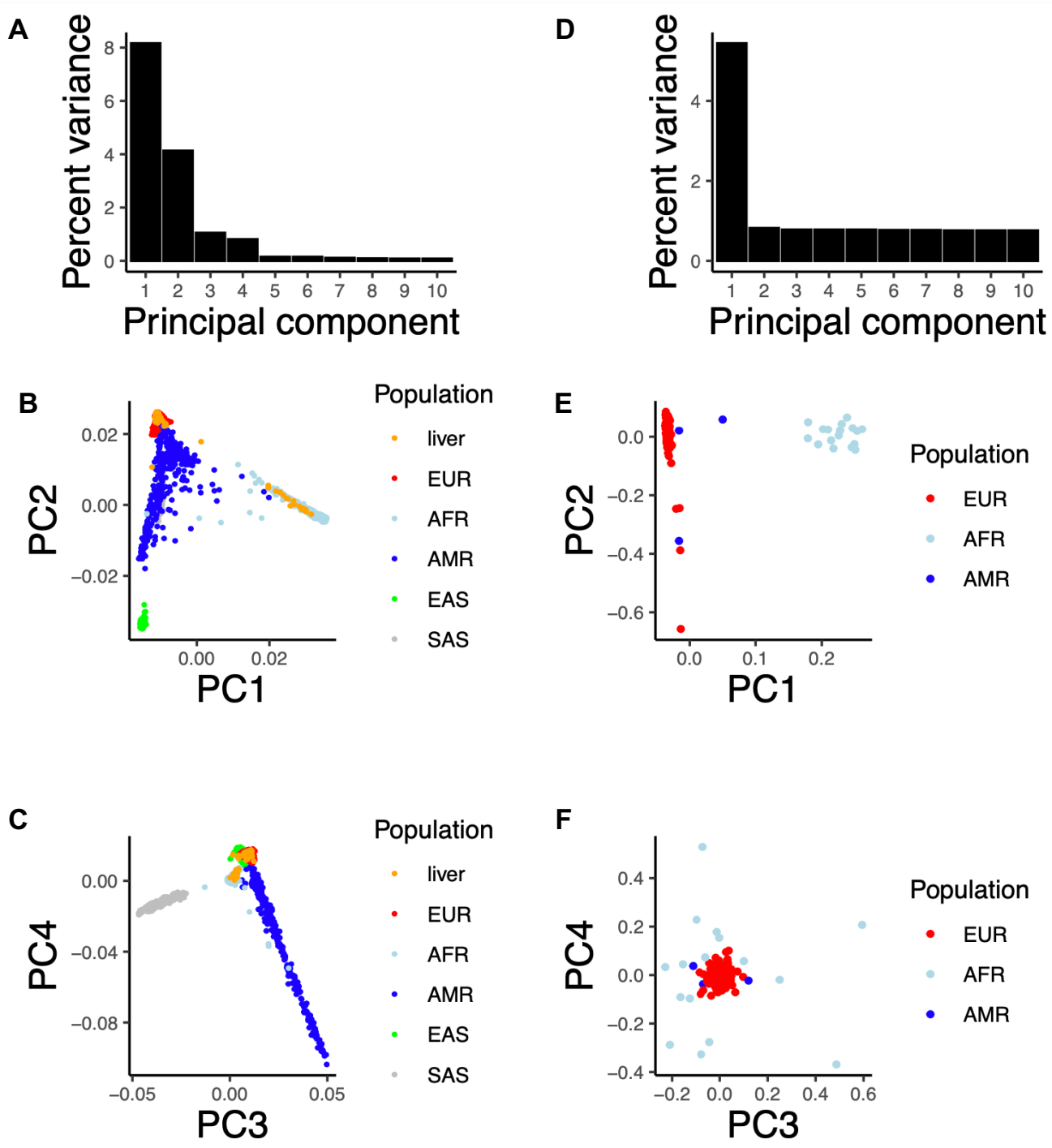
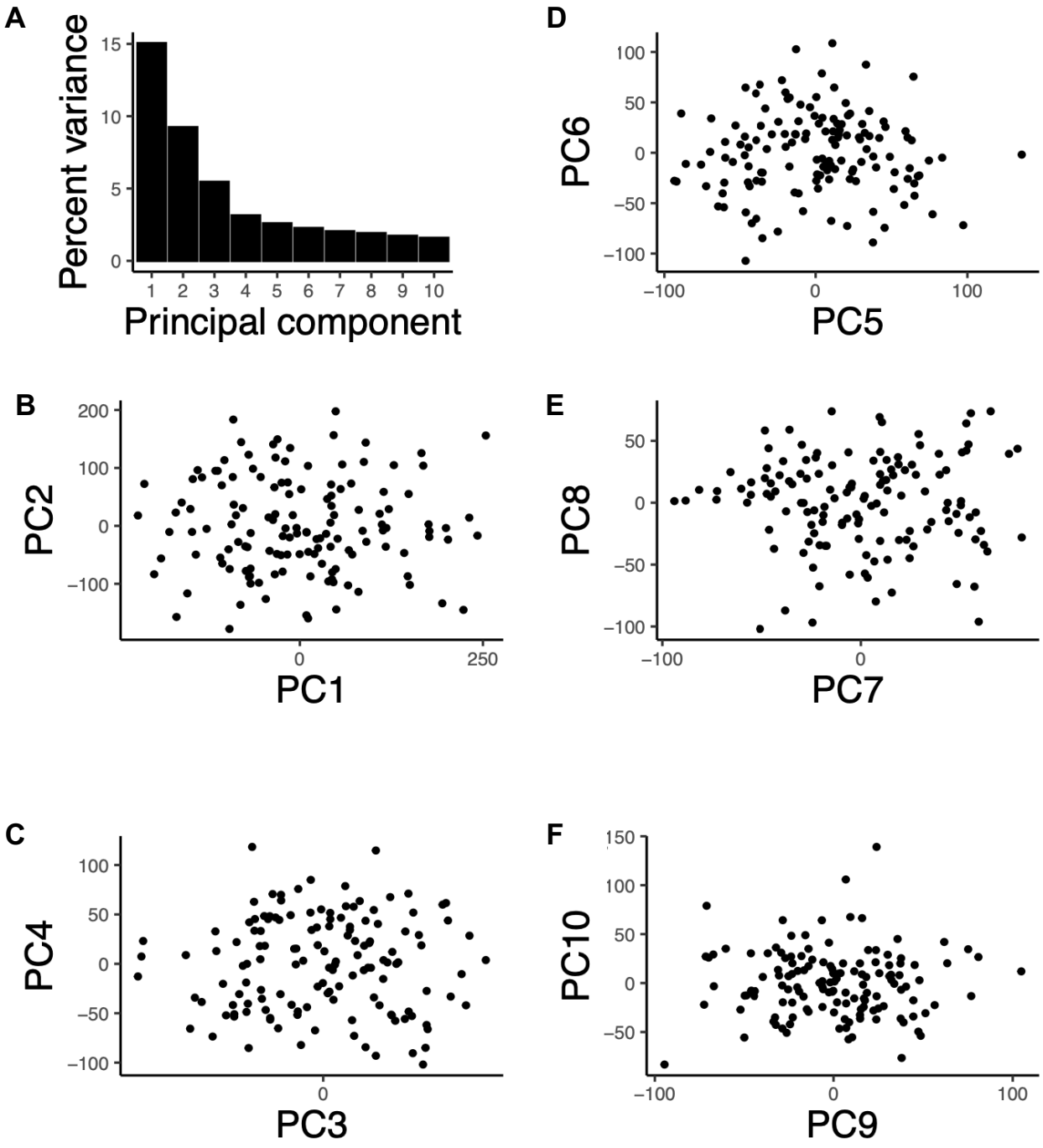
