Supplementary figures

Supplementary Figure 1. Percentages for purity yield of sorted cells from two porcine males using BD FACSymphony™.

A)

B)

C)

D)

E)

- Eukaryotic translation elongation
- T cell activation
- Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell
- T cell receptor signaling pathway
- Adaptive immune response
- Class A/1 (Rhodopsin-like receptors)
- Chemotaxis
- Pathways in cancer
- PID CD8 TCR PATHWAY
- Second-messenger-mediated signaling
- Cell killing
- Leukocyte migration
- Wound healing
- NABA MATRISOME ASSOCIATED
- Positive regulation of cell death
- Positive regulation of cytosolic calcium ion concentration involved in phospholipase C-activating G protein-coupled signaling pathway
- Cell adhesion molecule binding
- Hematopoietic cell lineage
- Peptidyl-tyrosine phosphorylation
- Positive regulation of kinase activity

F)

- Adaptive immune response
- T cell receptor signaling pathway
- Pathways in cancer
- Th17 cell differentiation
- Chemotaxis
- T cell selection
- Hemostasis
- Cytokine-cytokine receptor interaction
- Positive regulation of cytokine production
- Regulation of T cell differentiation
- Calcium ion binding
- Blood vessel development
- Transmembrane receptor protein tyrosine kinase signaling pathway
- Second-messenger-mediated signaling
- Neuroactive ligand-receptor interaction
- Leukocytes migration
- Cell adhesion molecules
- Rap1 signaling pathway
- Peptidyl-tyrosine phosphorylation
- Ras guanyl-nucleotide exchange factor activity

G)

- Regulation of cell adhesion
- Cytokine-cytokine receptor interaction
- Hematopoietic cell lineage
- Th17 cell differentiation
- Chemotaxis
- Leukocytes migration
- Blood vessel development
- Interferon-gamma production
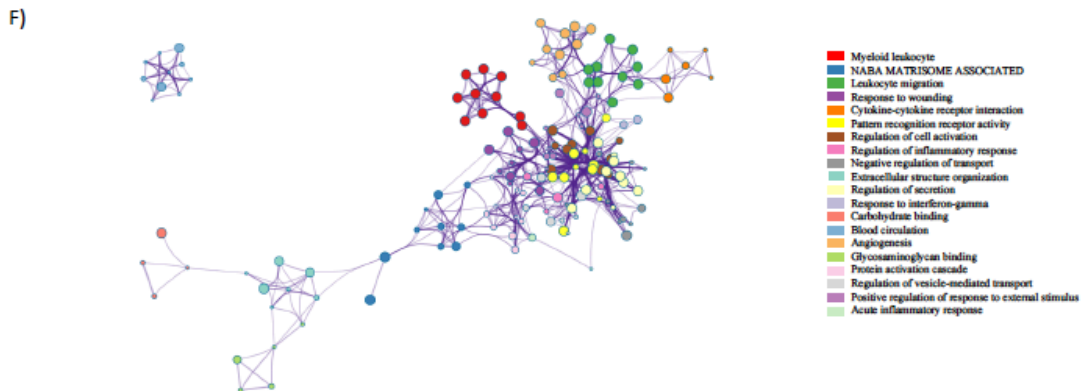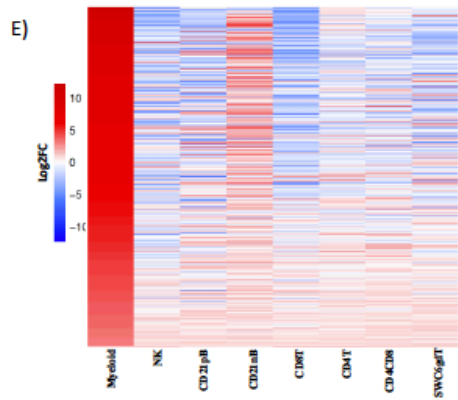- Alpha-beta T cell activation
- Cell adhesion molecules (CAMs)
- T cell selection
- Cell-substrate adhesion
- Wound healing
- Inflammatory bowel disease
- Positive regulation of cytosolic calcium ion concentration involved in phospholipase C-activating G protein-coupled signaling pathway
- Peptidyl-tyrosine phosphorylation
- Positive regulation of MAPK cascade
- Integrin binding
- Transmembrane receptor protein tyrosine kinase signaling pathway
- Positive regulation of response to external stimulus

H)

- Regulation of cell adhesion
- Cell morphogenesis involved in differentiation
- Second-messenger-mediated signaling
- Chemotaxis
- Regulation of system process
- Pathways in cancer
- Angiogenesis
- Cell-substrate adhesion
- Transmembrane receptor protein tyrosine kinase signaling pathway
- Hemostasis
- Peptidyl-tyrosine phosphorylation
- GPCR ligand binding
- Cytokine-cytokine receptor interaction
- Actin cytoskeleton organization
- Th17 cell differentiation
- Regulation of actin cytoskeleton
- Guanyl-nucleotide exchange factor activity
- G protein-coupled receptor signaling pathway
- Modulation of chemical synaptic transmission
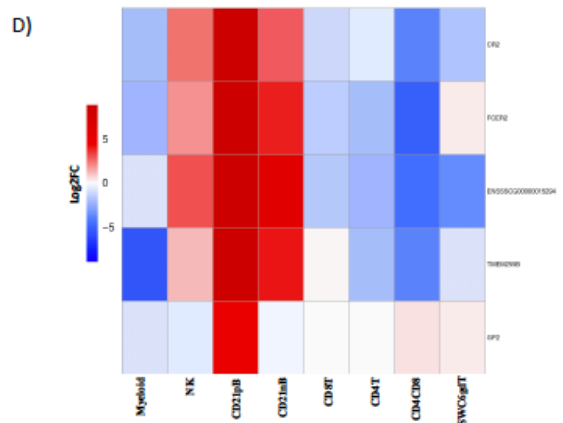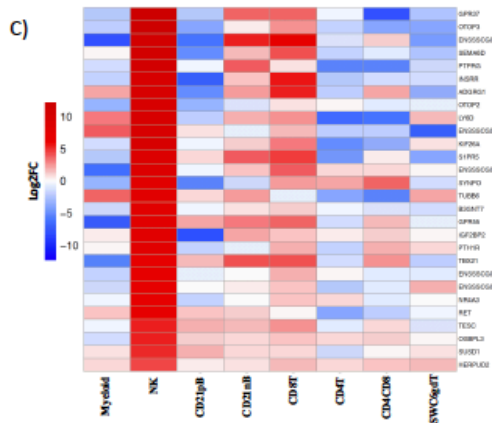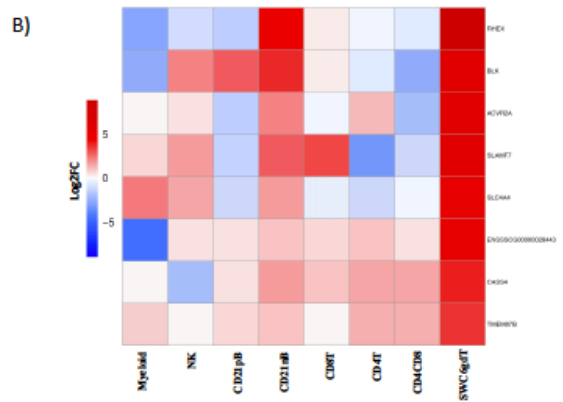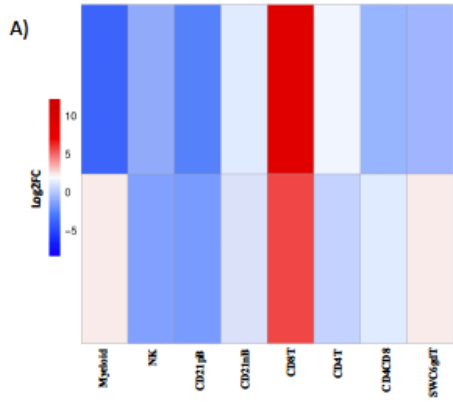- Protein kinase activity

