

Isolating the perturbation response of gene regulatory networks in the presence of biological variability and technical noise

Clayton W. Seitz

May 12, 2022

Drug-induced reprogramming as a mode of cancer drug resistance

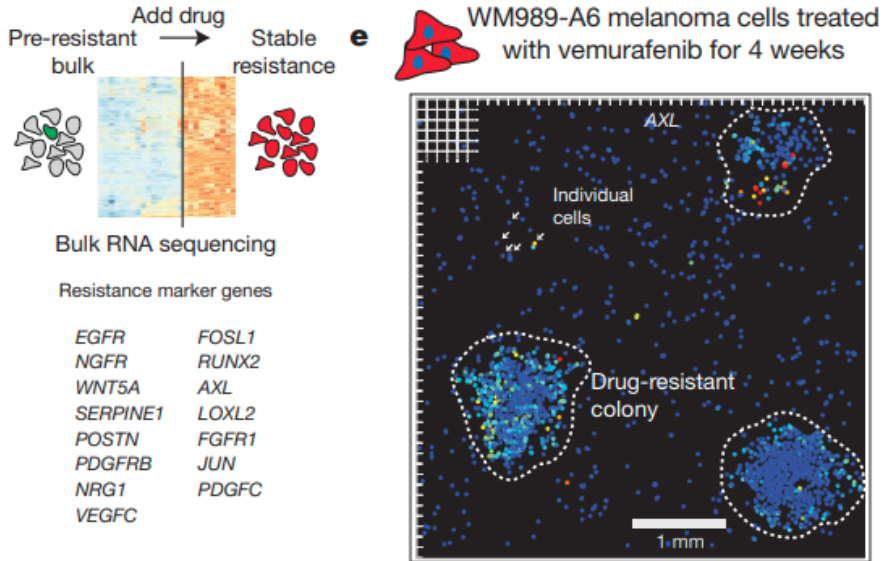
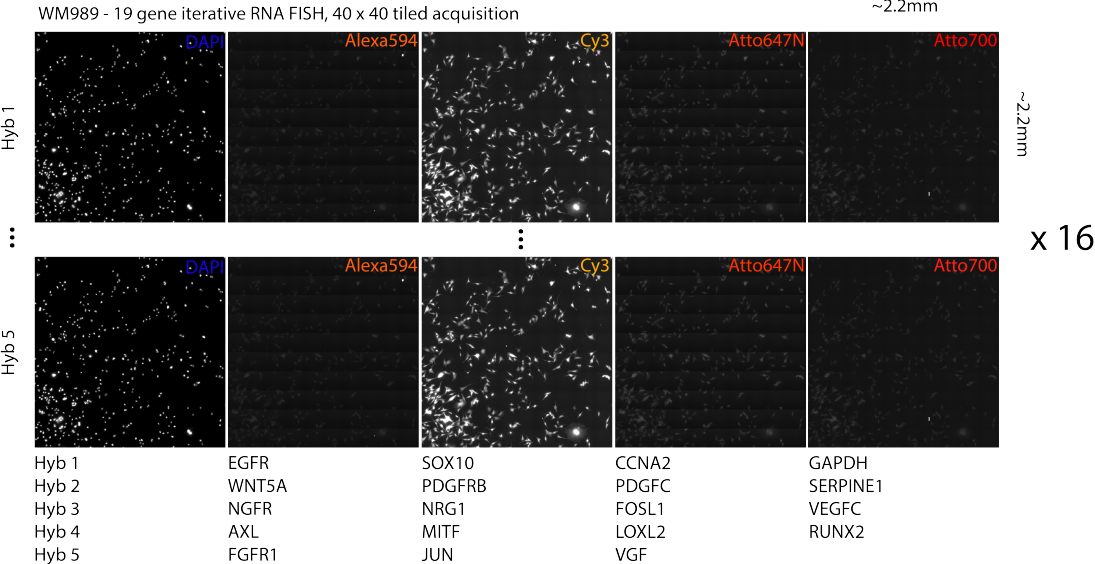


Figure 1: Shaffer et al., Nature 2017

WM989-A6 RNA-FISH data summary



Graphical abstract

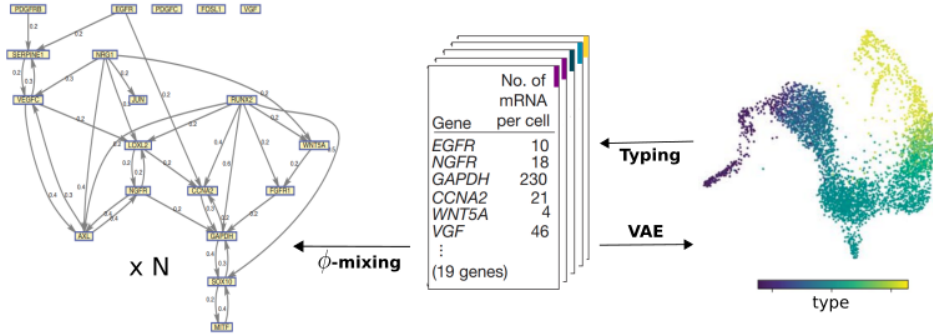
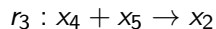
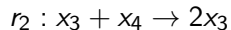
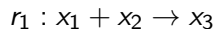