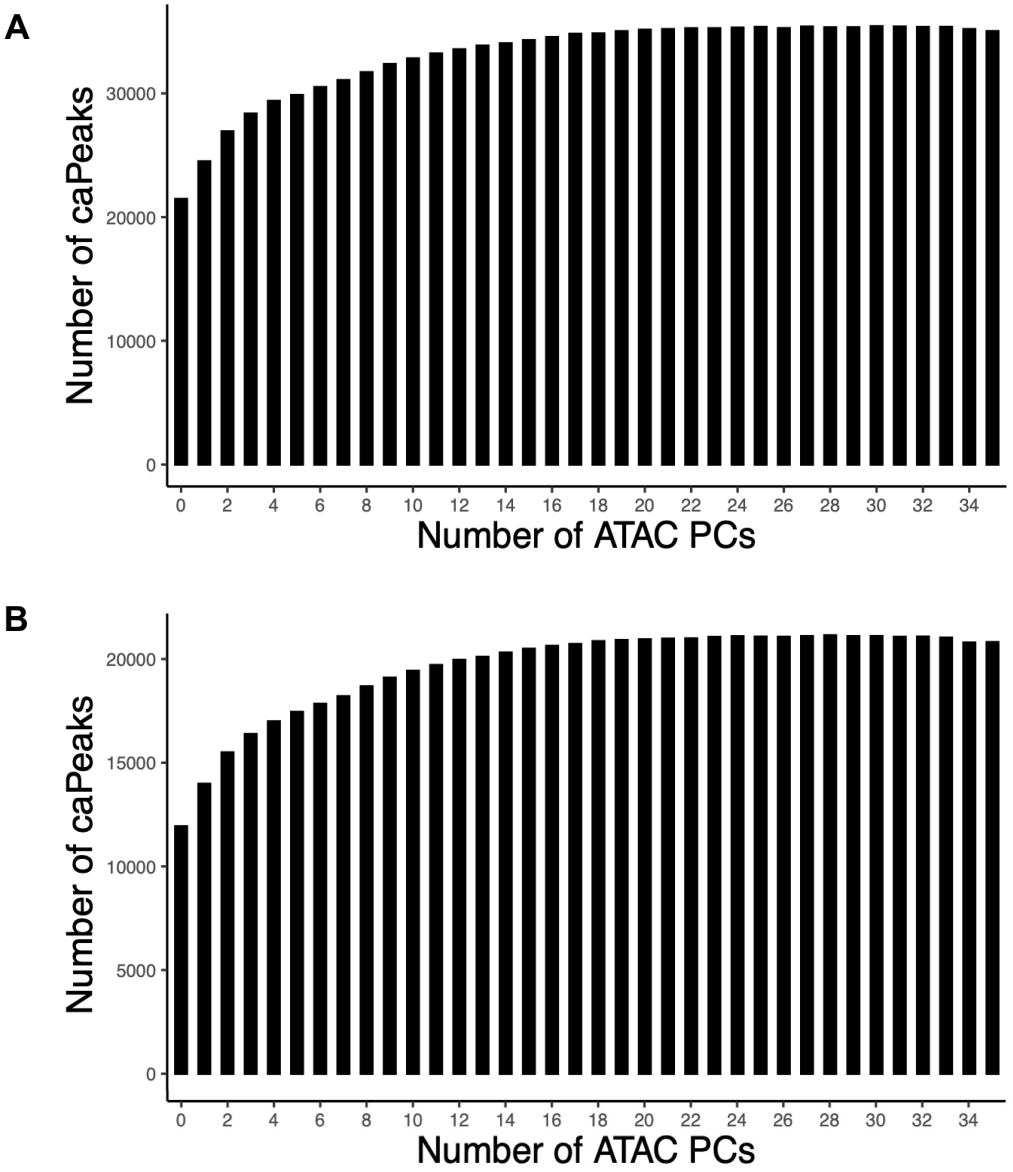


Supplemental Fig. 1: Genotype PCA performed on liver donors combined with 1000G populations (A-C) and using only the 138 liver donors (D-F).

