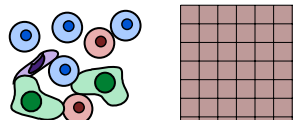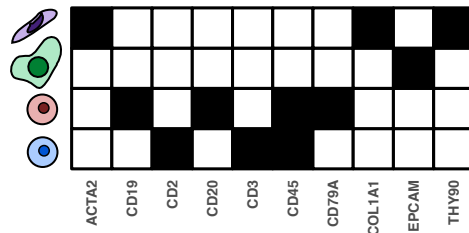# Expression count matrix for heterogeneous population

Prior knowledge of marker genes

T cell: CD2, CD3, CD45

B cell: CD19, CD20, CD79A, CD45

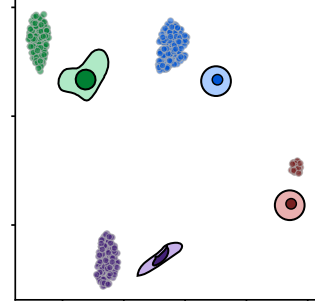Tumor cell (epithelial): EPCAM

Fibroblast: COL1A1, ACTA2, THY90

Binarization

CellAssign: automated probabilistic assignment of cells to known cell types

$$\log \mu_{ngc} = \log s_n + \delta_{gc}\rho_{gc} + \beta_{g0} + \sum_{p=1}^{P} \beta_{gp}x_{pn}$$

Cell type assignments

Probabilities