

Chaining cell state classifiers for cancer analysis

Daniel Carlin and James Durbin
UC-Santa Cruz Biomolecular
Engineering

Motivation

Module Map of Stem Cell Genes Guides Creation of Epithelial Cancer Stem Cells

David J. Wong¹, Helen Liu¹, Todd W. Ridky¹, David Cassarino², Eran Segal³,  , Howard Y. Chang¹,  



The Stem Cell Identity of Testicular Cancer

Amander T. Clark

Nature Reviews Cancer **6**, 846-856 (November 2006) | doi:10.1038/nrc1991

Polycomb silencers control cell fate, development and cancer

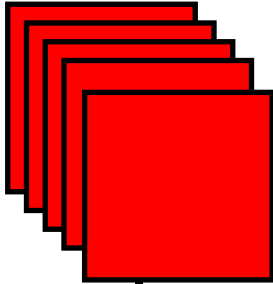
Anke Sparmann¹ & Maarten van Lohuizen¹ [About the authors](#)

Epithelial-Mesenchymal Transition in Breast Cancer Relates to the Basal-like Phenotype

David Sarrió¹, Socorro María Rodríguez-Pinilla¹, David Hardisson²,
Amparo Cano³, Gema Moreno-Bueno³, and José Palacios⁴

Classifier Chaining

Cancer Expression Datasets



Preprocessing: Mapping to standard set of genes and Quantile Normalization

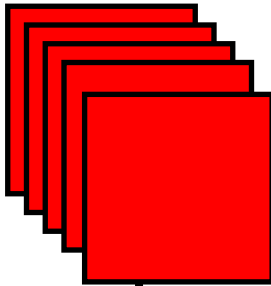
Stem Cell Expression (497 samples)
Early Multipotent vs ESC



Build Classifier
Choose TopModel

Classifier Chaining

Cancer Expression Datasets



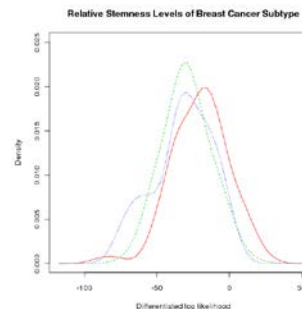
Preprocessing: Mapping to standard set of genes and Quantile Normalization