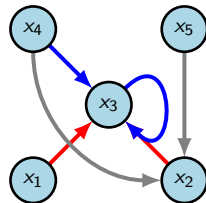
Figure 2: Gene expression matrices used as training data to learn a latent-space representation of gene expression, uncovering latent structure of the joint distribution and permitting cell typing, account for batch variability. Type information is then used for inference of the underlying regulatory network using the ϕ -mixing coefficient, which may differ across types

Modeling stochastic biochemical reaction networks

Three ways of representing a biochemical reaction network



$$\nu = \begin{pmatrix} -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 & -1 \end{pmatrix}$$

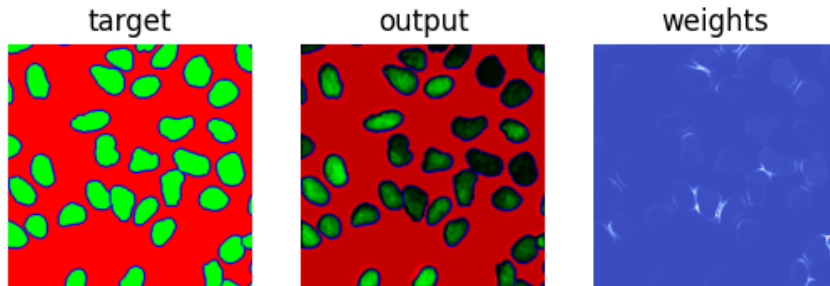


Bayesian parameter inference for gene regulation

Suppose we

Training on BBBC039 U2OS Cells

BBBC039: 200 images, 160 train + 40 validation, 256 x 256 random crop

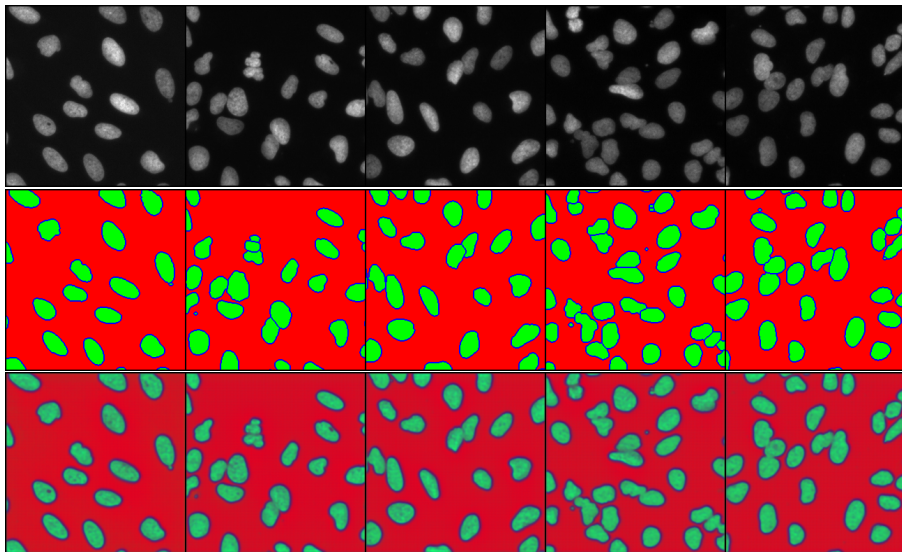


We train a 3-channel semantic segmentation model with **weighted** cross-entropy loss:

$$\mathcal{L} = \sum_{i,j} w_{ij} \log p_{ij}(\tilde{x}) = \sum_{i,j} w_{ij} \log \frac{\exp(-s_{ij}(\tilde{x}))}{\sum_{x \in \chi} \exp(-s_{ij}(\tilde{x}))}$$

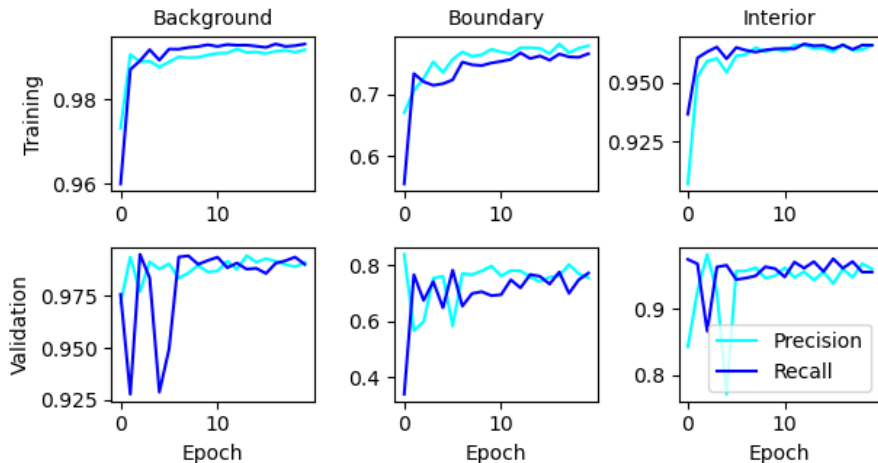
p_{ij} is the probability the model assigns a pixel to the true class $\tilde{x} \in \{a, b, c\}$

Training on BBBC039 U2OS Cells



Training on BBBC039 U2OS Cells

Learning rate $\eta = 0.01$, Batch-size $B = 5$ (32 train iterations, 8 validation)



References I