Bringing single cell genomics closer to the clinic for patients with leukemia

Scott Furlan, MD (he/his) - 7-29-22
Assistant Professor, FHCRC
Overview and Disclosures

• Introduce the concept of measurable residual disease (MRD) as told from the of one patient

• Provide rationale for the use of single cell genomics to potentially improve MRD assessment after transplant

• Share our preliminary data using single cell RNA seq in patients with relapsed leukemia
  • Highlight novel molecular and computational approaches

• Broader applicability to understanding the biology of acute leukemia and mechanisms of relapse

No Conflicts of Interest or relevant disclosures!
Thank you ___
A bedside to bench approach to a clinical conundrum

- 12 yo with a history of Myelodysplastic Syndrome diagnosed in late 2018
- Evolved to Acute Myeloid Leukemia (AML) shortly thereafter
- Matched Unrelated Bone Marrow Transplant in April 2019
- Two years later (April 2021), patient developed low blood counts on routine monitoring -> AML

A bedside to bench approach to a clinical conundrum

• After relapse, underwent reinduction chemotherapy and was being considered for a second transplant using cord blood.

• Flow cytometry 1-2%

• But his TP53mut (R248Q), IDH1mut (R132C) > 10% VAF.
A bedside to bench approach to a clinical conundrum

- After relapse, underwent reinduction chemotherapy and was being considered for a second transplant using cord blood.
- Flow cytometry 1-2%
- But his TP53mut (R248Q), IDH1mut (R132C) > 10% VAF.

A bedside to bench approach to a clinical conundrum

• After relapse, underwent reinduction chemotherapy and was being considered for a second transplant using cord blood.

• Flow cytometry 1-2%

• But his TP53mut (R248Q), IDH1mut (R132C) > 10% VAF.

What is the current status of MRD testing?

- PCR
- NGS
- MFC
- Chimerism Testing
What is the current status of MRD testing?

- PCR based tests: Fusion transcripts
  - High sensitivity, but are limited in their applicability (A priori)
The ‘long tail’ of acute leukemia is a problem for molecular assays

- PCR based tests: Fusion transcripts
  - High sensitivity, but are limited in their applicability (A priori)
The ‘long tail’ of acute leukemia is a problem for molecular assays

- **PCR**
  - PCR based tests: Fusion transcripts
    - High sensitivity, but are limited in their applicability (A priori)

- **NGS**
  - Bulk NGS: Mutations
    - Similarly limited, less sensitive
What is the current status of MRD testing?

**PCR**
- PCR based tests: Fusion transcripts
  - High sensitivity, but are limited in their applicability (A priori)

**NGS**
- Bulk NGS: Mutations
  - Similarly limited, less sensitive
  - Also limited applicability, but lower sensitivity

**MFC**
- Single cell methods - Multiparameter flow cytometry (MFC):
  - More broadly applicable
  - Limited in sensitivity
  - Challenging to standardize and difficult to interpret
What is the current status of MRD testing?

**PCR**
- PCR based tests: Fusion transcripts
  - High sensitivity, but are limited in their applicability (A priori)

**NGS**
- Bulk NGS: Mutations
  - Similarly limited, less sensitive
  - Also limited applicability, but lower sensitivity

**MFC**
- Single cell methods - Multiparameter flow cytometry (MFC):
  - More broadly applicable
  - Limited in sensitivity
  - Challenging to standardize and difficult to interpret
What is the current status of MRD testing?

**PCR**
- PCR based tests: Fusion transcripts
  - High sensitivity, but are limited in their applicability (A priori)

**NGS**
- Bulk NGS: Mutations
  - Similarly limited, less sensitive
  - Also limited applicability, but lower sensitivity

**MFC**
- Single cell methods - Multiparameter flow cytometry (MFC):
  - More broadly applicable
  - Limited in sensitivity
  - Challenging to standardize and difficult to interpret

**Chimerism Testing**
- Not currently sensitive enough (+/- 5%)
How can we improve?

SINGLE CELL RNA SEQUENCING MAY BE ABLE TO MEET THIS NEED!