Stem Expression (497 samples)
Early Multipotent vs ESC



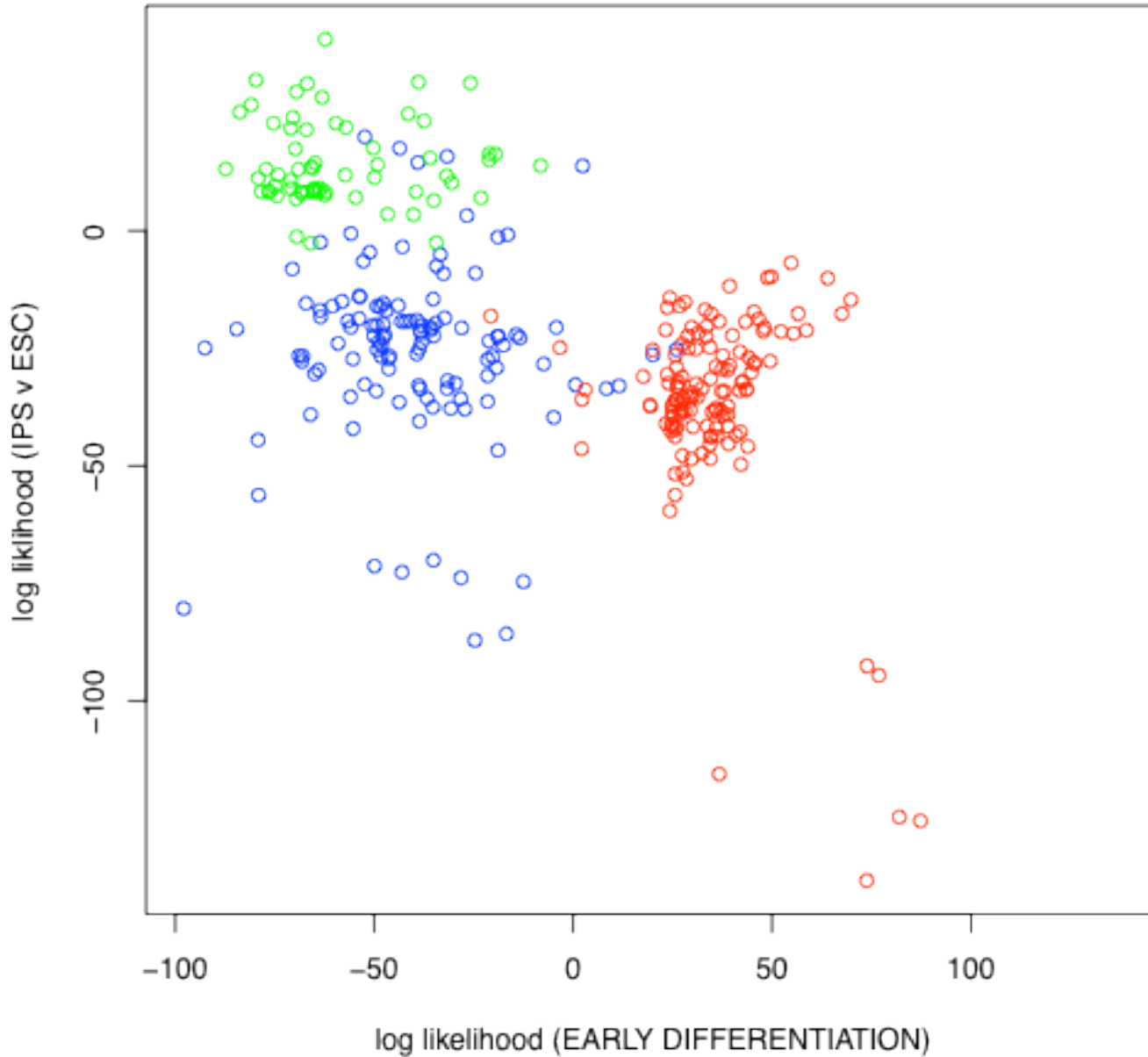
Build Classifier
Choose TopModel

Classify Cancer
using the
stemness
signature



Estimates global stemness for each cancer sample

Robust stem cell classifiers

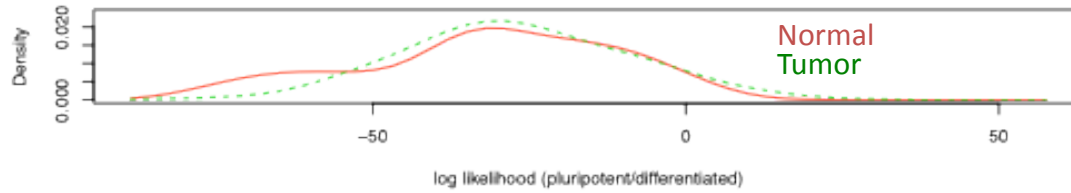