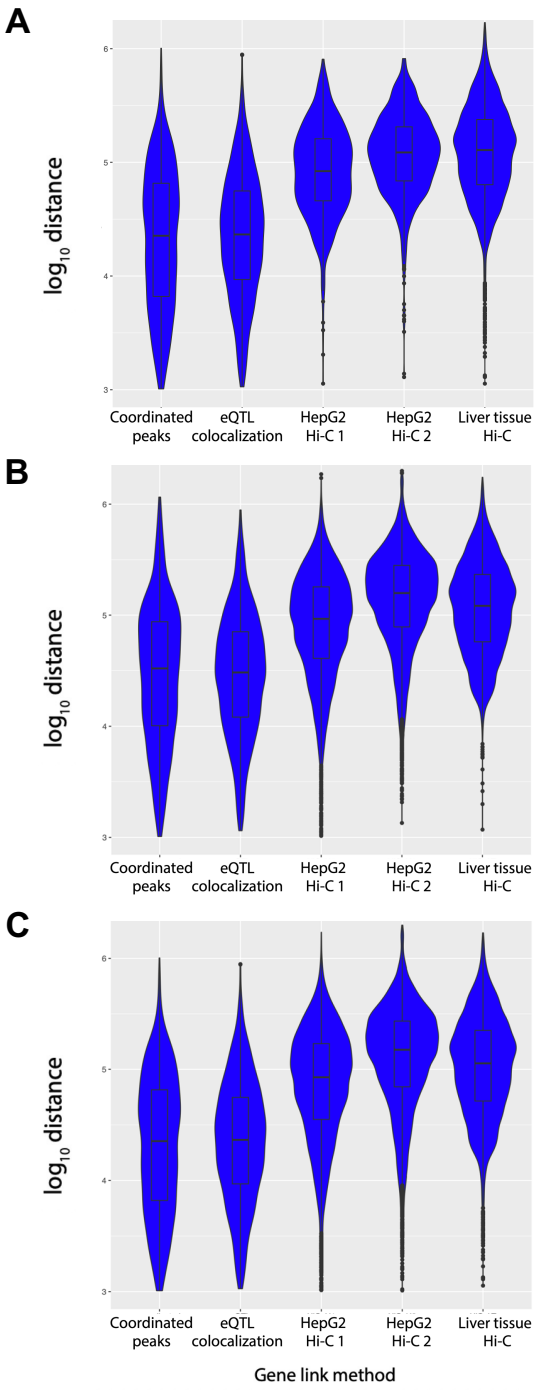
(A) The percent variation explained by the first 10 PCs for liver donors combined with 1000G. (B) PC1 vs. PC2 for liver donors combined with 1000G. (C) PC3 vs. PC4 for liver donors combined with 1000G. (D) The percent variation explained by the first 10 PCs for liver donors only. (E) PC1 vs. PC2 for liver donors only. Population assignments were from the weighted K nearest neighbors assignments using 1000G populations as the reference. (F) PC3 vs. PC4 for liver donors only



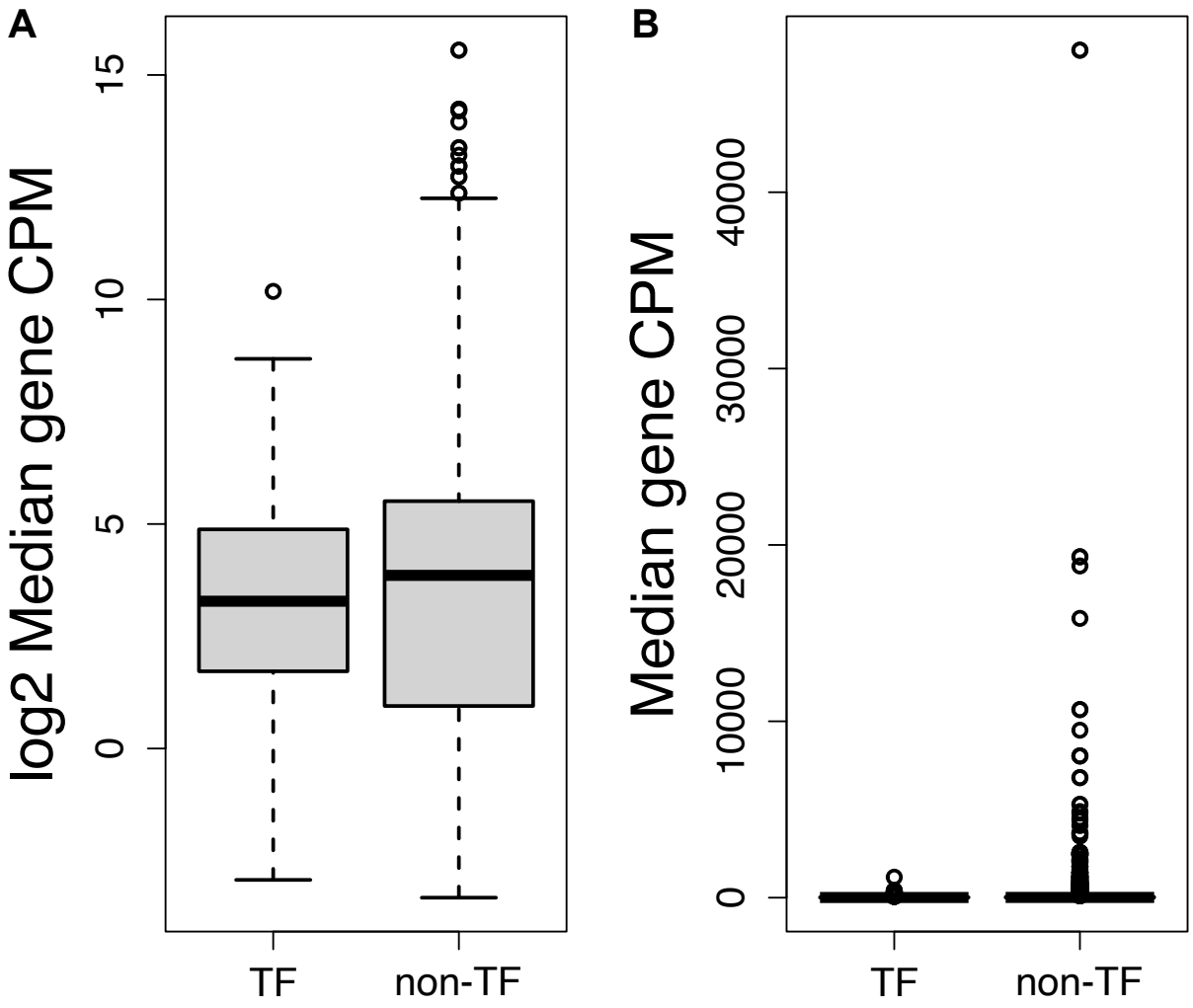