Deliverable:
More CONFIDENT assessment of MRD
Overview of our first experiment

Bone Marrow Sample → RBC Lysis Ficoll CD34+ Enrichment → Single Cell Capture 3’ and 5’ → Sequence
A note on feasibility

Sample collected 2:50pm Thurs.
Sample capture 5:35pm Thurs.
Library preparation starts 10:15am Fri.
Library preparation complete 6:35pm Fri.
Sequencing begins 8:04am Mon.
Sequencing ends 9:25am Tue.
Demux ends 4:40pm Tue.
Cellranger ends 7:45pm Tue.
MRD ENUMERATION
droplet partitioning single-cell RNA sequencing

cell capture

Microfluidic Device

Shruti Bhise MS  Sami Kanaan PhD
Higher than expected numbers of cells expressing CD34

RBC Lysis
Ficoll
CD34+ Enrichment
3’ Data only
Higher than expected numbers of cells expressing CD34

CD34 Enriched

Gradient Centrifugation

RBC Lysis

3’ Data only

About 10%
Co-embedding patient cells with healthy atlas

RBC Lysis
Ficoll
CD34+ Enrichment

3’ Data only
Seurat
RPCA Integration

GSE139369
(Atlas of Healthy Marrow, Cord, PBMC)
Granja et al.
Nature Biotech 2019

HSC
Early Erythroid
Late Erythroid
Myeloid Progenitor
Lymphoid Progenitor
pDC
cDC
CD14+ Monocyte
CD16+ Monocyte
Other
Pre B
B
Plasma
T
NK
Patient Sample

B cells
Monos, DCs, and progenitors

T/NK cells
RBCs
Annotation of cell types

RBC Lysis
Ficoll
CD34+ Enrichment

3’ Data only
UMAP Transform

Atlas (GSE139369)

- HSC
- Early Erythroid
- Late Erythroid
- Myeloid Progenitor
- Lymphoid Progenitor
- pDC
- cDC
- CD14+ Monocyte
- CD16+ Monocyte
- Other
- Pre B
- B
- T
- NK
- Atlas (GSE139369)

RBC Lysis
Ficoll
CD34+ Enrichment

3’ Data only
UMAP Transform

Atlas (GSE139369)

UMAP transform

B cells
Monos, DCs, and progenitors
T/NK cells
RBCs
Celltype classification needs to be quick (and accurate)
Motivation for a new cell classifier

- Wish list:
  - R interactive session
  - One line of code
  - Fast
  - Modular
  - UMI count-based (not have to embed)
Viewmaster - Softmax Regression

(Most Variable Genes)
~2500

Reference Dataset

Input features
Bias units

Net Input function

One-hot targets

argmax

Celltype


Softmax Regression

Cross-Entropy

One-Hot True Labels
Benchmarking Viewmaster

Reference Dataset 80% Train + Viewmaster Softmax Regression → Model → Query Dataset

Reference Dataset 20% Test

Predicted Celltype

‘Actual’ Celltype
Granja et al. Nature Biotech 2019
Viewmaster

GSE139369 Reference Dataset
Seurat ‘bmcite’ Query Dataset

Accuracy on training data: 98.98
Accuracy on testing data: 98.60

<table>
<thead>
<tr>
<th>user</th>
<th>system</th>
<th>elapsed</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.432</td>
<td>3.325</td>
<td>29.724</td>
</tr>
</tbody>
</table>

'Actual'

Celltype

Predicted Celltype
Rare cell types may confound softmax regression

GSE139369

Seurat ‘bmcite’

Harmonized Labels
- HSC
- Early Erythroid
- Late Erythroid
- Myeloid Progenitor
- Lymphoid Progenitor
- pDC
- cDC
- CD14+ Monocyte
- CD16+ Monocyte
- Other
- Pre B
- B
- Plasma
- T
- NK

Granja et al. Nature Biotech 2019
Viewmaster

GSE139369  Reference Dataset
Seurat ‘bmcite’  Query Dataset

Accuracy on training data: 98.98
Accuracy on testing data: 98.60

user  system elapsed
7.432  3.325  29.724

Actual Celltype

GSE139369  Query Dataset
Seurat ‘bmcite’  Reference Dataset

Accuracy on training data: 98.99
Accuracy on testing data: 97.49

user  system elapsed
8.405  7.795  35.641

Predicted Celltype

Olivia Waltner BA
Can we leverage natural genetic variation to improve leukemia detection?
Can we leverage natural genetic variation to improve leukemia detection?

