 150kb away.

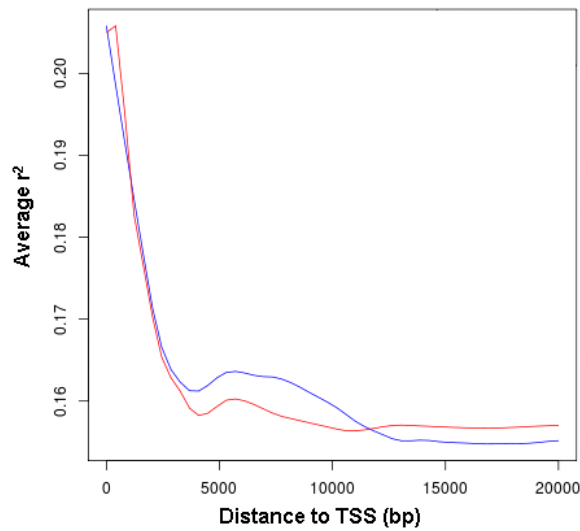


Figure 3: Assessment of gene expression prediction for different CRE positions relative to the TSS. The mean r^2 between observed and predicted expression decreases as we consider CREs further away from the TSS (blue line). Accuracy of prediction also decreases if the CRE described in the model is the closest to the TSS (red line).

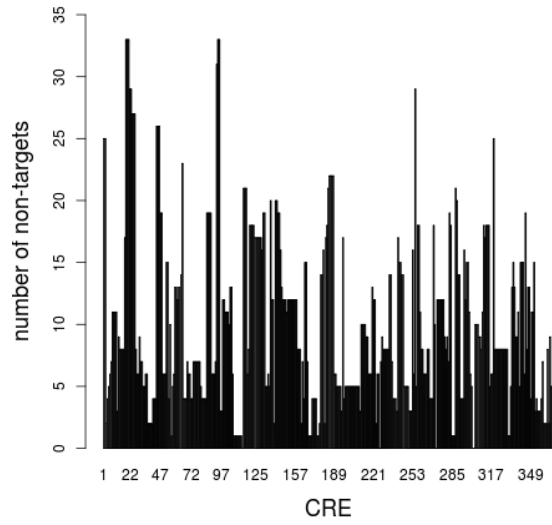


Figure 4: Distribution of non-target genes mapped to CREs. Genes 1Mb upstream and downstream of each CRE were analyzed. For training and testing of target gene prediction, genes not associated with any SNP in the CREs are classified as non-targets.

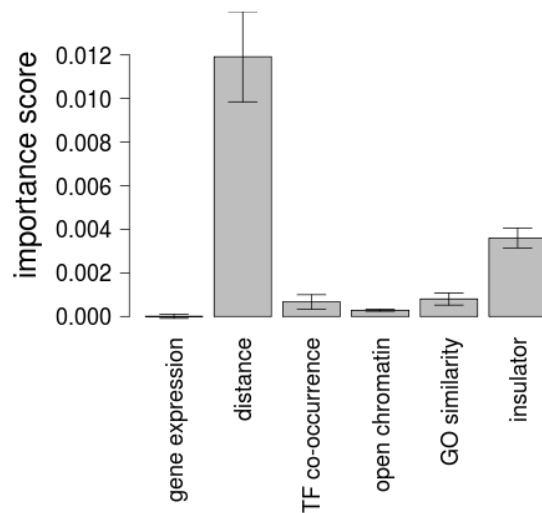


Figure 5: Conditional permutation importance scores of the six features used in the random forests classifier. Mean scores and standard errors from the different training sets used in cross validation are shown.

