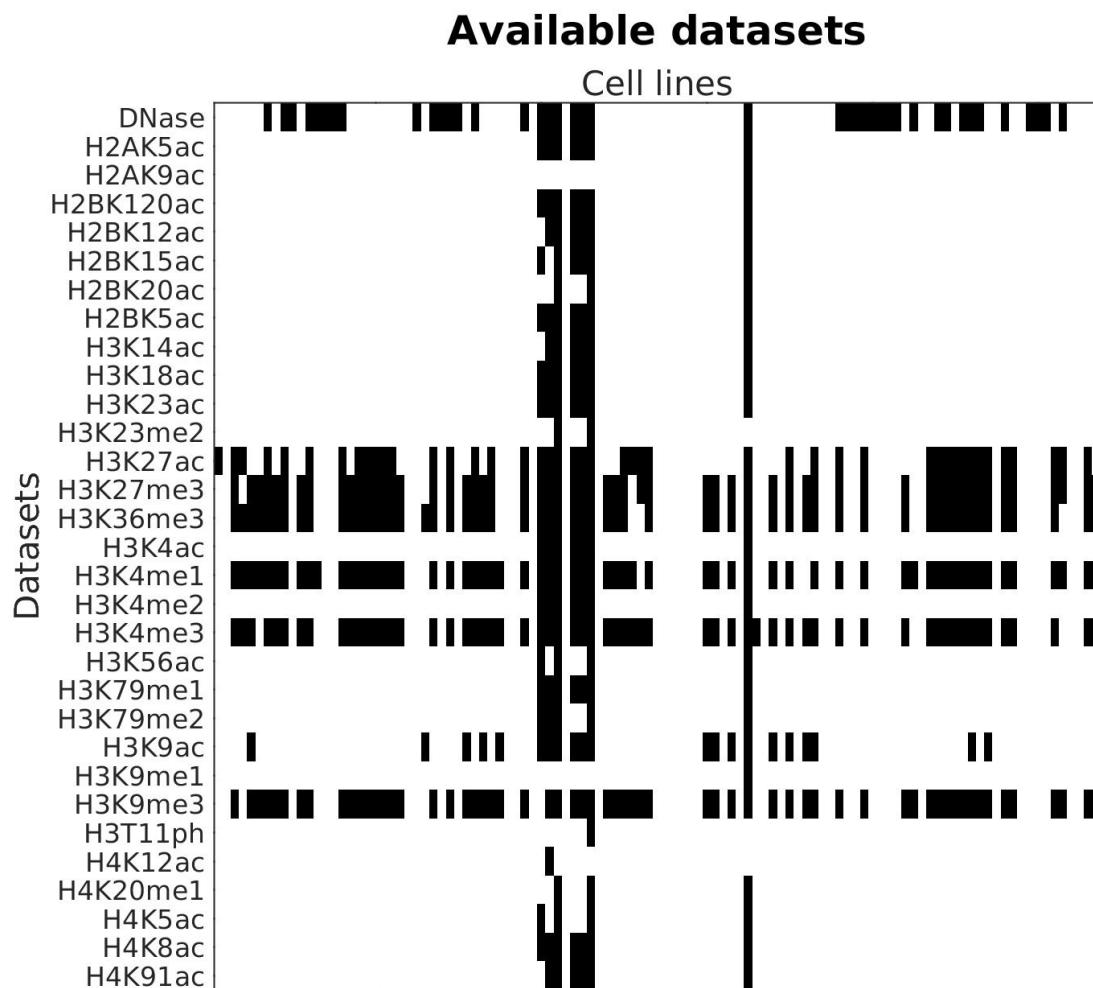
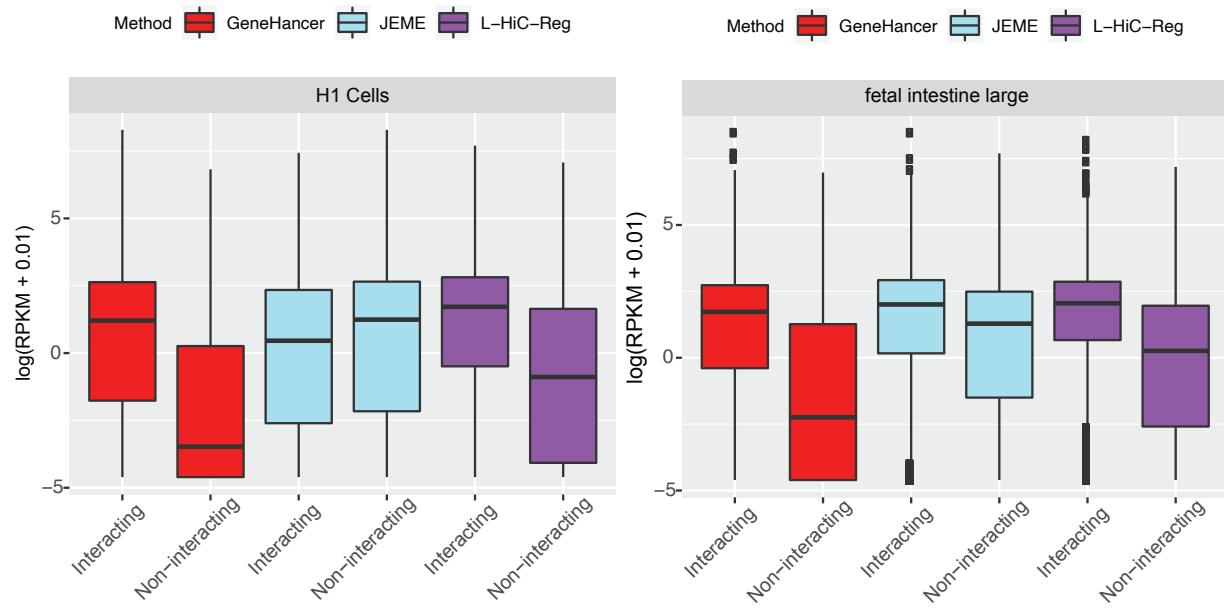


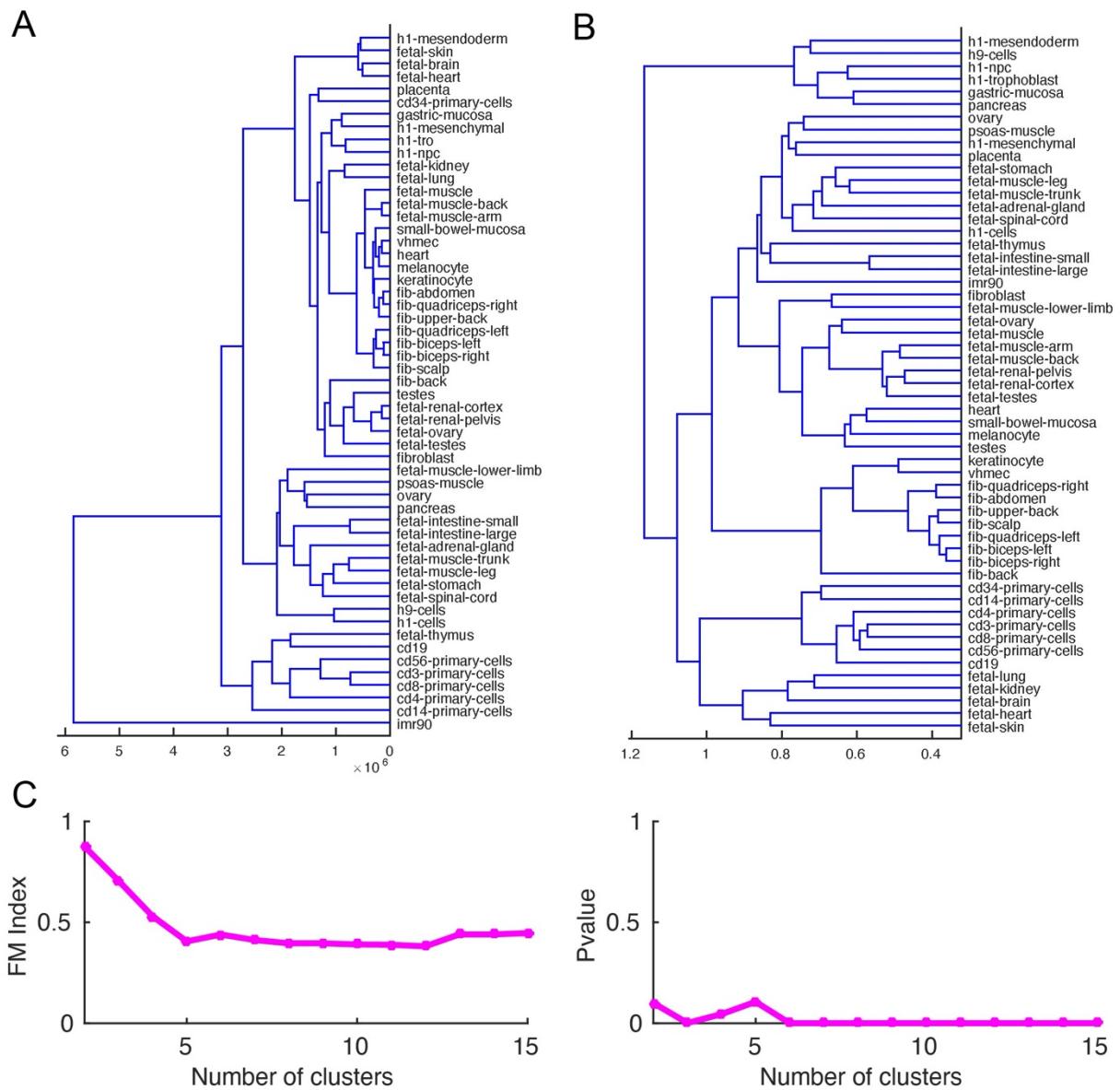