Induced pluripotent stem cells

Embryonic stem cells

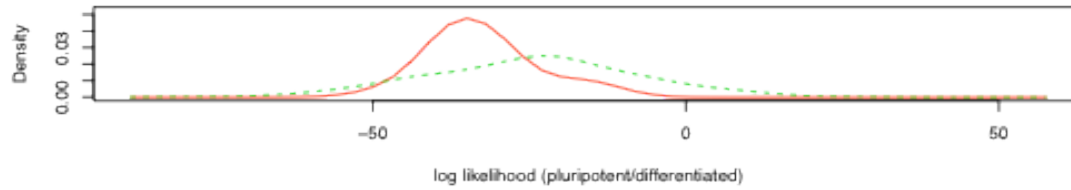
Adult stem cells and early progenitors

Stemness Levels of BRCA



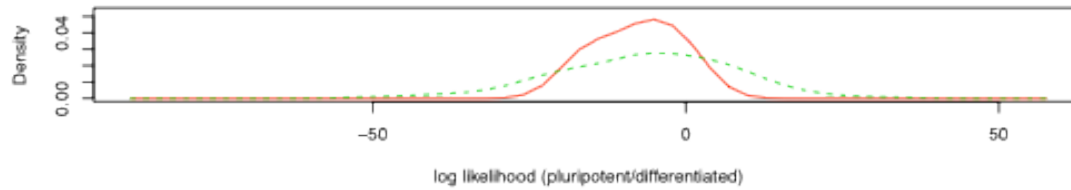
p-val=0.07

Stemness Levels of COAD



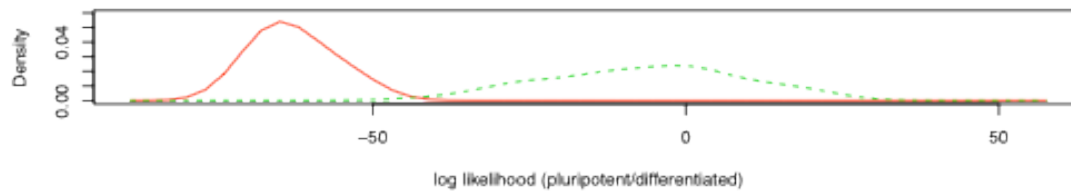
