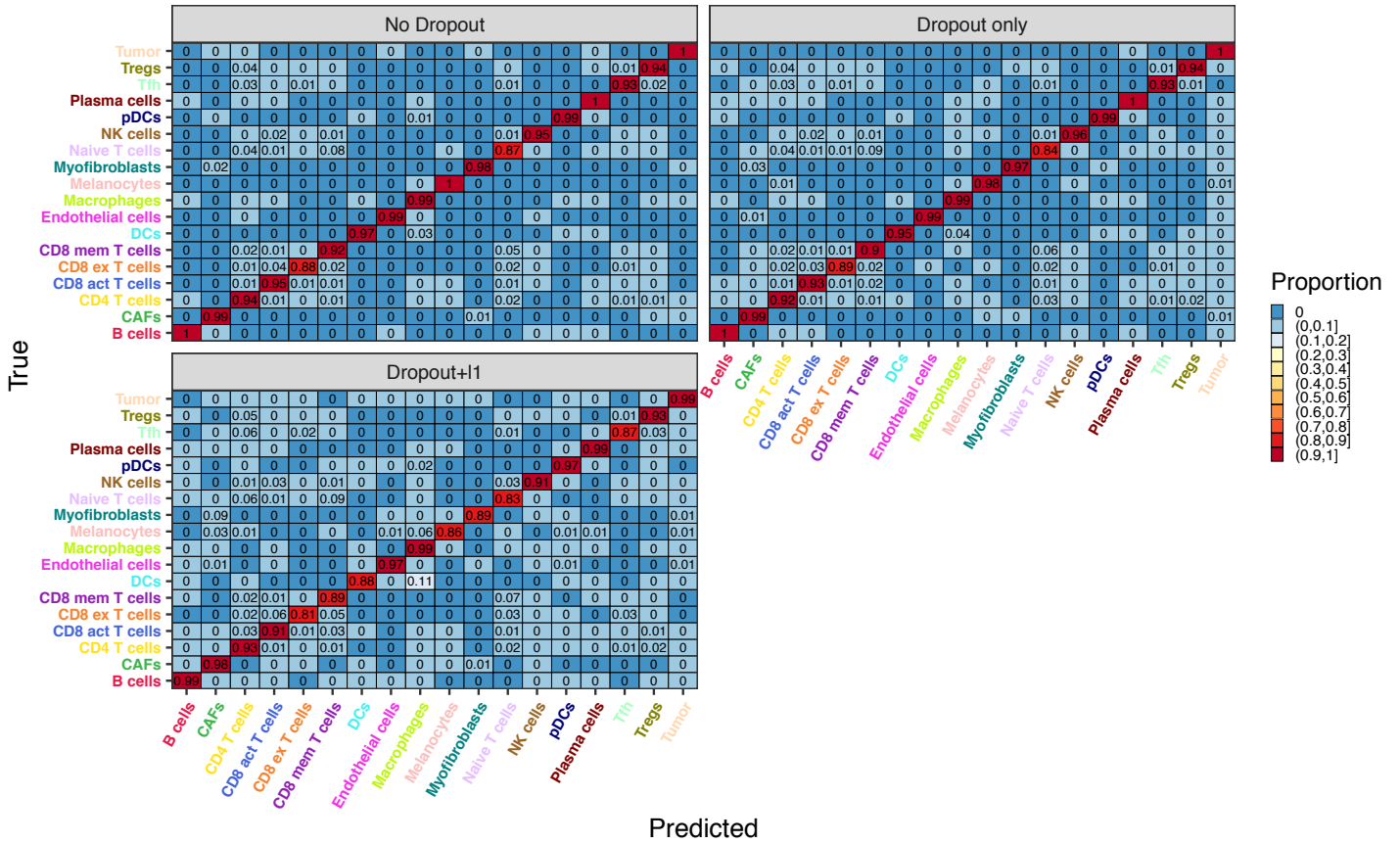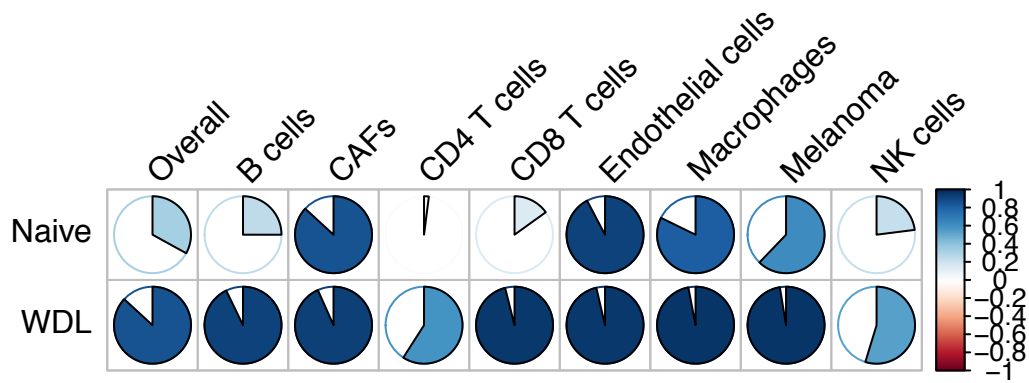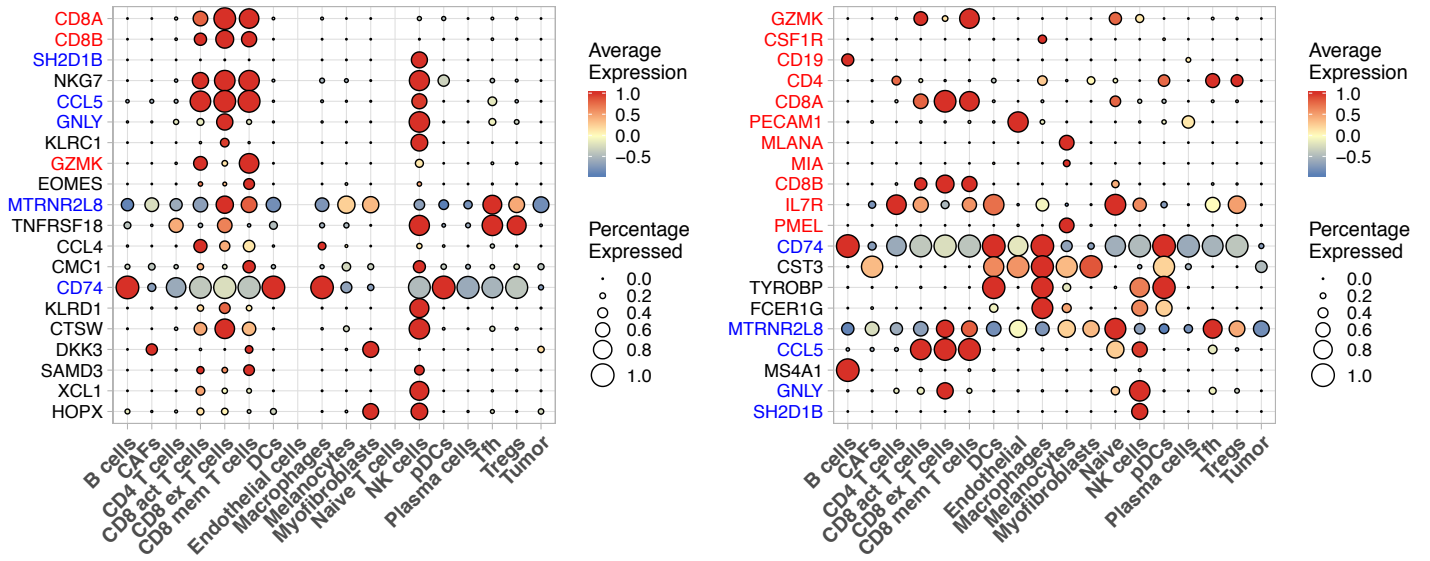Supplementary Figure 1. UMAP of the Yost et al. basal cell carcinoma single cell RNA-sequencing dataset colored by cell types.