Supplementary figures



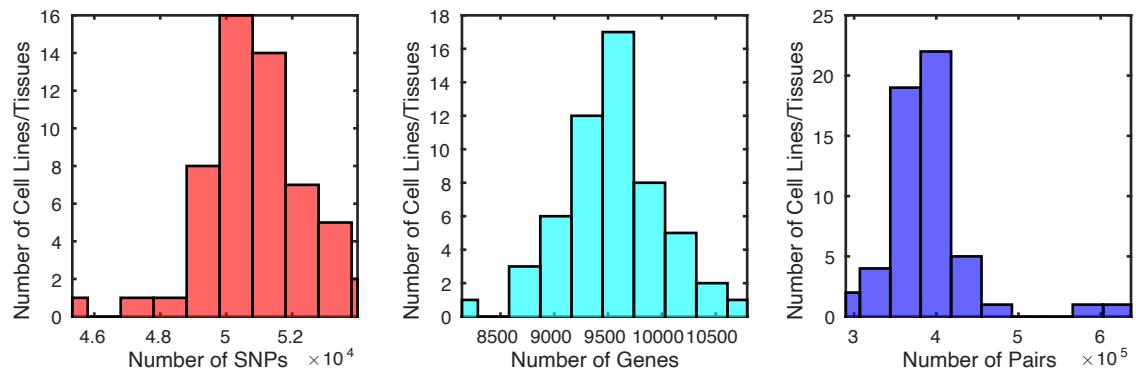
Supplemental Figure 1. Datasets available from the Roadmap epigenomics database across cell lines (columns). Black – dataset exists in Roadmap, White – dataset absent from Roadmap.