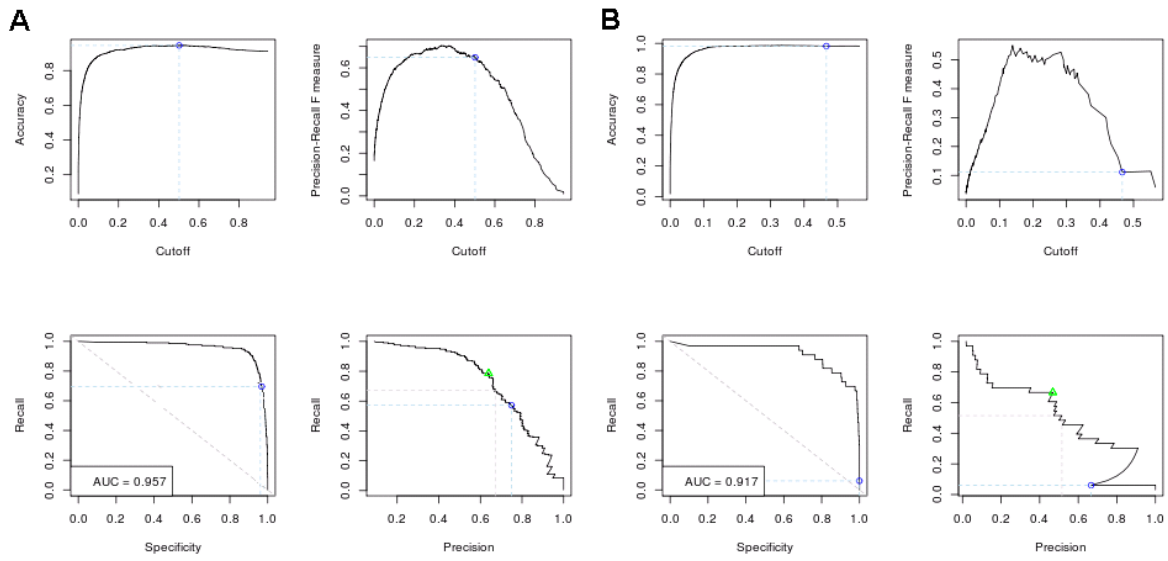


Figure 6: ROC curves for (A) classification of genes within 1Mb of CREs and (B) genes at least 150kb from CREs are compared to other measures of classifier performance. Dotted lines indicate the target probability cutoff $\Theta = 0.5$.

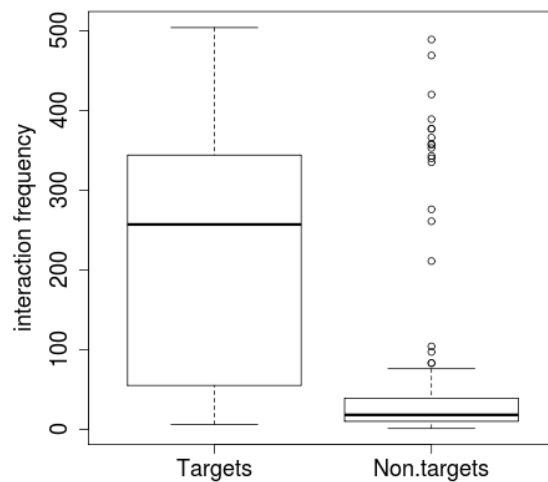


Figure 7: Comparison of intrachromosomal interactions detected by HiC between CREs and the two classes of genes. Genomic regions containing target genes have a higher frequency of being ligated (interaction frequency) to regions containing CREs.

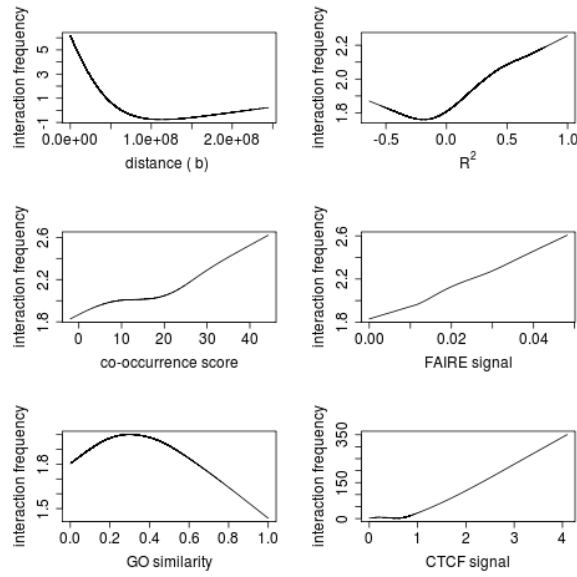


Figure 8: Features used in the classification of genes in relation to the chromosomal interactions between loci containing the CRE and gene. The average number of observed contacts estimated by ligation products (interaction frequency) between the locus containing the CRE and the locus containing the gene is compared to variations in genomic distance, gene co-expression (R^2), TF co-occurrence (co-occurrence score), open chromatin (FAIRE signal), gene functional similarity (GO similarity), and insulators (CTCF signal).