Supplementary Figure 2. Top 25% highly enriched genes in CD3$^+$ sorted cells. Heatmap showing in decreasing order the top 25% of highly enriched genes in **A**) CD8T, **B**) CD4T, **C**) CD4CD8T and **D**) SWC6gdT cells. Ontology enrichment clusters of the top 25% highly enriched genes of **E**) CD8T, **F**) CD4T, **G**) CD4CD8T and **H**) SWC6gdT cells. The most statistically significant term within similar term clusters was chosen to represent the cluster. Term color is given by cluster ID and the size of the terms is given by –log10 P-value. The stronger the similarity among terms, the thicker the edges between them.
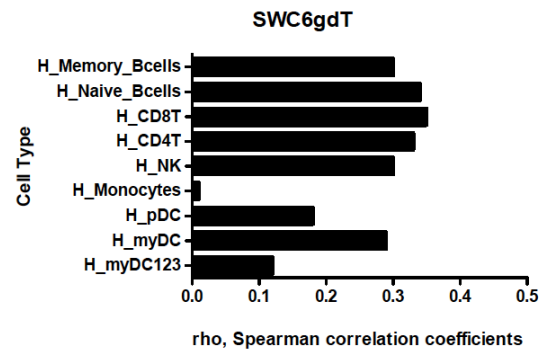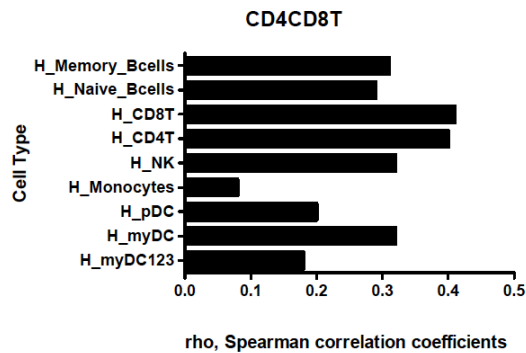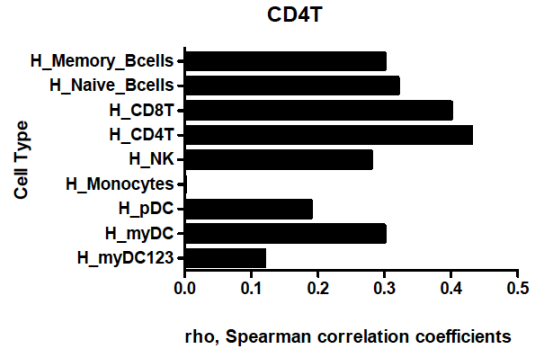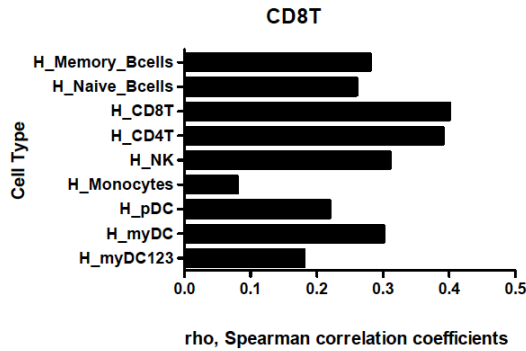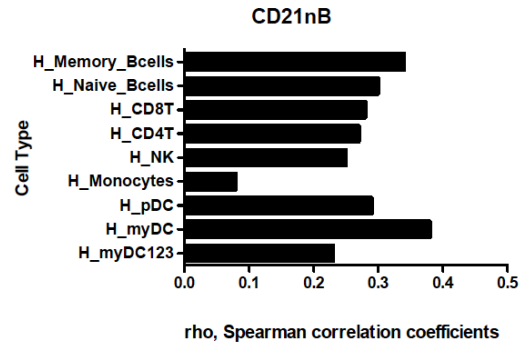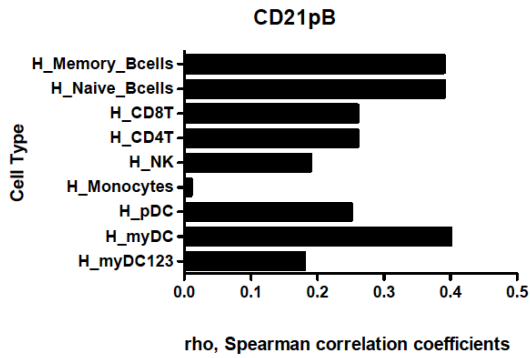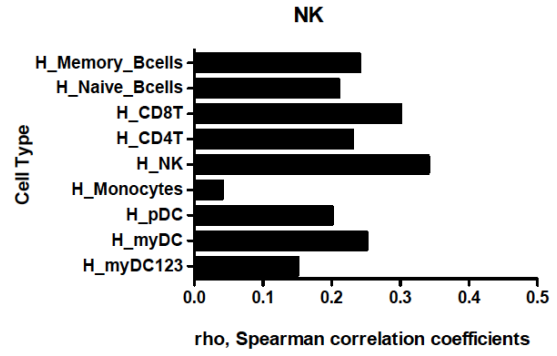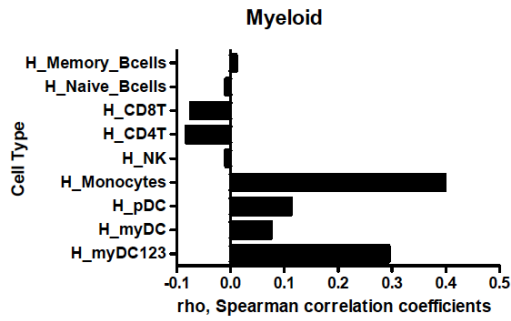
Supplementary Figure 3. Spearman rank correlation between TPM values of expressed genes in porcine sorted immune cells and their orthologous genes in human sorted immune cells.
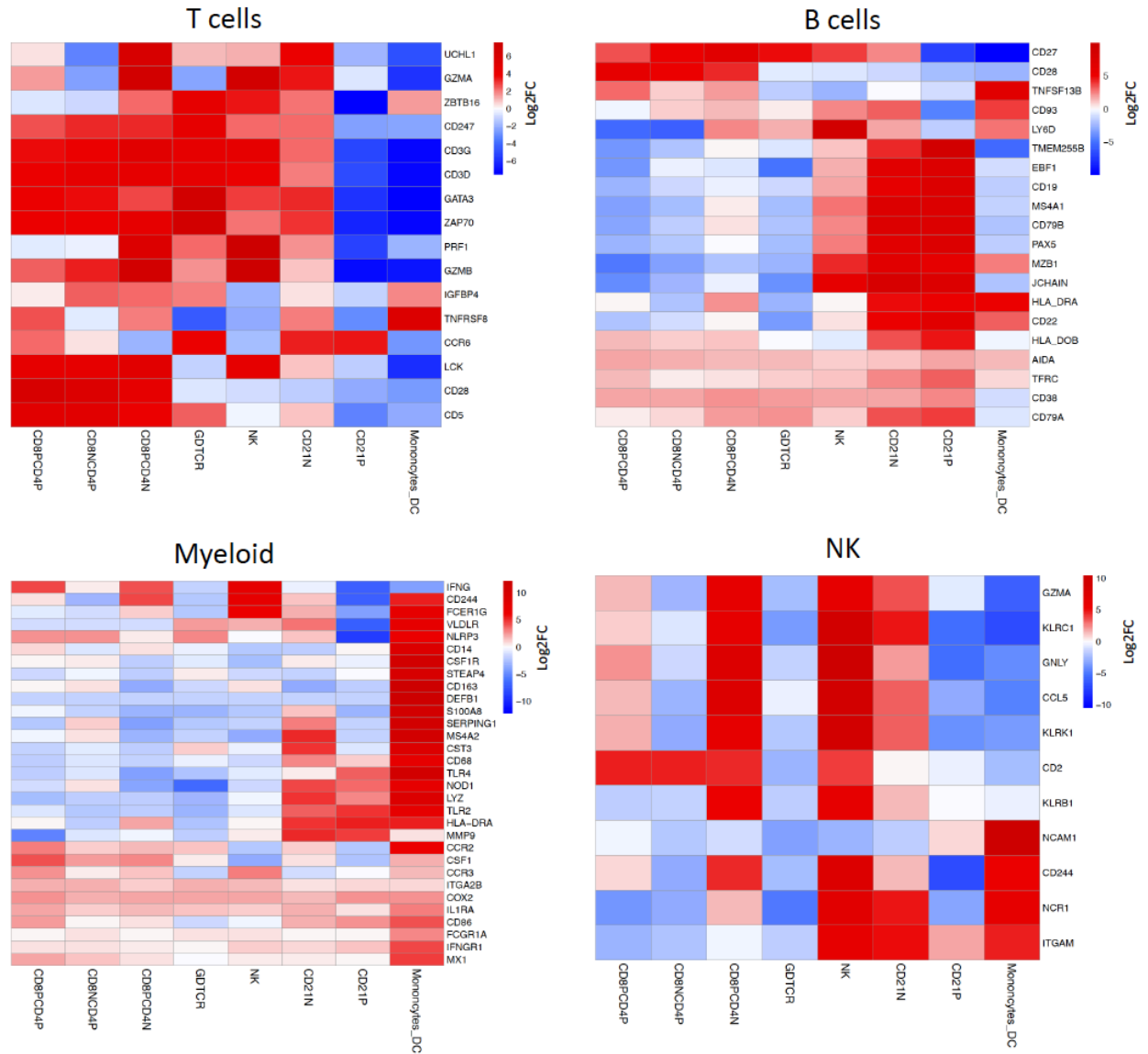
Supplementary Figure 4. Gene enrichment heatmap of highly expressed genes and cell markers reported in human and mouse databases in T cells, B cells, myeloid and NK cells.