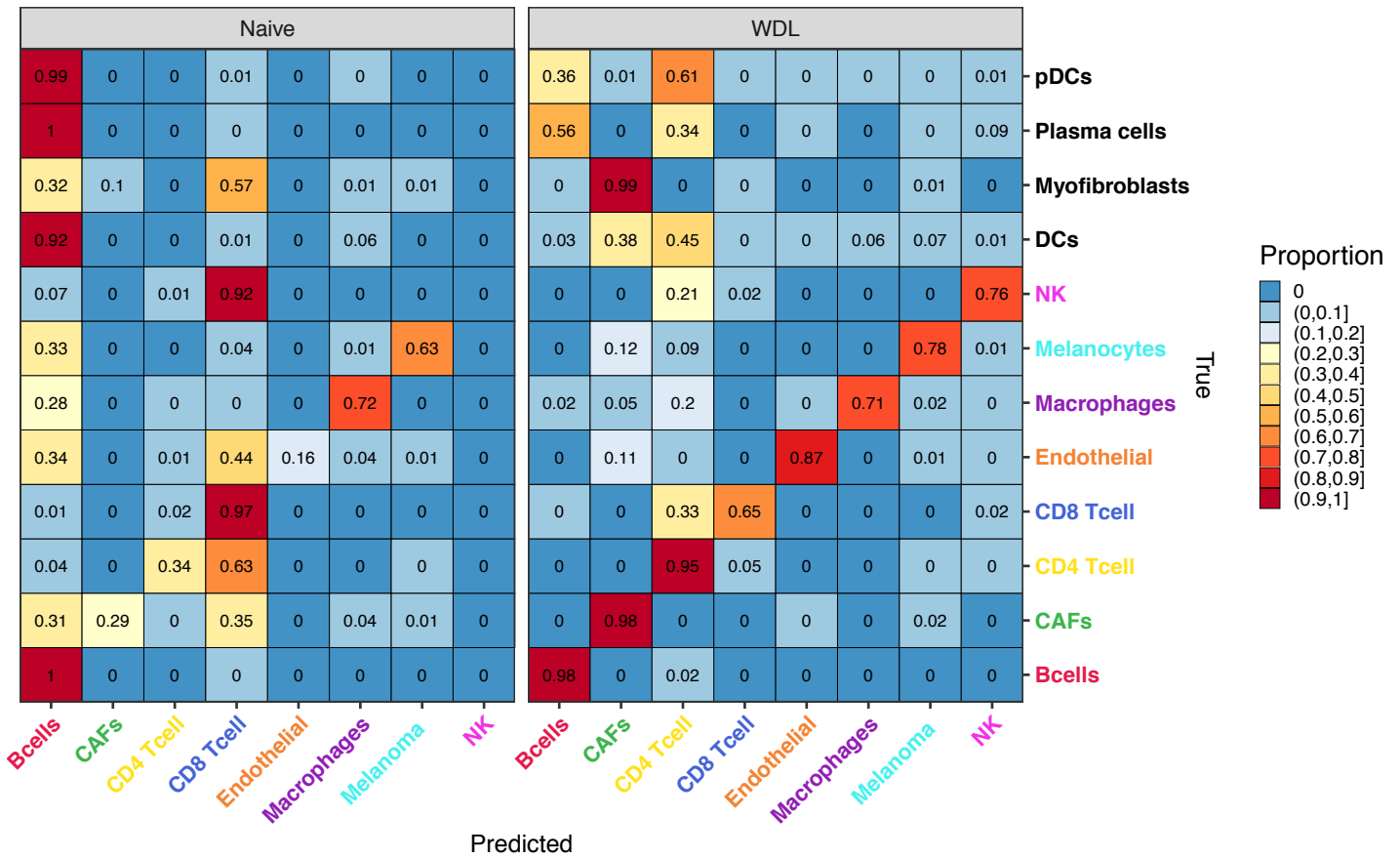
Supplementary Figure 2. Heatmaps of accuracy by cell type of models trained with no dropout (No Dropout), 20% dropout (Dropout Only), and 20% dropout and $l1$ regularization (Dropout + $l1$) using Yost et al. data. Accuracy was calculated by averaging across 10 replications in each model.

Supplementary Figure 3: Pie chart representations of the overall and by cell type accuracies for both naive and WDL model using Yost el al. data as training and Tirosh et al. data as testing. Accuracy was calculated by averaging across 10 replications in each model.

Supplementary Figure 4: Dot plots for the Yost el al. data showing the expression of top 20 markers in naive (A) and WDL (B) models, related to Figure 5. The genes names highlighted in red correspond to genes that were included in the wide part of the WDL model, and the blue gene names correspond to the genes that were not in the wide part yet were influential in both the naive and WDL models.

Supplementary Figure 5. Heatmaps of accuracy by cell type of models trained using Tirosh et al. and tested on Yost et al. data. Accuracy was calculated by averaging across 10 replications in each model. Naïve model, Dropout+l2 regularization without wide elements; WDL model, Dropout+l2 regularization with wide elements as genes listed in Figure 1.

Supplementary Figure 6. Heatmaps of the accuracy by cell types of Garnett, CHETAH, and SingleR in both intra-dataset (left) and inter-dataset (right) experiments. Row, true cell type; Column, predicted categories. The predicted categories that exist in testing data are labeled in colors on Y-axis; the predicted categories that do not exist in testing data, as well as the intermediate and unassigned/unknown categories, are labeled in black on Y-axis. Within each true cell type, percentage of cells predicted into different categories are shown in color, where red indicates higher percentage and blue indicate lower percentage. The diagonals show the percentage of cells that were correctly predicted by the models.