p-val=1.4E-4

Stemness Levels of GBM



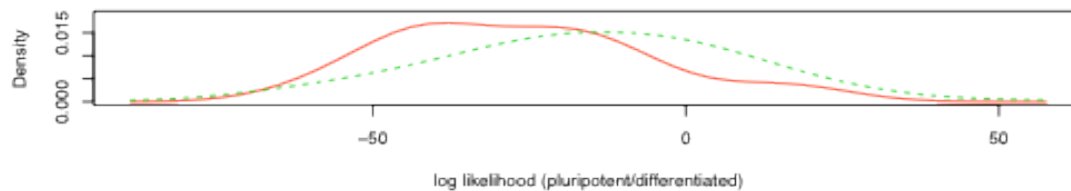
p-val=0.14

Stemness Levels of LUSC



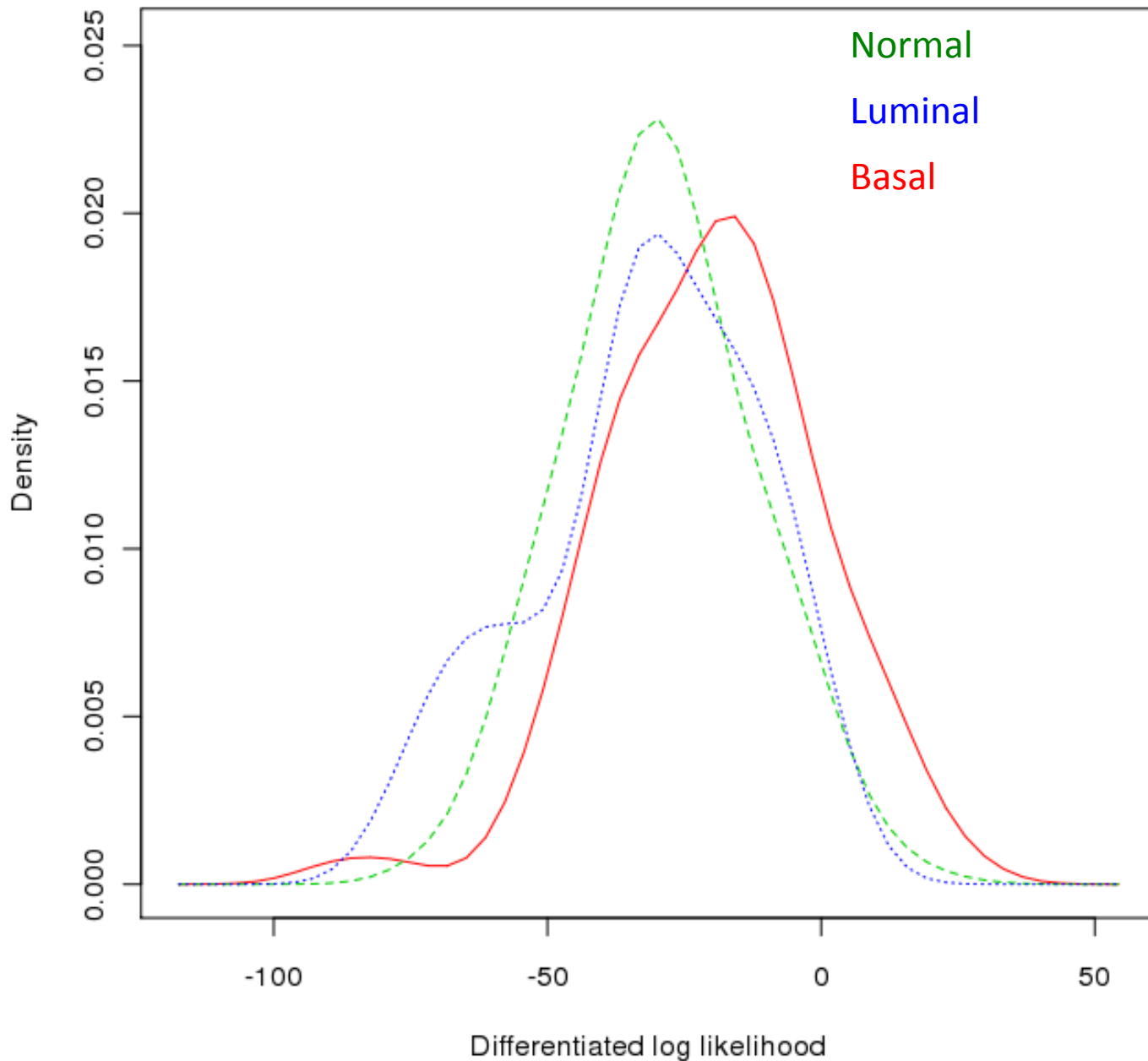
p-val=6.2E-5

Stemness Levels of OVCA



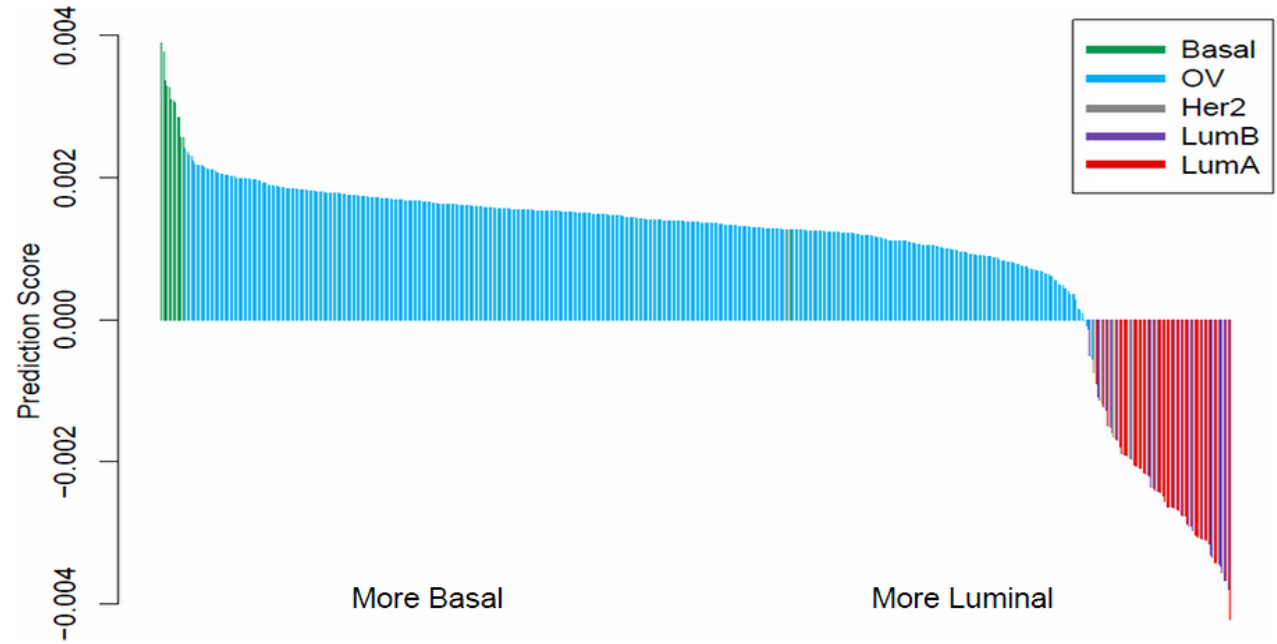
p-val=0.13

Relative Stemness Levels of Breast Cancer Subtype