A)

B)

C)

D)

E)

F)

Myeloid leukocyte
NABA MATRISOME ASSOCIATED
Leukocyte migration
Response to wounding
Cytokine-cytokine receptor interaction
Pattern recognition receptor activity
Regulation of cell activation
Regulation of inflammatory response
Negative regulation of transport
Extracellular structure organization
Regulation of secretion
Response to interferon-gamma
Carbohydrate binding
Blood circulation
Angiogenesis
Glycosaminoglycan binding
Protein activation cascade
Regulation of vesicle-mediated transport
Positive regulation of response to external stimulus
Acute inflammatory response

Supplementary Figure 5. Cell-type specific genes in CD8T, SWC6gdT, NK, CD21pB and myeloid cells. Heatmap showing in decreasing order the cell-type specific genes in **A)** CD8T, **B)** SWC6gdT, **C)** NK, **D)** CD21pB and **E)** myeloid. Ontology enrichment clusters cell-type specific genes of **F)** myeloid cells. The most statistically significant term within similar term cluster was chosen to represent the cluster. The most statistically significant term within similar term clusters was chosen to represent the cluster. Term color is given by cluster ID and the size of the terms is given by –log10 P-value. The stronger the similarity among terms, the thicker the edges between them.
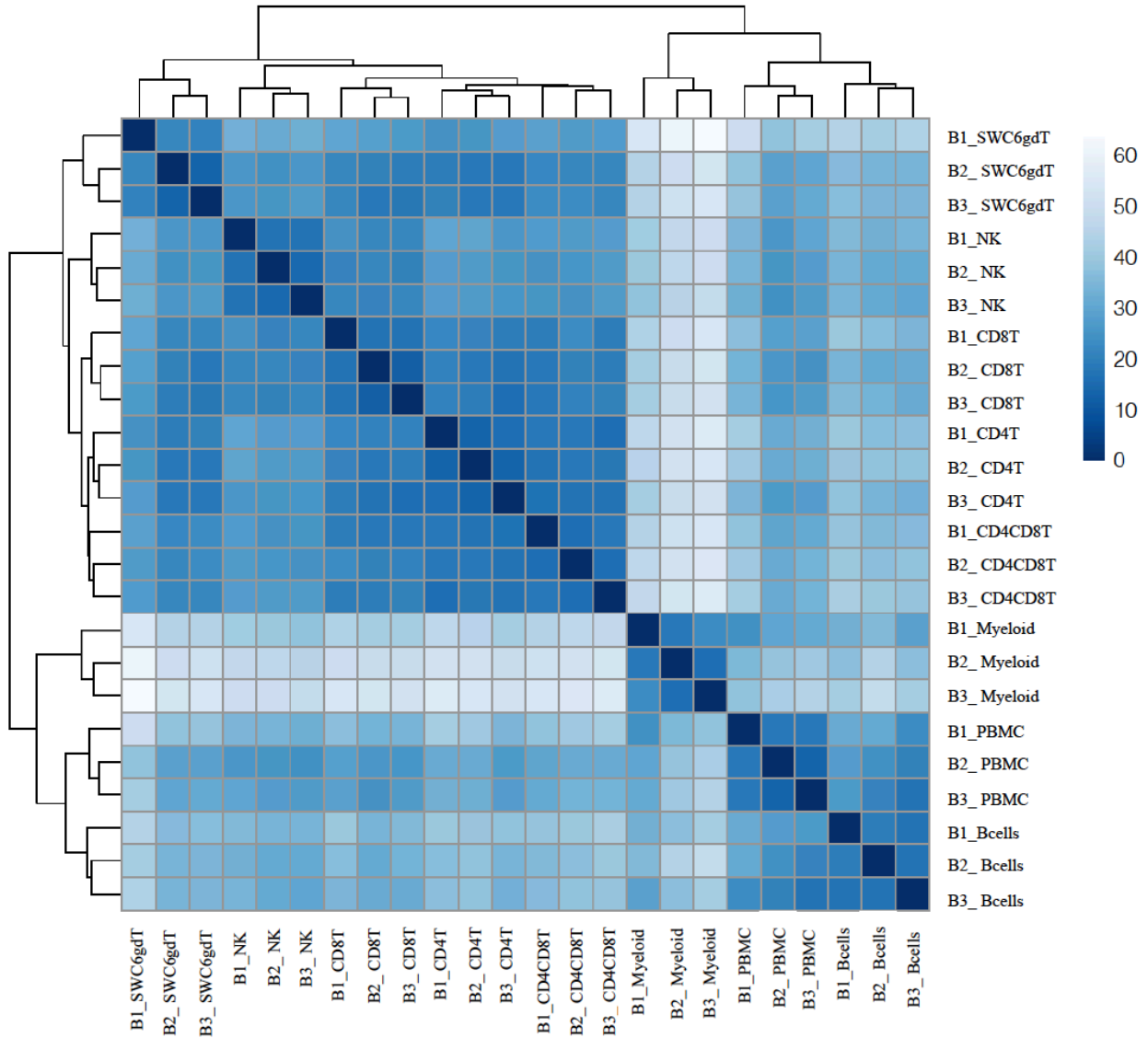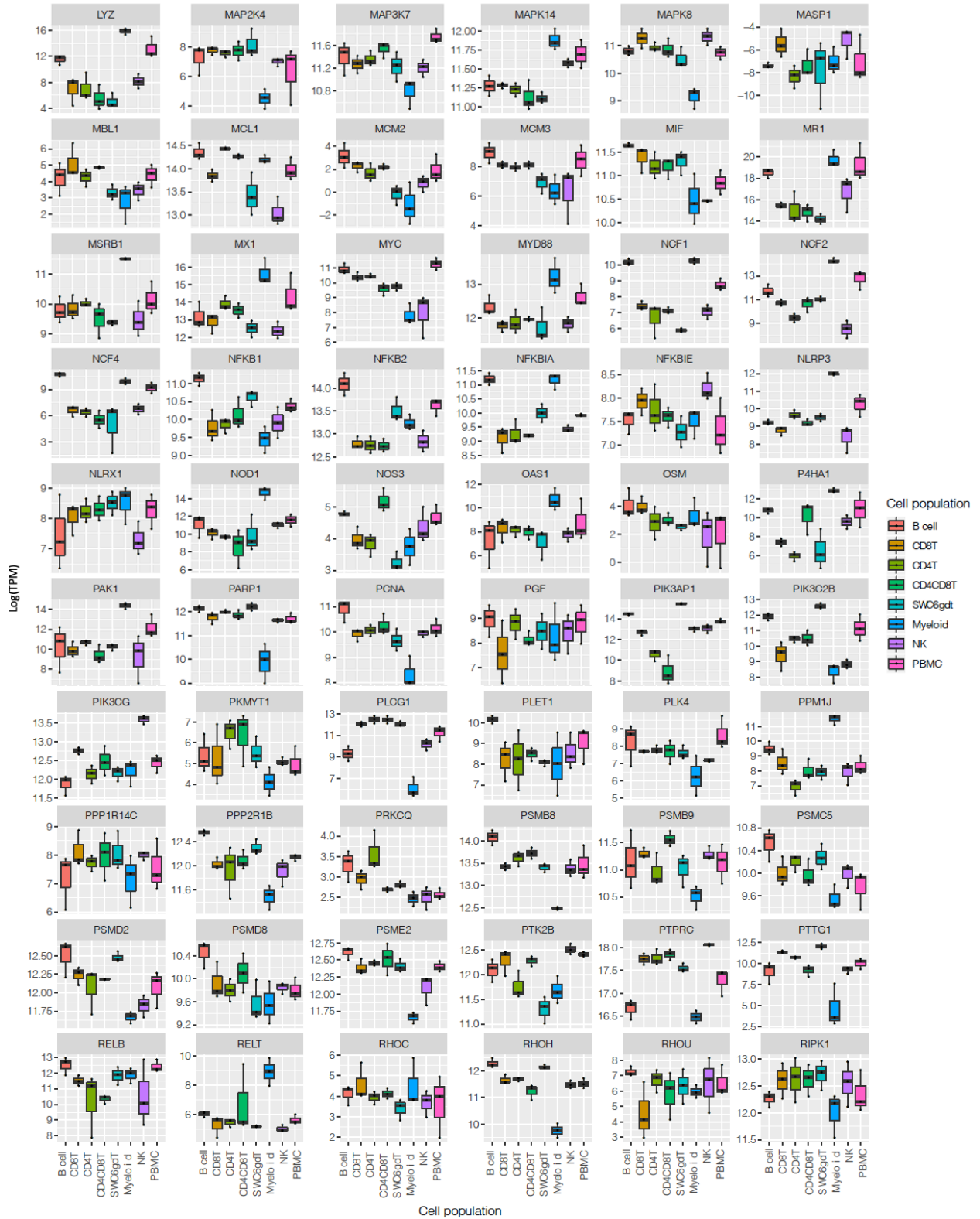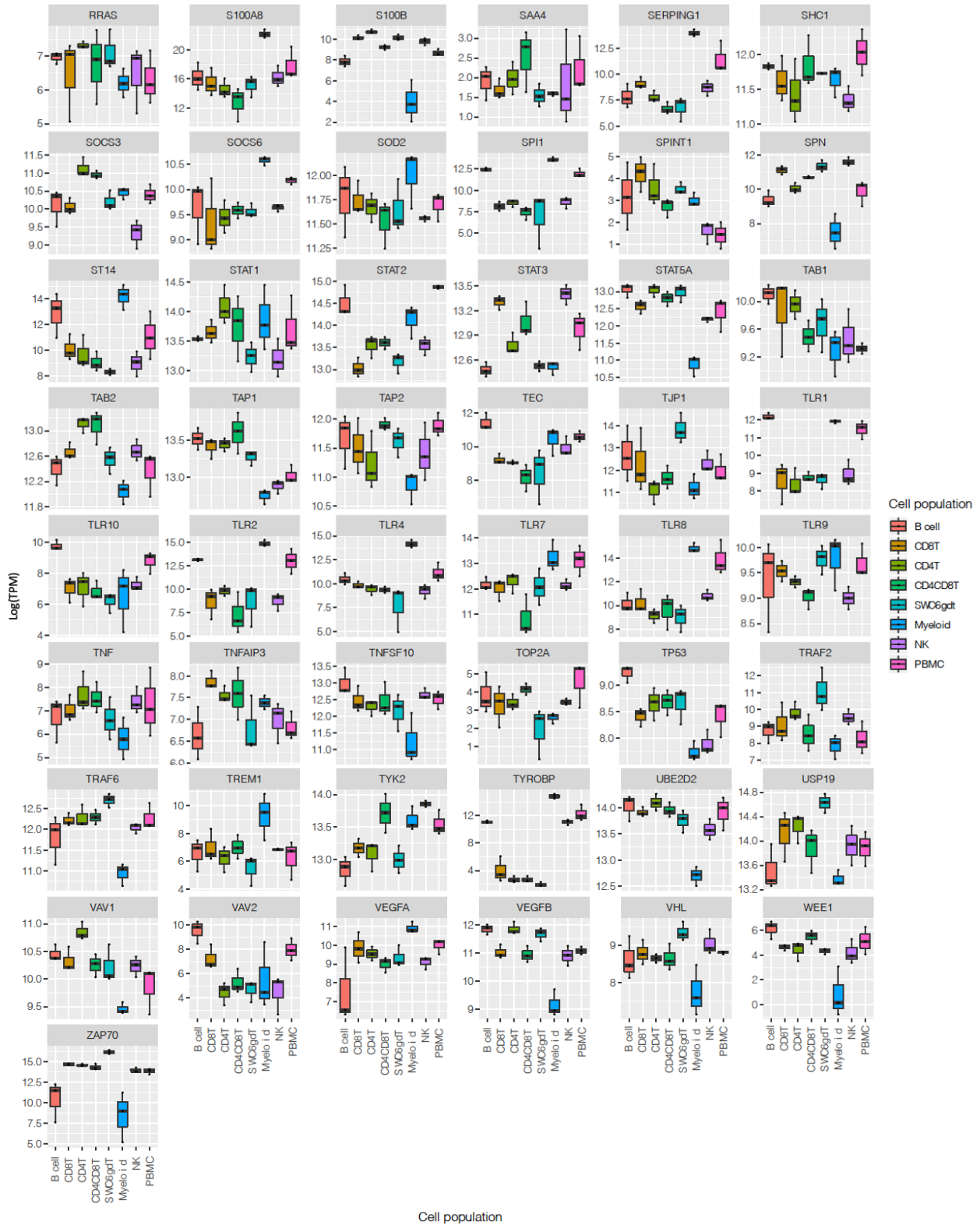
Supplementary Figure 6. Gene expression patterns of sorted immune cells showed closer clustering than those share far common progenitors