Defining novel multilineage progenitor populations using single-cell RNA-Seq

Nathan Salomonis
Department of Biomedical Informatics, CCHMC
Uncovering Cell Heterogeneity from Single-Cell RNA-Seq

- **Major Objectives**
  - Define major cell populations.
  - Identify ultra-rare cells.
  - Identify hidden cellular heterogeneity and transition states.
Within the bone marrow, stem cells differentiate and give rise to diverse blood cell types and functions. Currently, hematopoietic progenitors are defined using surface markers combined with functional assays that are not directly linked with in vivo differentiation potential or gene regulatory mechanisms. Here, we comprehensively map myeloid progenitor sub-populations by transcriptional sorting of single cells from the bone marrow. We describe multiple progenitor subgroups, showing unexpected transcriptional priming toward seven differentiation fates but no progenitors with a mixed state. Transcriptional differ-
Analyze using the 96-cell Fluidigm platform.
Average 3 million reads of PE 75nt reads.
Exclude outliers (depth, alignment %).
Automated Single Cell Analysis in AltAnalyze: Iterative Clustering and Guide-gene Selection

15K Genes

Pioneering Round

Filter

Correlation

Reduce Cluster Size

Select Initial Guide Genes

Select Additional Guide Genes

Optimize Clusters

Final Cluster Using All Guide Genes

Correlate Guide genes to all genes

Optionally Exclude Cell Cycle Effects

Integrate Cell Type Predictions

All detected genes

Kallisto in AltAnalyze

FASTQ files or pre-processed
Molecular Dissection of Hematopoiesis from scRNA-Seq using ICGS

Monocyte and Granulocyte Progenitors Defined by Opposing Transcription Factors

Gfi1 and Irf8 Interact on Promoters to Regulate Myeloid Specification

Intermediate Gfi1 and Irf8 Define a Metastable Bipotential Progenitor Population

1. scRNA-Seq combined with TF-deletion and ChIP-Seq can define transcriptional regulatory networks.

2. Appears to require deep RNA-Seq.

3. Non-HSC, Multi-Lineage progenitors are frequently found in CMP and GMP gated cell populations.

4. These cells are primarily defined by multi-lineage gene priming and only weakly defined by unique marker genes.

5. Multi-Lin’s can be captured and enriched by sorting for progenitors with dual lineage programs (Gfi1 and Irf8 expression).

6. Genetic deletion of these factors traps cells in an undecided state.
Controversies and Questions from scRNA-Seq Predictions

1. Other myeloid biologists argue Multi-Lin’s are technical artifacts (doublets).
2. Bi-Potential progenitors enriched but not purified.
3. Multi-Lineage progenitors and MEPs not identified.
Optimized Isolation of CMP Multi-Potential Progenitors

Manuscript in preparation
New Algorithms to Predict Multi-Lineage States from scRNA-Seq (Schrodinger)

Step 1
ICGS Guide-Results

Step 2
ICGS Ordered Marker Genes

Step 3
Coincident Lineage Markers

Step 4
Aggregate Coincident Lineage Profiles

Step 5
Predicted Mixed-Lineage Cells and Cell States

Multilineage Score
Olsson

Z-score

0 2 4 6 8

- Multi-Lin
- MDP
- Meg
- Meg-2
- Eryth
- HSCP-2
- NK
- Nk
Identification of similar Multi-lineage states from other scRNA-Seq datasets (ICGS + examination of lineage markers).

>200k reads/library
1.8k cells
Schrodinger Detection of Multi-Lineage States in ~500 Mouse Bone Marrow Progenitors (Nestorowa et al.)

Also validated in data from:
Can This Approach Be Used to Find Similar Cell Populations Across Technologies?

E14 Kidney Dissociation

Multi-Platform Single-Cell Analysis

DropSeq

Fluidigm

10X Genomics

Unsupervised Analysis

(Cell Cycle Effects Excluded)

Supervised Classification

Interactive Technology Comparison

Cellular Interactions

Population Description

Magela et al. 2017 Developmental Biology
Consistent Schrodinger Prediction of Multi-Lineage States in Embryonic Kidney
Acknowledgements

Salomonis Lab

Meenakshi Venkatasubramaniam
Kashish Chetal

Grimes Lab (CCHMC)

H. Leighton Grimes
Andre Olsson

Singh Lab (CCHMC)

Harinder Singh
Virendra Chaudhri

Aronow Lab (CCHMC)

Bruce Aronow
Phil Dexheimer

Potter Lab (CCHMC)

Steve Potter
Mike Adam
Bliss Magella

Funding

NIH (R01HL122661 – Grimes
CCHMC Center for Pediatric Genomics Award
Identifying Bi-Potential Megakaryocyte-Erythroid Intermediates from Human scRNA-Seq
Monocle Trajectory Analysis of MarkerFinder
Results Finds Distinct MEP Subsets
Schrodinger Accurately Predicts that MEPs are Mixed-Lineage Progenitors

Mixed-c1: SLC44A1, ACTN1, TTC27, PDLIM1, FERMT3, TMSB4X, ITGA2B, ABCB1, MYH9, PTGS1, STOM, CTR9, SH2B3, C6orf25, CXCL8, TUBA4A, PKM, PLEK, CD9, GSN

Mixed-c2: CA1, MPC2, SF3B3, CD55, FBXO7, PRDX2, CD36, SDCBP, SPTA1, ANXA2, APOC1, RB1CC1, CALM2, WDR48, ACSM3, GOLGA4, ELL2, HBS1L, FAM45A, IARS, DLD, AHI1, SLC39A8, SKIL, ACSM1, BLVRB, SEC22C, CXADR, IRF1, ZDHHC2, USP12, RHOBTB3, EZR, KIT, FBXO34, ANK1, YBX1, STARD7, PDZD8

ERP: CNRIP1, TMEM14C, SLC40A1, FAM118A, RYR3, CASP3, U1, ZFP36L1, ELOVL6, HERC2P2, TRIB2, MYC, RREB1, P2RX5, SNORD3A

MEP: HSD17B11, FCER1A, RPS3AP47, CPA3, TESPA1, MEIS1, FREM1, SERPINB1, PBX1

MKP: CD52, CD74, FNB1, IDS, CD37, KIAA0125, SORL1, AJ006998.2, AHNAK, EGR1, KLF4, KLF2, NPR3, CRHB, PROM1, ADAM28, SMIM24, CLEC2B, ID2, ID3, VIM, ATP8B4