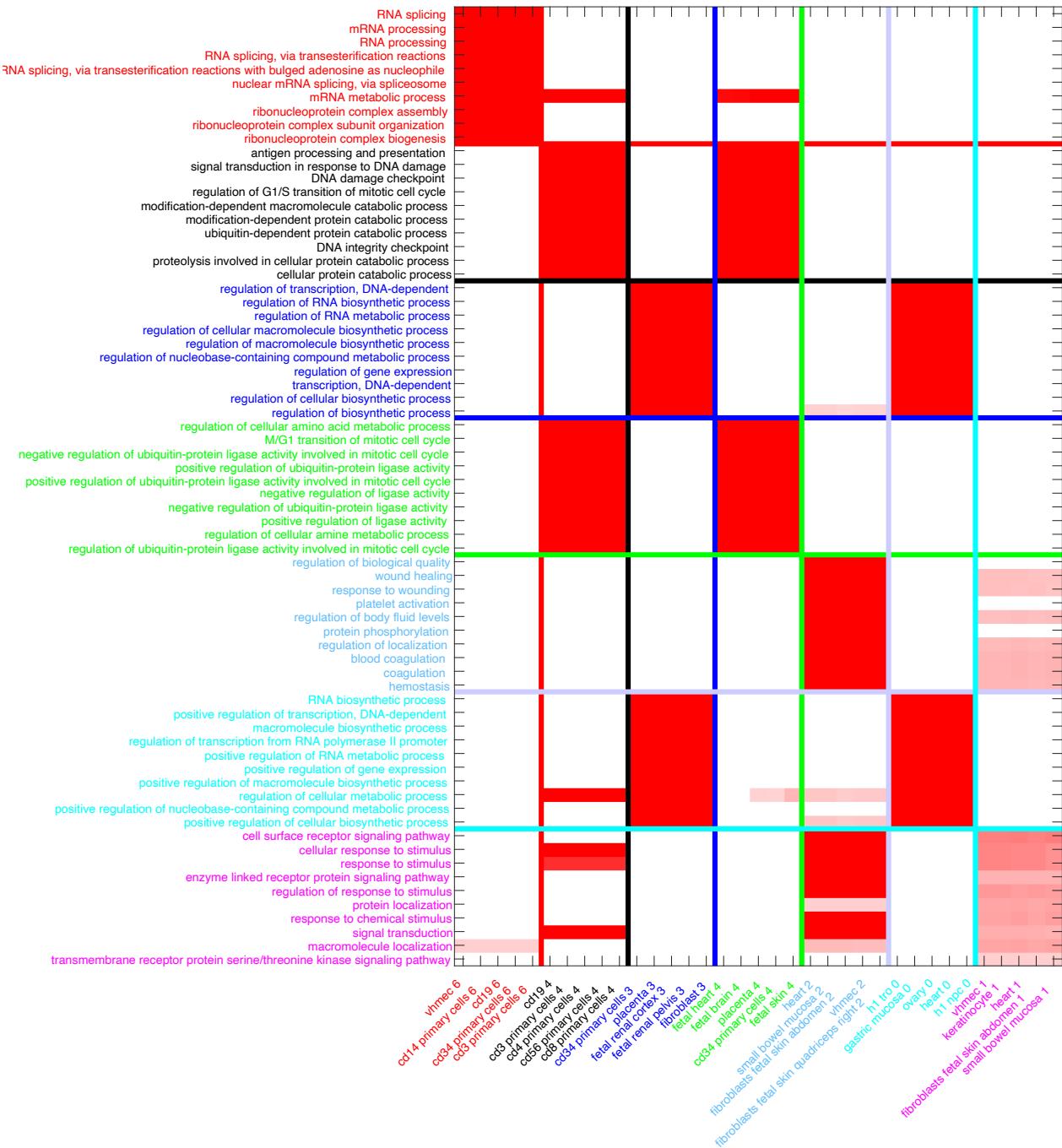
Supplemental Figure 2. Expression of interacting versus non-interacting genes for additional cell lines with expression data and predictions in both JEME and L-HiC-Reg. Related to Main Figure 3.