during hematopoiesis. Principal component analysis of transformed reads counts for selected genes in Nanostring. Axis indicate component scores.
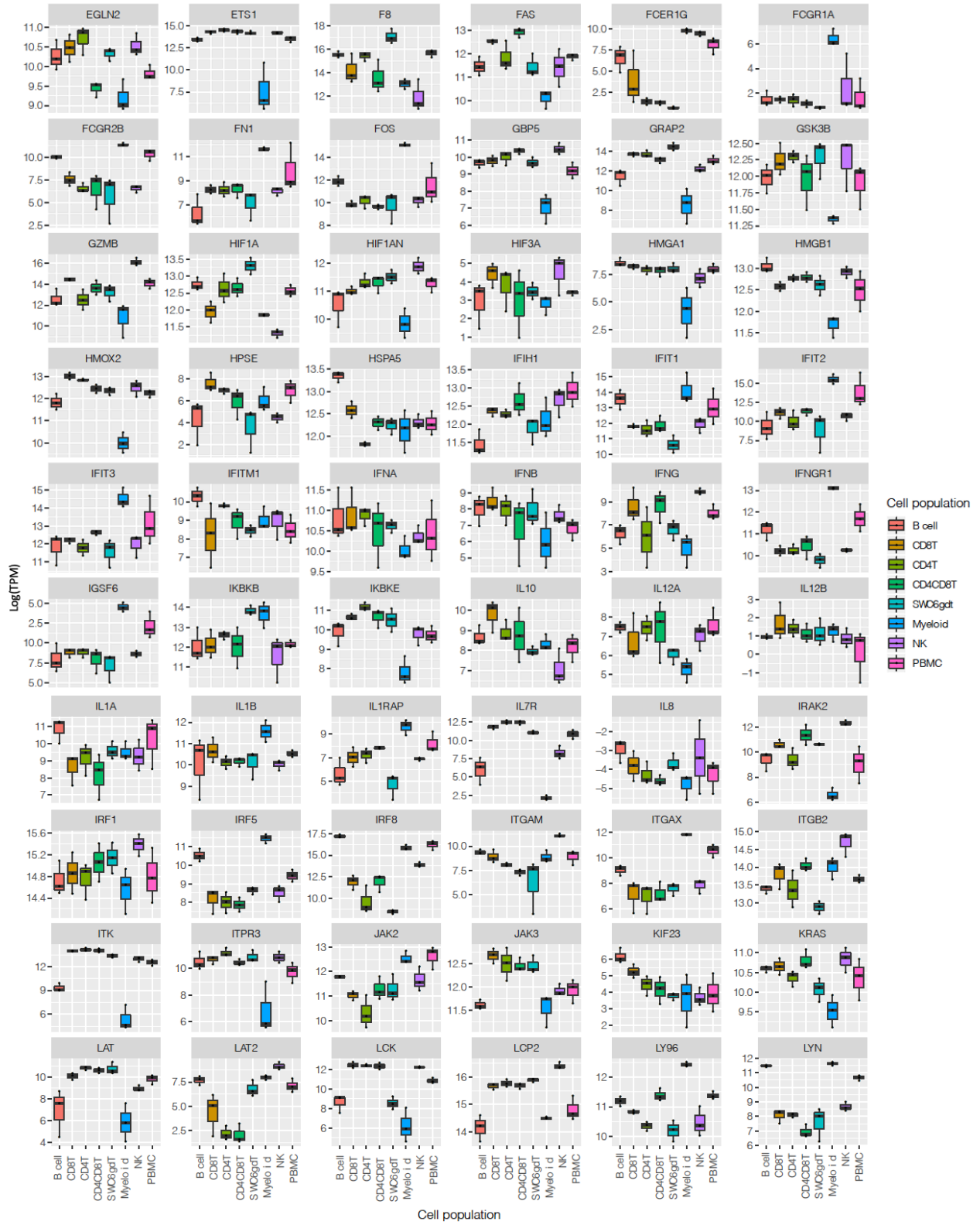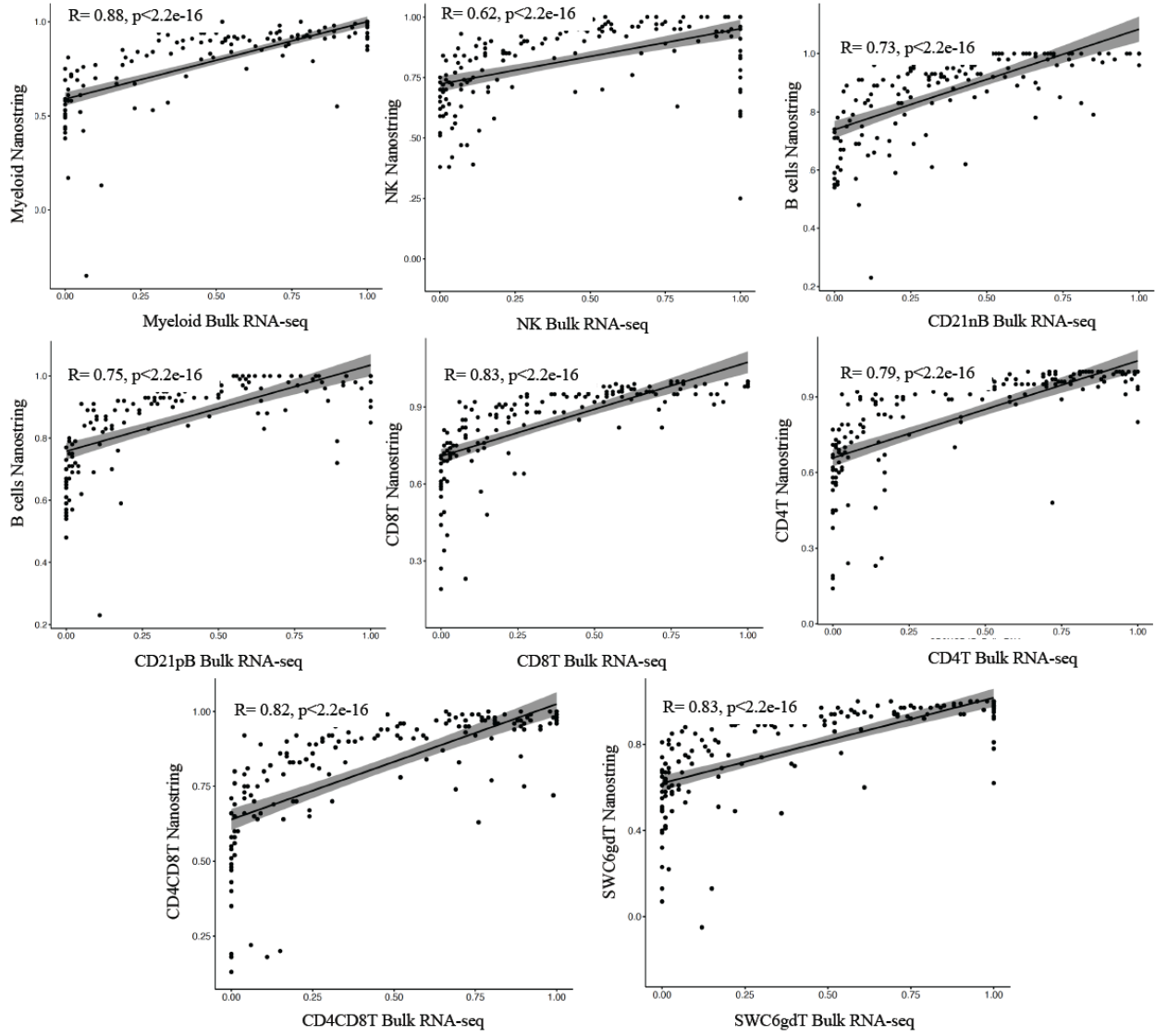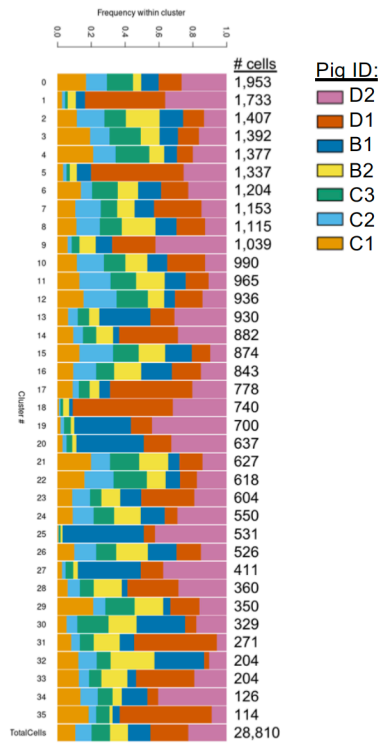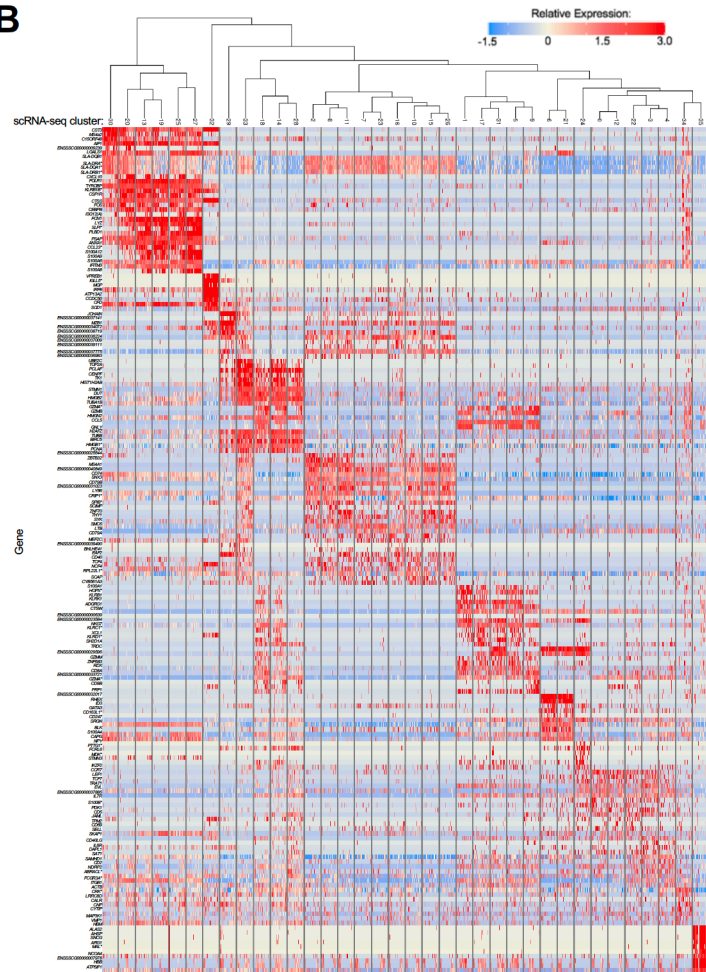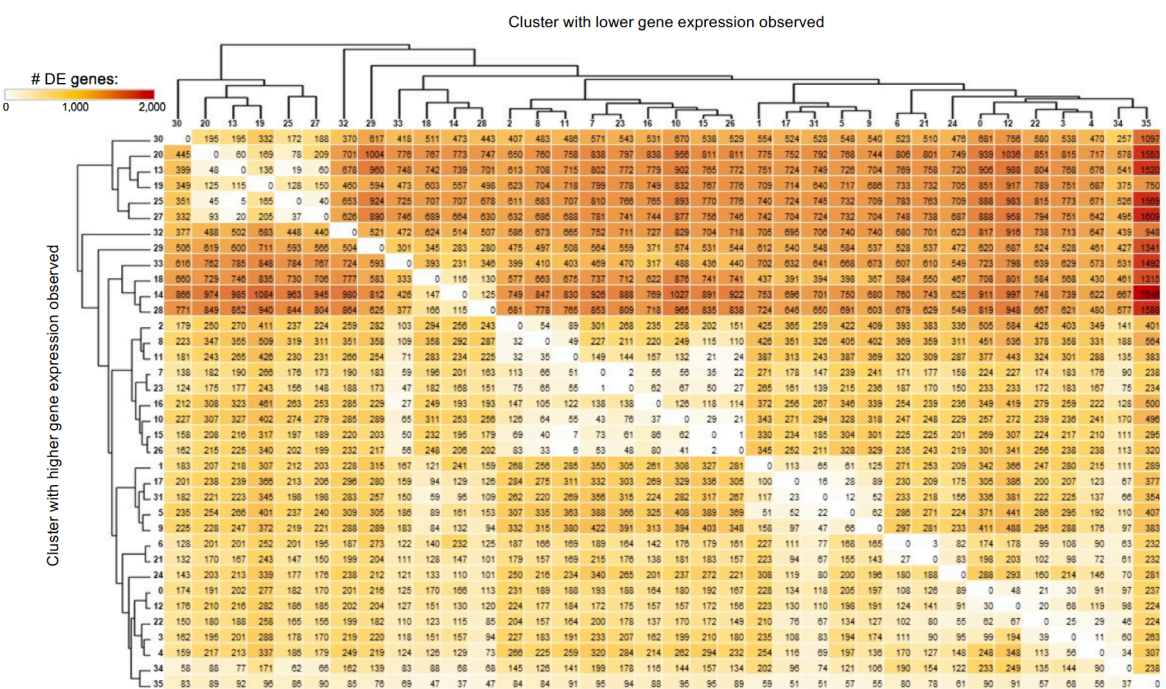
Supplementary Figure 7. RNA abundance of Nanostring selected genes. Box plot showing log(TPM) expression values of Nanostring selected genes of porcine sorted cells.