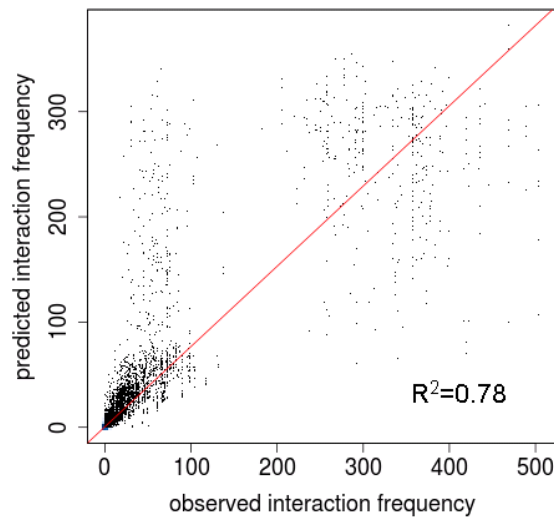


Figure 9: Predicted versus observed number of intrachromosomal interactions using the random forests classifier. We extracted features for each gene-CRE pair and predicted the interaction frequency of the gene's locus with the CRE's locus.

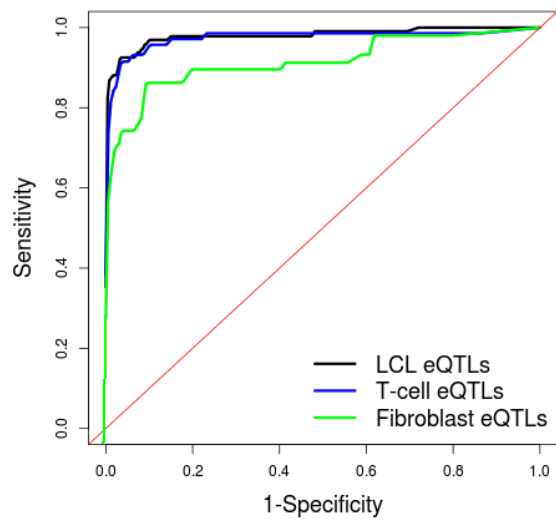


Figure 10: Comparison of ROC curves for predicting targets of eQTLs in LCLs (black), T-cells (blue), and fibroblasts (green). Each curve represents the average of ROC curves from 10-fold cross-validation on eQTL data from each cell type.

Table 1: Transcription factor ChIP-seq data downloaded from ENCODE

Source	TFs	Number of binding sites
Hudson Alpha	ATF3	2028
Hudson Alpha	BCL11A	25330
Hudson Alpha	BCL3	22503
Hudson Alpha	EGR1	487
Hudson Alpha	ELF1	17931
Hudson Alpha	ETS1	7544
Yale/UC-Davis/Harvard	FOS	3284
Hudson Alpha	GABP	2384
Hudson Alpha	IRF4	22281
Yale/UC-Davis/Harvard	JUN	82655
Yale/UC-Davis/Harvard	JUND	12958
Yale/UC-Davis/Harvard	MAX	6546
Hudson Alpha	MEF2A	18000
Yale/UC-Davis/Harvard	MYC	57629
Yale/UC-Davis/Harvard	NFKB	53740
Hudson Alpha	NRSF	2142
Hudson Alpha	PBX3	17435
Hudson Alpha	PU1	44755
Yale/UC-Davis/Harvard	RAD21	42564
Hudson Alpha	RXRA	5398
Hudson Alpha	SP1	26121
Hudson Alpha	SRF	658
Hudson Alpha	TAF1	4666
Hudson Alpha	TCF12	20356
Yale/UC-Davis/Harvard	TR4	10617
Yale/UC-Davis/Harvard	YY1	3437
Hudson Alpha	ZBTB33	2282
Hudson Alpha	ZEB1	4067
Yale/UC-Davis/Harvard	ZZZ3	8439
University of Washington	CTCF	69461

Table 2: CRE-target interactions in CD4+ T cells detected by ChIA-PET. Interaction between loci containing the CRE and target gene is measured by the number of paired-end reads mapped to the loci.

SNP	element chr	CRE start	CRE end	target gene	mapped tags
rs817353	chr20	62081649	62083030	NCRNA00176	8
rs817349	chr20	62080193	62080756	NCRNA00176	8
rs11680096	chr2	27127112	27127937	KHK	8
rs11126836	chr2	27126608	27126608	KHK	8
rs41417151	chr1	231126318	231126413	NTPCR	6
rs3812472	chr8	125555838	125556299	TRMT12	8
rs11778396	chr8	125556942	125556942	TRMT12	8
rs3740286	chr10	90740996	90741708	ACTA2	4
rs12359362	chr10	90742786	90743184.5	ACTA2	4
rs6480616	chr10	73727066	73728373	DDIT4	6