Supplemental Fig. 2: PCA performed on ATAC peak counts normalized for library size and variance-stabilized with DESeq2. Prior to PCA, variance-stabilized peak counts were adjusted for sex and two genotype PCs. (A) The percent variation explained by the first 10 PCs. (B-F) Scatterplots showing pairs of the first 10 PCs.



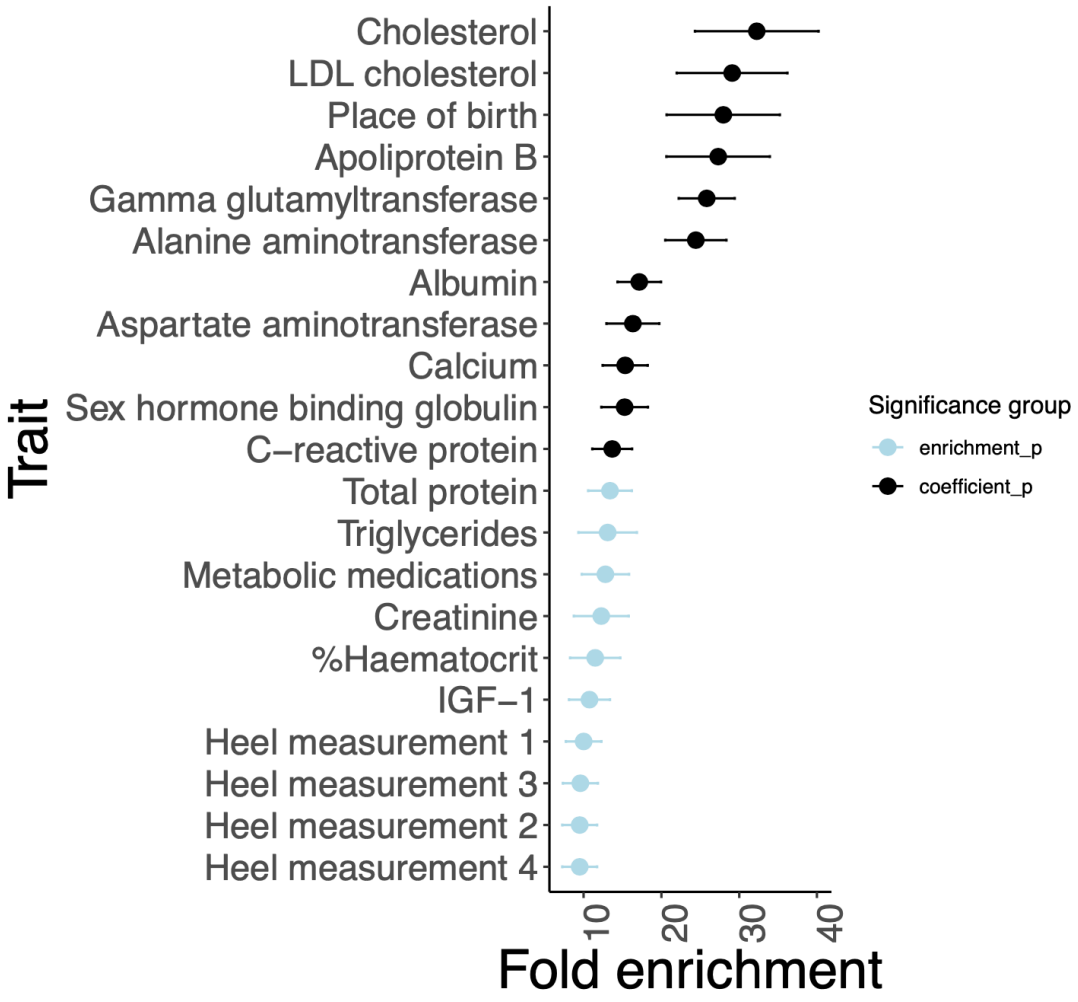
Supplemental Fig. 3. Number of caQTLs identified using various numbers of ATAC principal components (PCs) as covariates.
(A) caQTLs identified using variants within 1 kb of peak centers. (B) caQTLs identified using variants within 1 Mb of peak centers.



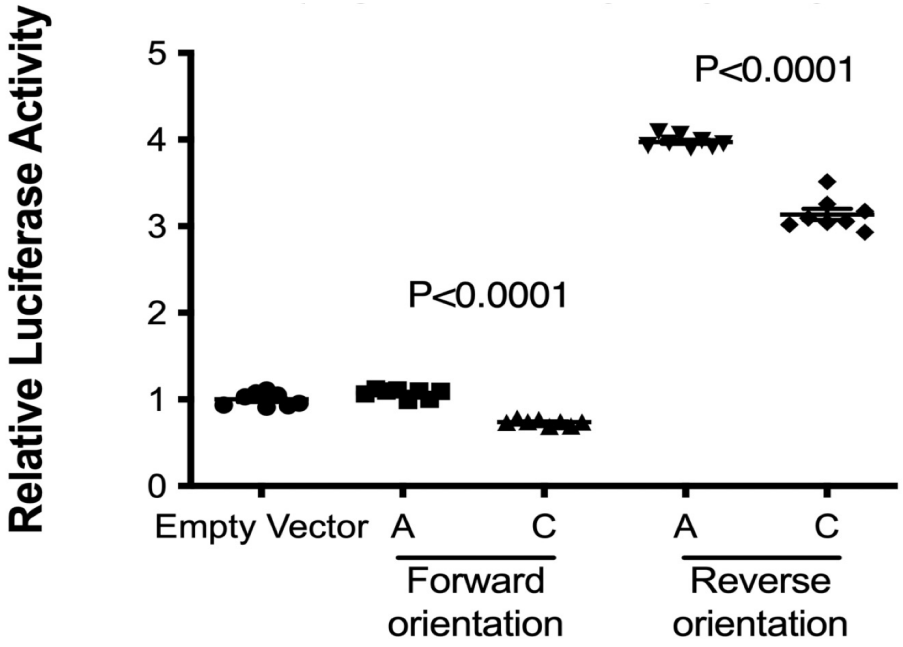
Supplemental Fig. 4. Distributions of log_{10} distances between caPeaks and the transcription start sites (TSS) of linked genes. The plot using distances calculated using the 5' most TSS per gene and Hi-C links between 15 kb and 2 Mb apart is shown in Fig. 3d. (A) Using the closest TSS to the linked caPeak to calculate distances and removing Hi-C links < 15 kb or > 2 Mb apart. (B) Using the 5' most TSS for a gene to calculate distances and including all Hi-C fragments < 2 Mb apart. (C) Using the closest TSS to the linked caPeak to calculate distances and including all Hi-C fragments < 2 Mb apart.