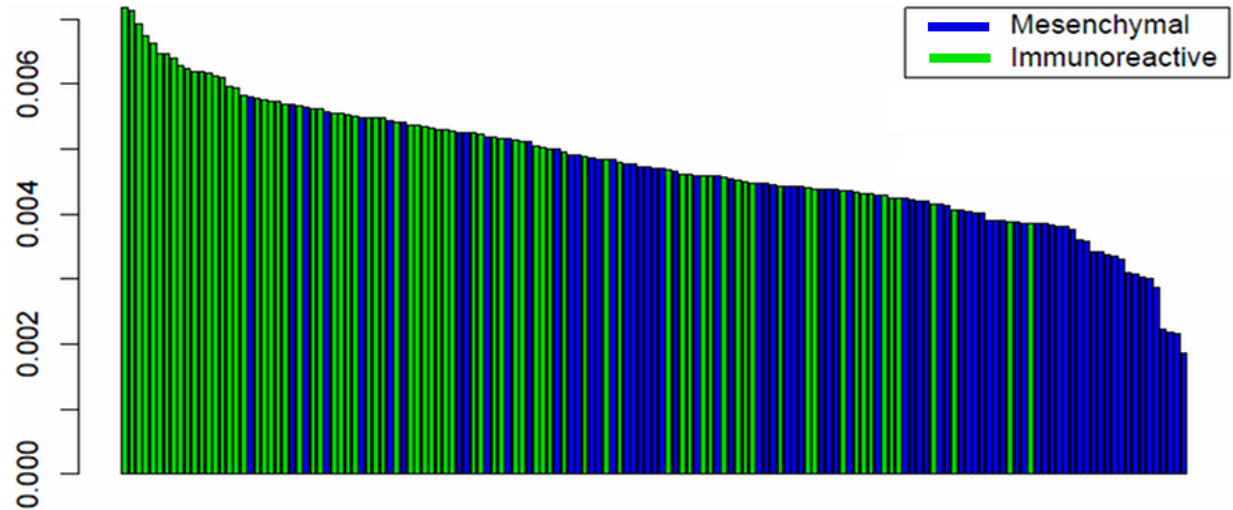


p-val for
difference in
distribution
between Basal
and Normal:
0.004242

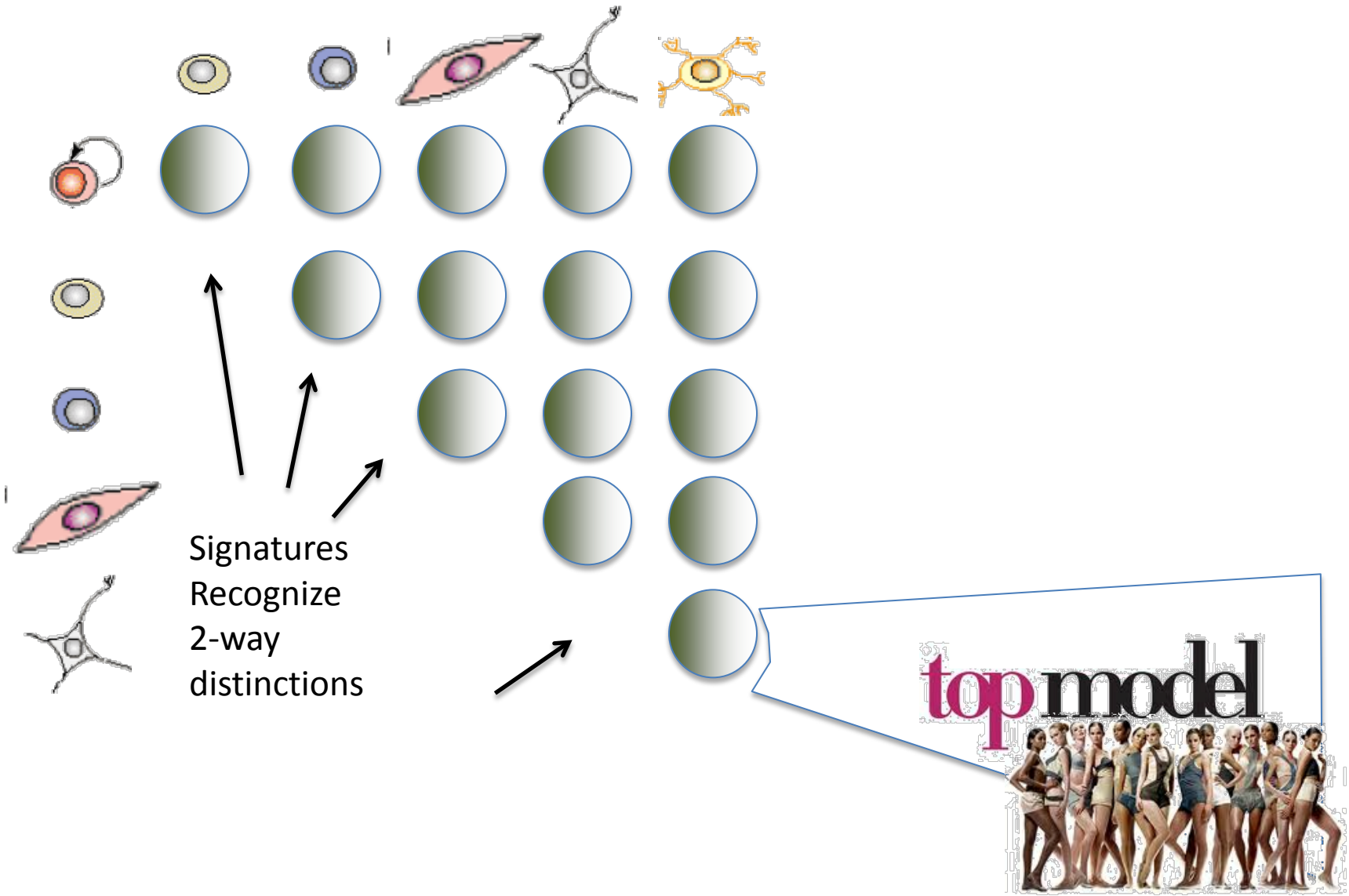
Trained on 80% of BRCA data to learn Basal vs Luminal, then applied to the remaining 20% BRCA and all OVCA



