

Supplementary figures

Figure 1 Graphical model for PRRR and nn-PRRR.

Figure 2 nn-PRRR coefficients for pancreatic cell types. Heatmaps showing the full coefficient matrix UV^T for nn-PRRR (left is original, and right is on a log scale). Cell types are shown on the rows and genes on the columns. In the left panel, white cells indicate values near zero, implying that this coefficient matrix is highly sparse.

Figure 3 Marker genes identified by PRRR for pancreatic cell types. For each cell type, the ten genes with the highest coefficients in the matrix UV^T were extracted for each cell type. Some cell types share the same ten marker genes, which corresponds with our observation that the cell types are largely overlapping in a PCA plot of the gene expression data (Supplementary Figure 4).

Figure 4 PCA plot of pancreas scRNA-seq data. The first two principal components (PCs) are plotted. Each point corresponds to a single cell and is colored by its annotated cell type.

Figure 5 nn-PRRR coefficients for GTEx eQTL mapping. Left: U matrix showing SNPs on the rows and latent factors on the columns. Right: V matrix showing genes on the rows and latent factors on the columns.