Supplementary Figure 8. Bulk RNA and Nanostring correlation. Spearman rank correlation of absolute normalize TPM values of common genes between bulk RNA-seq and Nanostring analysis.

A

Frequency within cluster

Pig ID:
- D2
- D1
- B1
- B2
- C3
- C2
- C1

B

Relative Expression: -1.5 0 1.5 3.0

scRNA-seq cluster:

Gene

Cluster with lower gene expression observed

C

DE genes: 0 1,000 2,000

Cluster with higher gene expression observed

Supplementary Figure 9. Unique characteristics of scRNA-seq clusters. A) Stacked barplot showing proportion of cells within each cluster derived from different PBMC samples. Bar color corresponds to the animal ID of a sample. Cluster IDs are shown on the left, while the number of cells within each cluster are listed on the right. Proportions of total cells are also listed in the bottom bar. **B)** Heatmap of the top DEGs within scRNA-seq clusters. Clusters are shown on the x-axis, along with a tree showing hierarchical clustering. Genes are listed on the y-axis. Up to the top five differentially expressed genes with highest positive logFC values for each cluster are listed. **C)** Heatmap showing the number of DEGs detected between scRNA-seq clusters by pairwise DGE analyses. Clusters are ordered on the x and y axis according to hierarchical relationships that were also shown in B. The cluster with higher gene expression for a pairwise comparison is shown on the y-axis, while the cluster with lower gene expression by pairwise comparison is shown on the x-axis. Seven PBMC samples used for scRNA-seq analysis were derived from each of three separate experiments (experiment B, n = 2; experiment C, n = 3; experiment D, n = 2) * Refer to 'Gene name replacement' methods.
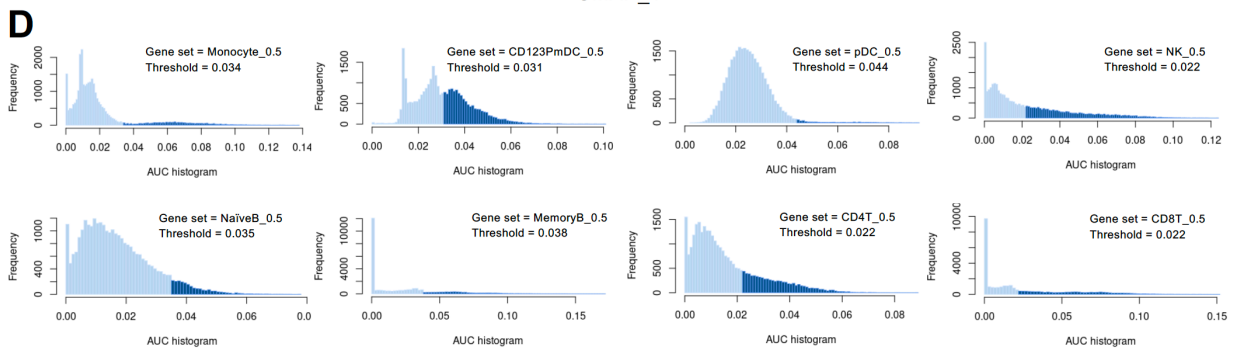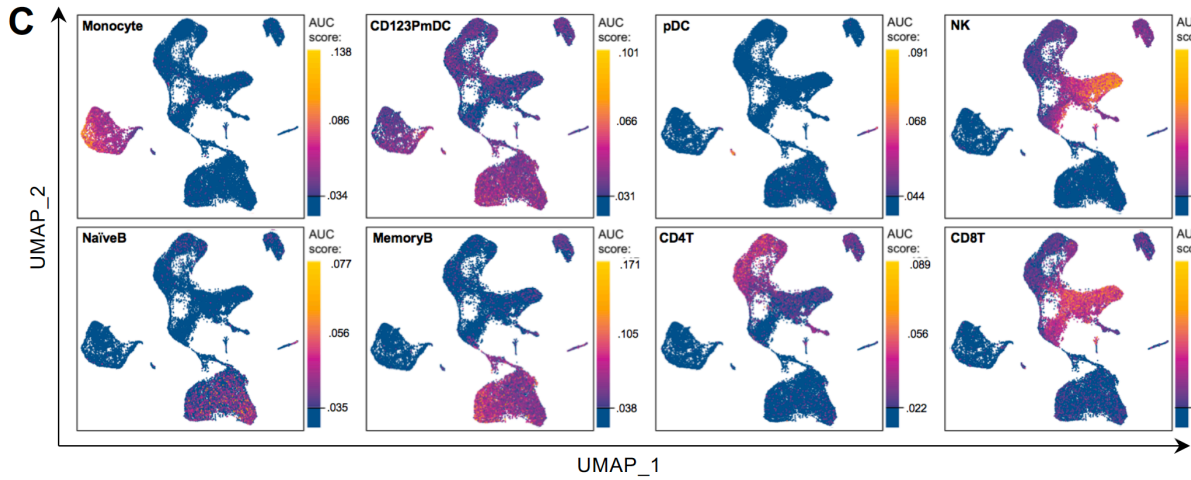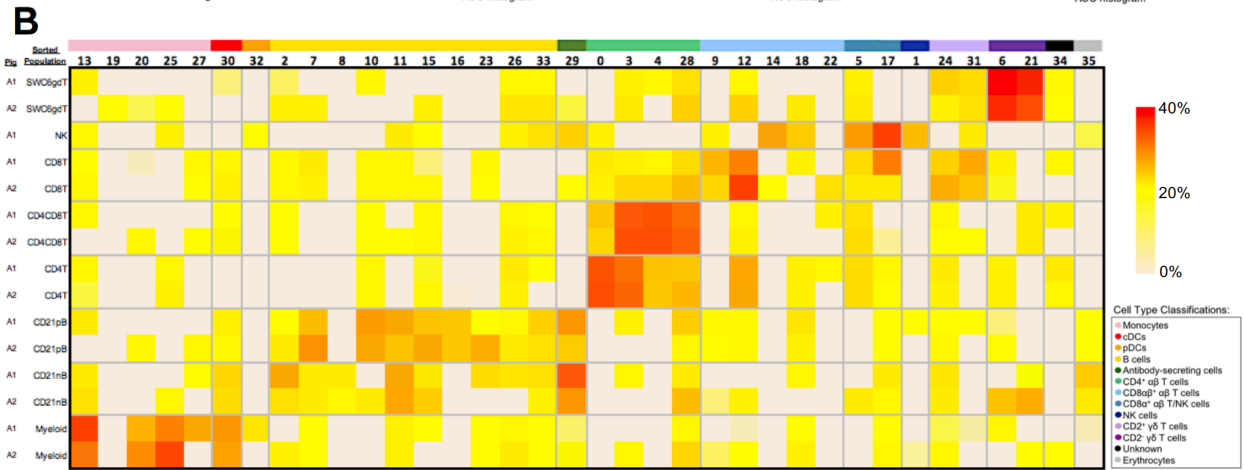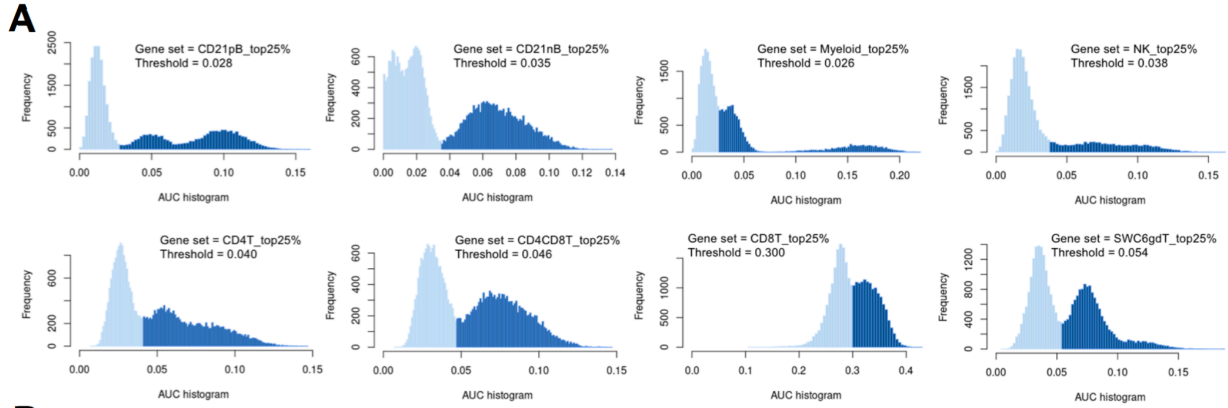
Supplementary Figure 10. Additional comparison between bulk and scRNA-seq data. **A)** Histogram of distribution of AUC scores for top 25% gene sets from porcine bulk RNA-seq sorted cell populations depicted in Figure 5 A-B. Frequency (y-axis) corresponds to the number of cells from the porcine scRNA-seq dataset having an AUC score value specified on the x-axis. Light blue bars represent cells falling below a threshold of positive AUC score detection, while dark blue bars represent cells with AUC scores above the threshold of detection used for plots in Figure 5A. Gene set and threshold of detection are provided for each respective histogram. **B)** Heatmap