

Identifying Phenotypically Similar Clusters Between Single Cell RNA-Seq Samples

Guilherme de Sena Brandine

May 10, 2017

1 Introduction

With the increasingly large throughput of single cell RNA-Seq technologies, it is possible to understand the transcriptomes heterogeneous populations of cells in tissues in single-cell resolution. These transcriptomes can be used to identify the phenotype of each sampled cell through diverse methods, such as the expression of specific known marker genes or unsupervised clustering methods based on expression similarity.

A question that arises with these technologies is how biological systems differ between two samples, for instance: The differences between the phenotypes of a healthy and a diseased tissues or the differences in a tissue on distinct timepoints of development. We would like to identify which phenotypes are present in both samples and which are specific to a single one. This could be useful not only to better understand the changing compositions of cell types but also to interrogate the difference between the gene expression of cells of the same phenotype when in different environments.

We consider two groups of cells in two different samples to be phenotypically equivalent if their expression profiles are similar enough that we can

2 Clustering Methods Used in Single Cell Analysis

2.1 Louvain Modularity

2.2 DBSCAN

2.3 Gaussian Mixture Models

2.4 Spectral Clustering

2.5 K-Means

2.6 Hierarchical Clustering

3 Cluster Comparison Methods

3.1 CCLUMP

3.2 CODENSE

4 Data

Paper	Organism	Clustering	# Clusters	# Cells	# Cells Details
Yan	hg19	Biological	7	129	Oocyte, Zygote, 2-cell, 4-cell, 8-cell, Morula, Blastocyst
Biase	mm10	Biological	3	9+10+5 = 24	hESC, 2-cell, 4-cell
Goolam	mm10	Biological	5	40	2-cell, 4-cell, 8-cell, Morula, Blastocyst
Deng	mm10	Biological	10		Oocyte, Zygote, 2-Cell(E,M,L), 4-cell, 8-cell, 16-cell, Blastocyst(E,M,L)
Pollen	hg19	Biological	11	301	Oocyte, Zygote, 2-cell, 4-cell, 8-cell, Morula, Blastocyst
Kolodziejczyk	hg19		7		Oocyte, Zygote, 2-cell, 4-cell, 8-cell, Morula, Blastocyst
Klein	mm10				
Macosko	hg19				
Shekhar	hg19				
Zeisel	mm10				