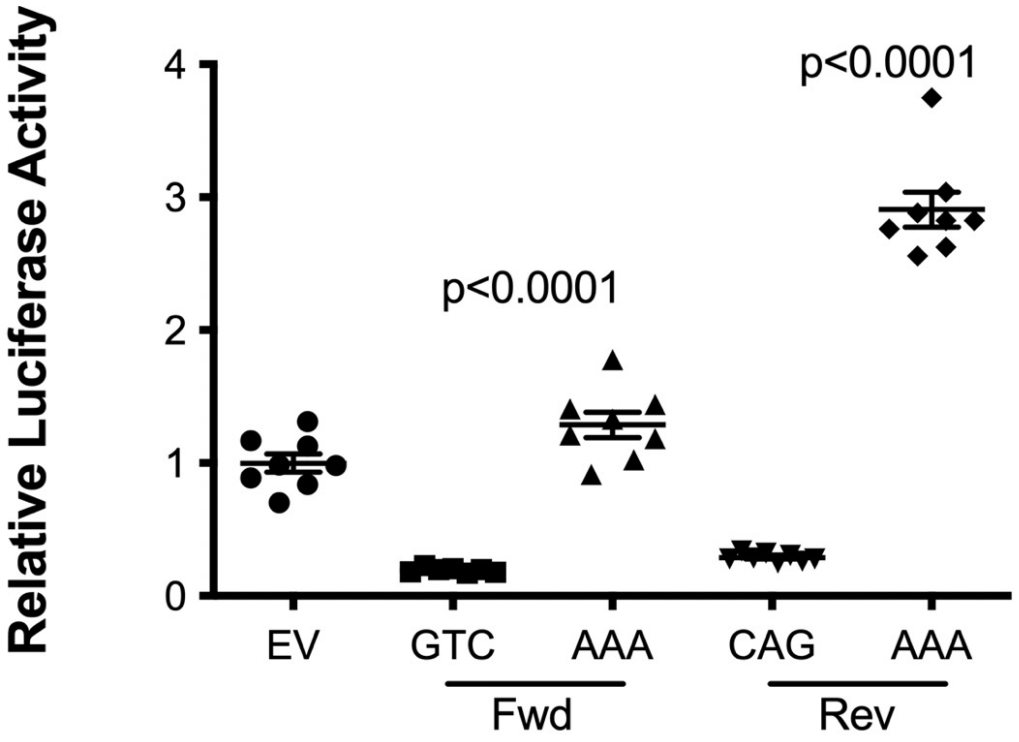
Supplemental Fig. 5. Comparison of gene expression between transcription factor and non-transcription factor genes in GTEx liver tissue. Only genes linked to caPeaks through any of the 4 gene linking methods were included. Gene expression values were normalized with the CPM function in edgeR, using the TMM method for library size adjustment. We then took the median normalized count per gene across GTEx liver donors. (A) Log_2-transformed median normalized gene counts with a pseudo count of 0.1 added. (B) Normalized median gene expression without log transformation.



Supplemental Fig. 6. Heritability enrichment of GWAS traits in caPeaks using LDSC. Traits are shown if they exhibited significant heritability enrichment in caPeaks (FDR<5%) using enrichment_p (red) or coefficient_p (black). All traits that are significant by coefficient_p are also significant by enrichment_p. Dots represent the enrichment fold change estimated by LDSC and error bars are enrichment standard error.



Supplemental Fig. 7. Replicate *TENM2* transcriptional reporter experiments performed on a separate day. Transcriptional activity in HepG2 cells of a 329-bp DNA element spanning caPeak273749 and containing rs7726117. The DNA element was tested in both orientations relative to the genome. EV, empty vector. Symbols represent two independently transfected wells for each of 4 independent clones for each allele; bars indicate mean and standard deviation; p-values from t-tests of allelic differences.



Supplemental Fig. 8. Replicate *RALGPS2* transcriptional activity experiments performed on a separate day. Transcriptional activity in HepG2 cells of a 323-bp DNA element spanning caPeak21014 and containing rs17361251, rs17276513, and rs17276527. The DNA element was tested in both orientations relative to the genome. EV, empty vector. Symbols represent the average of two transfected wells for each of 8 independent clones for each haplotype; bars indicate mean and standard deviation; p-values from t-tests of haplotype differences.