of CIBERSORTx deconvolution results. Porcine sorted cell bulk RNA-seq populations for each animal are listed on the y-axis, and porcine scRNA-seq clusters are listed on the x-axis. A color bar above scRNA-seq cluster IDs on the x-axis indicates the cell type classification, as according to Figure 4D. Heatmap color represents the proportion of bulk RNA-seq cells that correspond to each of the scRNA-seq clusters, with totals across rows for each bulk RNA-seq sample equaling 100%. **C)** Gene set enrichment scores calculated by AUCell analysis of enriched gene sets created from genes with high expression scores > 0.5 for each respective sorted population of human cells overlaid onto cells of the porcine scRNA-seq dataset visualized in 2-dimensional UMAP plot. Each dot represents a single cell. The color of the dot corresponds to the AUC score for each respective cell. Higher AUC scores correspond to greater gene set enrichment. A threshold for AUC score detection was set and indicated by a horizontal line on the gradient fill scale for each plot. The gene set mDC_0.5 was not included due to containing only a single gene used to calculate enrichment scores. **D)** Histogram of distribution of AUC scores from genes with high expression scores > 0.5 for each respective sorted population of human cells depicted in Figure 5 D & Supplementary Figure 10C. Frequency (y-axis) corresponds to the number of cells from the porcine scRNA-seq dataset having an AUC score value specified on the x-axis. Light blue bars represent cells falling below a threshold of positive AUC score detection, while dark blue bars represent cells with AUC scores above the threshold of detection used for plots in Figure 5A. Gene set and threshold of detection are provided for each respective histogram.
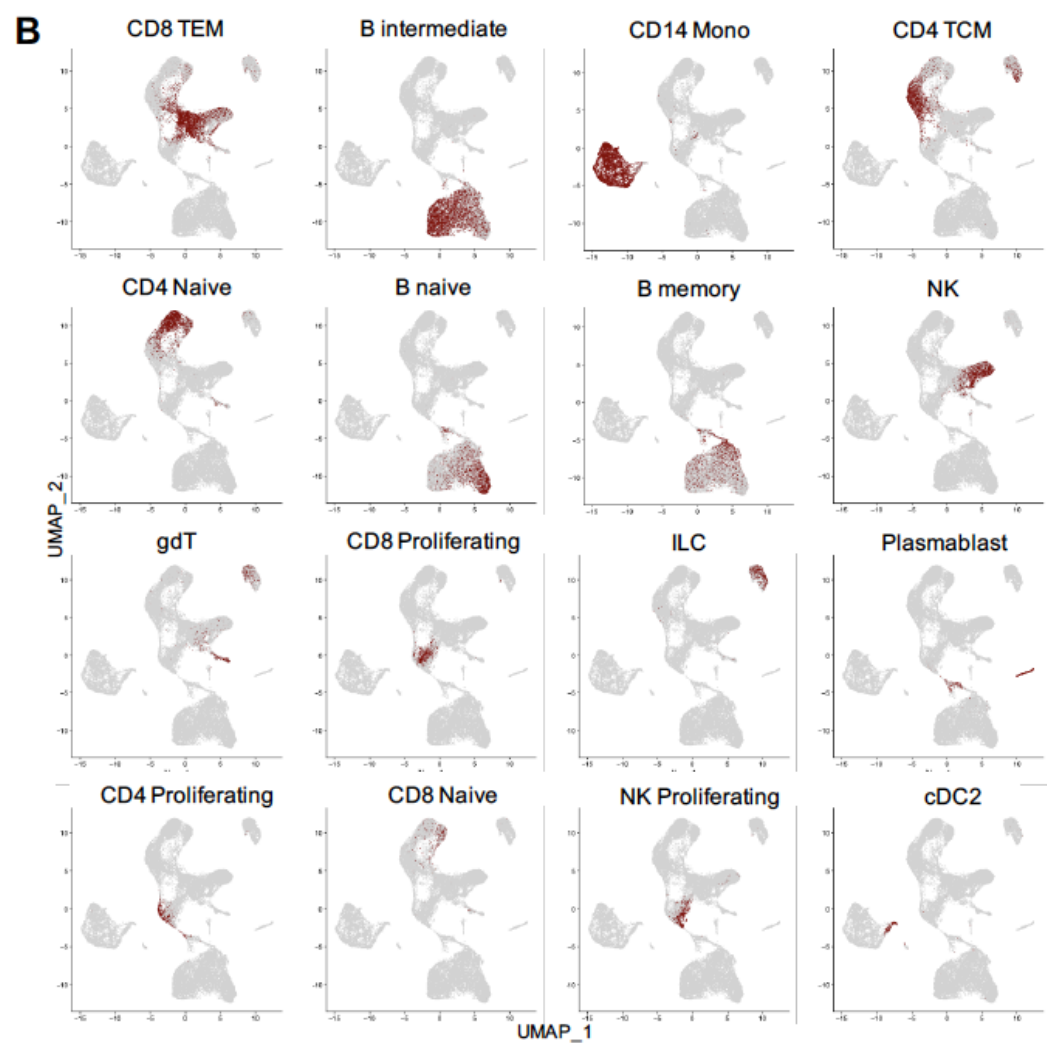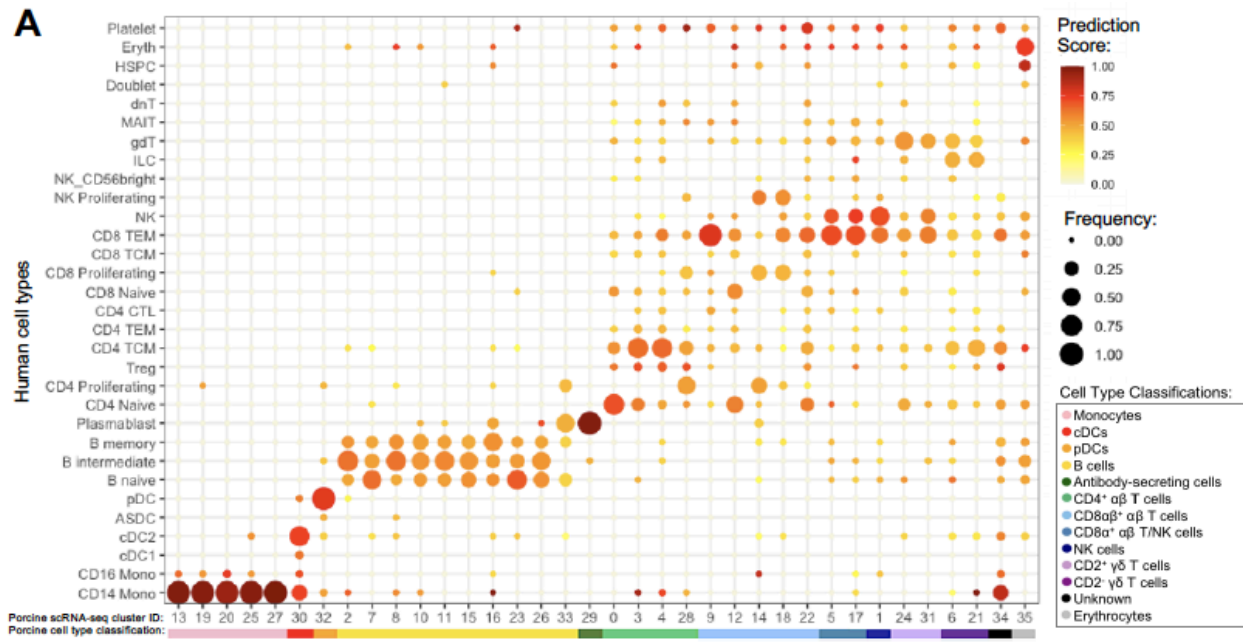
Supplementary Figure 10. Integration of porcine and human scRNA-seq datasets and transfer of human cell type labels. **A)** Prediction scores calculated to determine the confidence of cell type annotations. The human cell type specific frequency (size of the circle) and prediction score for that human cell type (color) are shown for each porcine scRNA-seq cluster. Porcine cell type classifications (color) are shown below the porcine scRNA-seq cluster IDs. **B)** The transferred human cell type labels are shown on the original porcine dataset UMAP. Each tile highlights a set of porcine cells annotated with their highest confidence level human cell type label in dark red.