Imperative to pair this with cell type
Souporcell genetically demultiplexes scRNAseq data

Souporcell identifies two genotypes in our sample.
Souporcell struggles with RBCs
Souporcell struggles with RBCs
Convincing evidence for relapse using souporcell

Atlas (GSE139369)

- HSC
- Early Erythroid
- Late Erythroid
- Myeloid Progenitor
- Lymphoid Progenitor
- pDC
- cDC
- CD14+ Monocyte
- CD16+ Monocyte
- Other
- Pre B
- B
- Plasma
- T
- NK
- Atlas (GSE139369)
Convincing evidence for relapse using souporcell
Can we detect AML mutations in scRNAseq data?
Can we detect AML mutations in scRNAseq data?

RNA molecule

EXON 1

G

AAA

EXON 2

polyA

CLINICALLY RELEVANT SNV

cb_sniffer

Petti et al. Nature Communications 2019
Can we detect AML mutations in scRNAseq data?

Petti et. al. Nature Communications 2019

cb_sniffer

TP53
- NotFound
- R248Q
- WT

Genotype 1
Genotype 2
Not sufficient coverage for the study of clonal heterogeneity
Can we augment coverage?

CD34+ Enrichment
Unfragmented

3' Data only

Hybridization capture with 1253 cancer gene panel

PCR, concatenation, SMRT sequencing
We can dramatically increase coverage using IsoSeq

CD34+ Enrichment
Unfragmented

3' Data only

IDH1 R132 mutation

IDH1 polyA
How does this extend to other patients?

- 2 yo with undifferentiated leukemia, with suspected relapse after unrelated cord blood transplant.
Can we leverage cell surface immunophenotype?

**Results**

- **Clinical Flow Cytometry**
  - 10-20 Cell surface antibodies

- **CITE Seq**
  - 150 Cell surface antibodies
Can we leverage cell surface immunophenotype?

Deliverable: Improve MFC!

10-20 Cell surface antibodies

150 Cell surface antibodies
Where are we now?

- Two patients with PCR+/MFC negative
- One cell of early myeloid lineage and recipient origin
Summary

• Inconsistencies in clinical assays are important motivators to improve diagnostics
• Integration of single expression data with genetic demultiplexing can provide a confident assessment of burden of relapsed leukemia
• Promising preliminary data suggesting that we can augment coverage of specific loci
• Immunoproteomic data show promise in recapitulating clinical flow cytometric data.
Future directions

• Increase sample numbers
• Working to detect fusion transcripts using PacBio sequencing
• Mechanisms of relapse
  • HLA expression / Antigen expression
  • T cell exhaustion
  • Myeloid suppressor cells
• Resources...
Data Scientist Positions at SCRI

- The Ben Towne Center for Childhood Cancer Research (BTCCCR) at SCRI is looking for Computational Biologists/Data Scientists

- Positions are available for Lead Data Scientist and Data Scientist I positions

- Experience with omics or other high dimensional data analysis and pipelines, e.g. next-generation sequencing, proteomics, metabolomics, epigenomics, phenotypic readouts, imaging, and clinical data, including experience developing and using statistical models and algorithms is required

- Work together with pediatric cancer researchers to improve the lives of children with cancer!

https://careers-seattlechildrens.icims.com/

Email: Sean.Taylor@seattlechildrens.org or jay.sarthy@seattlechildrens.org for more information
Acknowledgements

Sami Kanaan
Shruti Bhise
Olivia Waltner
Rula Green-Gladden

Jeffrey Stevens
Todd Cooper
Melinda Biernacki
Marie Bleakley
Monica Thakar

Cole Trapnell
Jason Underwood
Vijay Ramani

Soheil Meshinchi
Rhonda Reis
Jenny Smith

OUR PATIENTS AND THEIR FAMILIES
Questions?