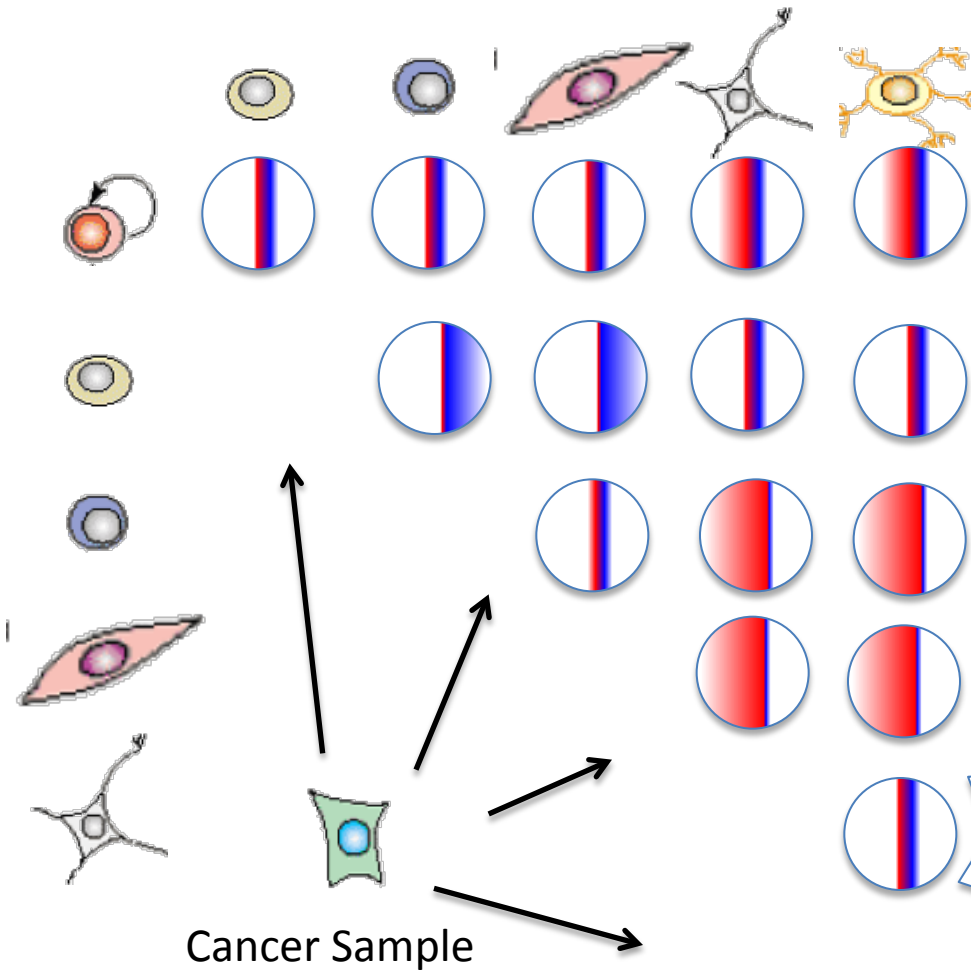
Just the OVCA:



