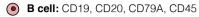
a

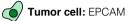
Expression count matrix for heterogeneous population



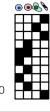
Prior knowledge of marker genes

T cell: CD2, CD3, CD45





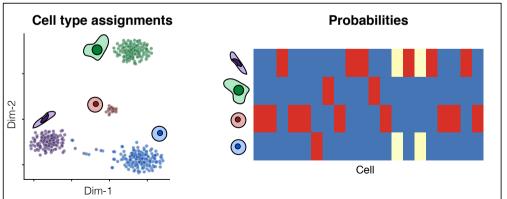
Fibroblast: COL1A1, ACTA2, THY90



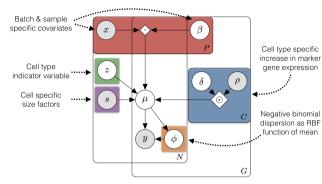


CellAssign: automated probabilistic assignment of cells to known cell types









Variable	Distribution	Description
y_{nc}	Negative Binomial	Single-cell counts
s_n	None	Cell size factor
z_n	Categorical	Cell type indicator
ϕ_{ngc}	Deterministic f ⁿ	Negative binomial dispersion
δ_{gc}	Log-Normal	Marker over-expression
$ ho_{qc}$	None	Marker-cell type matrix
x_{np}	None	Covariates (batch or sample)
β_{pg}	Gaussian	Covariate coefficients
a, b	None	Dispersion basis coefficients
		-