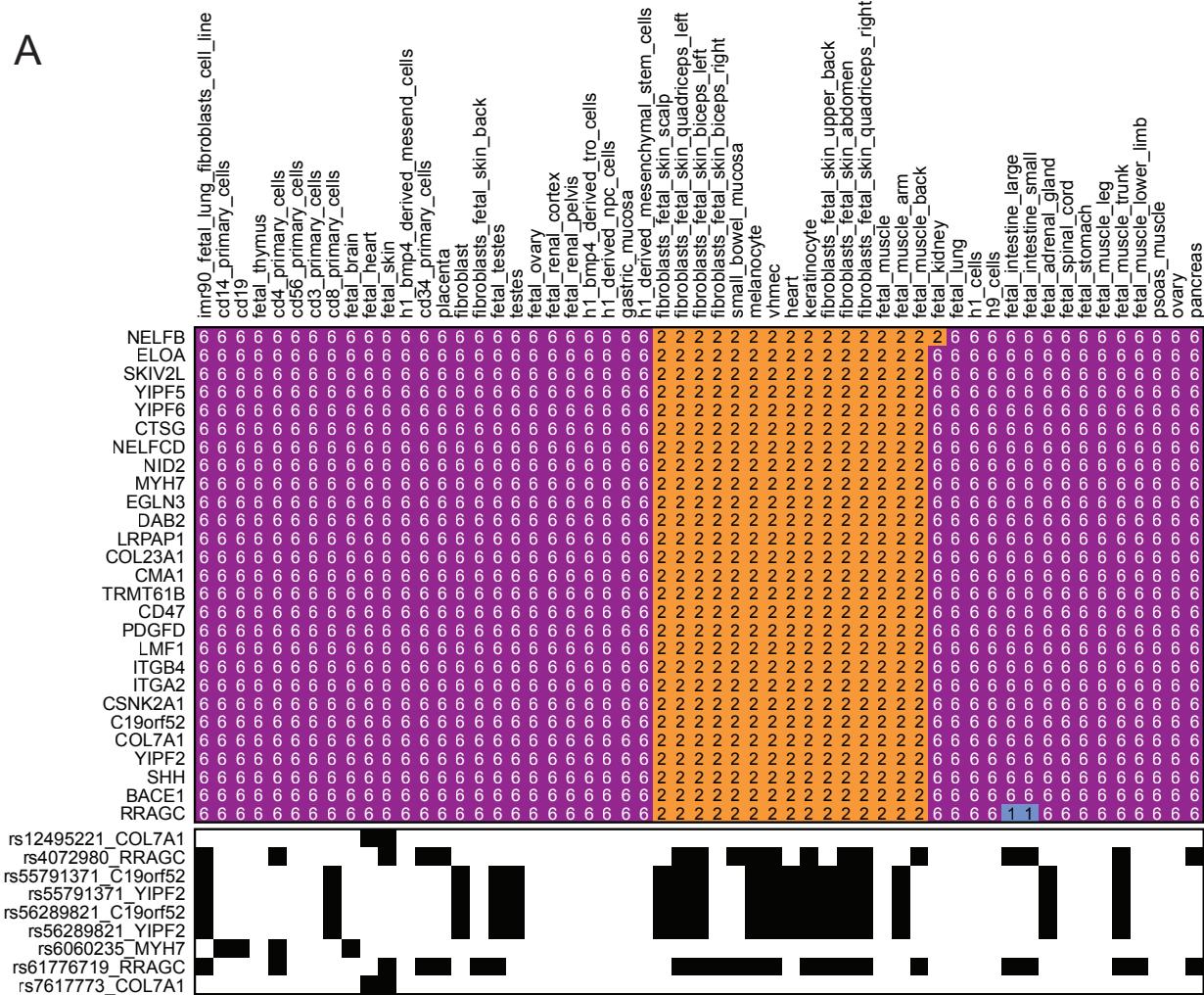
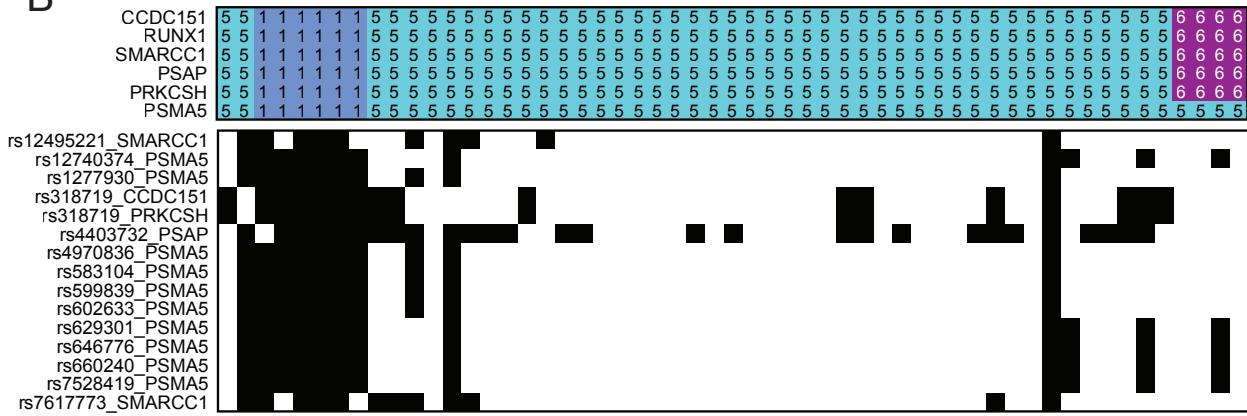
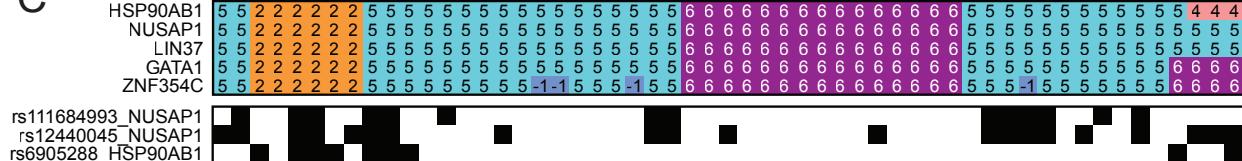
Supplemental Figure 3. **A.** Hierarchical clustering tree showing the similarity of H3K4me3 ChIP-Seq signal in the TSS of genes. This tree was used as input for the Multi-task Graph Clustering algorithm. **B.** Hierarchical clustering based on the F-score of shared interactions between cell lines. **C.** Fowlkes-Mallows index (left) comparing the similarity of two hierarchical clusterings at different levels of clustering with the associated P-value (right) based on random permutation.



Supplemental Figure 4. Distribution of the number of SNPs in LD with the GWAS non-coding SNPs from the NHGRI-EBI GWAS Catalog (left), the number of genes mapped to these SNPs (middle) and the number of pairs connect SNPs to genes across multiple cell lines (left).



Supplemental Figure 5. Top Gene Ontology terms and cell lines across the 7 multi-task graph clustering clusters. Numbers next to the cell line name represents the graph cluster. To enable extraction of meaningful annotation from our GO analysis, we applied Non-negative Matrix Factorization with seven low-dimensional factors on the -log(q-value) scores of all the GO terms across all cell lines and gene clusters. For each NMF factor, the top 10 GO terms and the top 5 cell lines were selected based on their values in the corresponding lower-dimensional factors.

A**B****C**

Supplemental Figure 6. Gene sets with changes in graph cluster assignment across cell lines that exhibit concordant presence-absence of SNP-gene interactions for Coronary Artery Disease (CAD). Shown are three gene sets, #37 (**A.**), #46 (**B.**) and #53 (**C.**). Color heatmaps represent the cluster assignments for all 55 cell lines and the black and white heatmaps represents the presence (black) or absence (white) of the SNP-gene interaction with specific genes in the gene set across the cell lines.