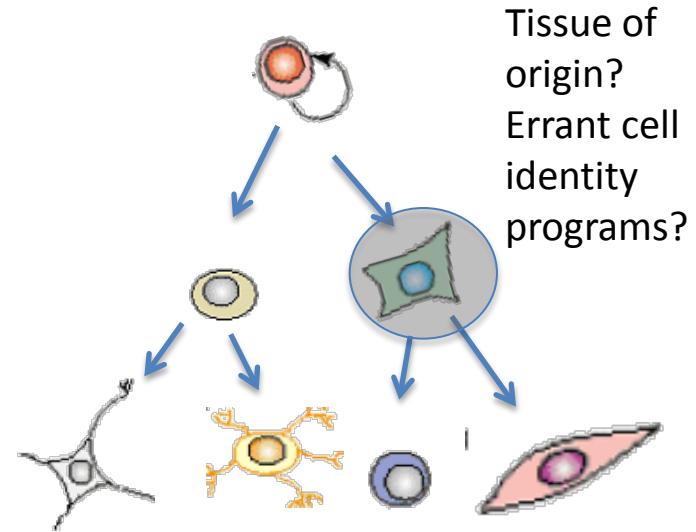
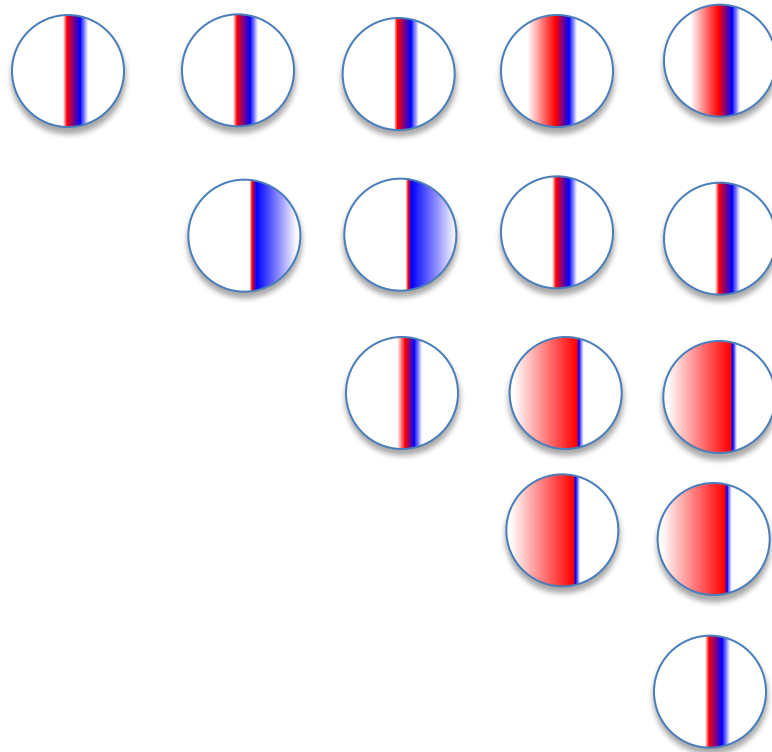
Gene signatures from binary classifiers



Gene signatures from binary classifiers



Cell signatures from gene signatures



Tissue of origin?
Errant cell
identity
programs?

Thanks

- Josh Stuart
- David Haussler
- Chris Szeto
- Artem Sokolov
- Sam Ng