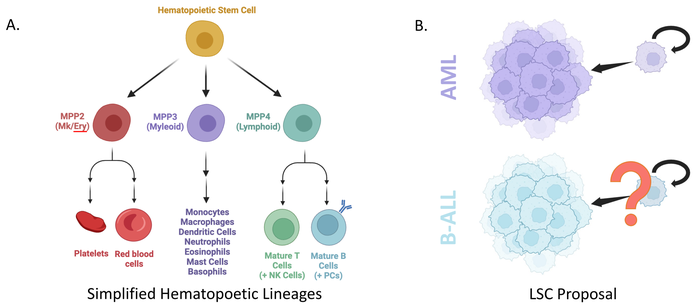
Identification of putative leukemic stem cells in B-Lymphoblastic Leukemia at a single cell resolution

DAVID MASTRO[1,2,3], Jennifer J. Chia[4,5,6], Dinesh S. Rao[4,5,6], Alexander Hoffmann[1,2,5]

[1]BIG Summer Program, Institute for Quantitative and Computational Biosciences, University of California, Los Angeles, CA
[2]Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, CA
[3]Department of Computer Science, University of California, Los Angeles, CA
[4]Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA
[5]Jonsson Comprehensive Cancer Center, University of California, Los Angeles, CA
[6]Broad Stem Cell Research Center, University of California, Los Angeles, CA

BACKGROUND & SIGNIFICANCE

- B-Lymphoblastic Leukemia (B-ALL) is a cancer in which malignant B-lymphoblasts clonally expand in the bone marrow, with relapse after initial treatment leading to high morbidity.
- In AML (Acute Myeloid Leukemia), a myeloid lineage leukemia, relapse can be attributed to leukemic stem cells (LSC), rare quiescent cells (~0.1%) that can replenish the leukemic clone.
- Whether LSCs contribute to relapse in B-ALL (~15% of pediatric cases) remains unknown.



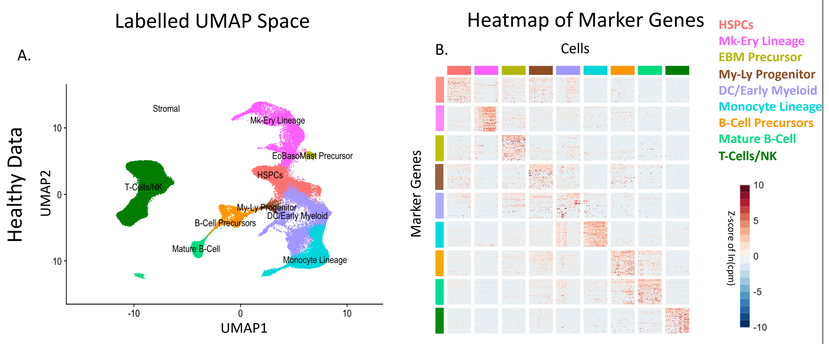
Simplified Hematopoetic Lineages | LSC Proposal

- LSCs have been identified in the literature by several transcriptomic signatures:
  - LSC17 gene set[4] – identified through a differential analysis of LSC enriched subpopulations against LSC depleted subpopulations
  - "ROS" gene set[5] – identified through a differential analysis of subpopulations with a high reactive oxygen species (ROS) vs low ROS, as LSCs have low ROS
  - OxPhos gene sets[2] - LSCs have low glycolytic activity and high oxidative phosphorylation
  - Quiescence gene sets – identifying quiescence can highlight these "dormant" LSCs
  - NF-κB induced gene sets – our lab recently found a correlation between high quiescence and NF-κB activity in HSCs across the age-span

RESEARCH QUESTIONS

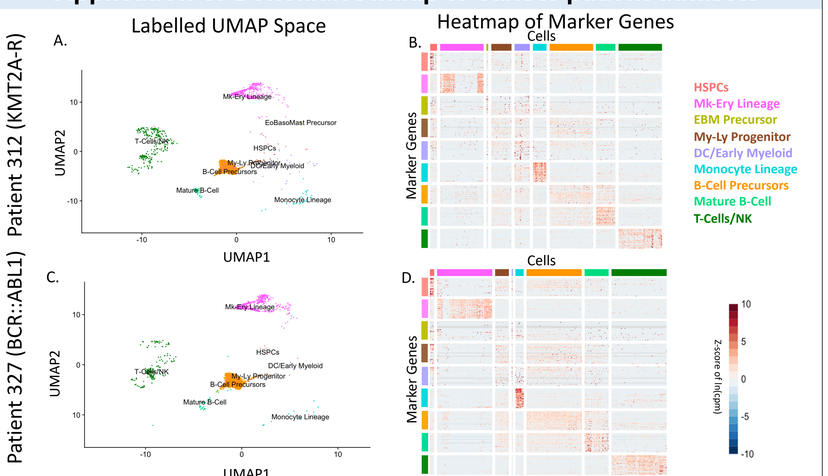
Can we identify LSCs in B-ALL leveraging single cell RNA sequencing (scRNA-seq) patient data?

BoneMarrowMap labelling shows high resolution of cell lineage



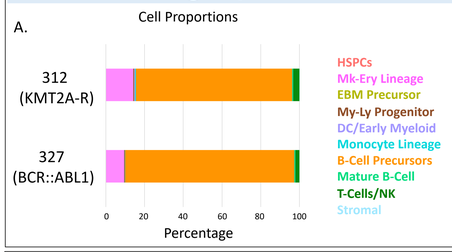
A. BoneMarrowMap (package) labelling uses K-means clustering against a reference healthy bone marrow to identify specific cell types. B. Cell type marker genes show the specificity of the labelling. Healthy data from Bone Marrow Atlas[6], a collection of 175 healthy scRNA-seq bone marrow samples.

Application of BoneMarrowMap to cancer patient datasets



B-ALL is subclassified by chromosome alterations, with KMT2A-rearranged (KMT2A-R, n=6) and BCR::ABL1 (n = 10) having poor prognoses. A-B. UMAP of representative datasets. C-D. Heatmap of marker genes.

Cell labeling captures expected expansion of B-Cell Precursors



A. Cell type labelling highlights the expansion of B-cell precursors at the expense of other hematopoietic cells, consistent with the diagnosis.

Healthy adults should have <5% of their bone marrow by B-cell precursors.

Ranked-based scoring allows for gene-set-enrichment analysis

A. Raw counts

	Cell 1	Cell 2	Cell 3
Gene A	1	3	5
Gene B	**3**	**6**	**3**
Gene C	2	4	6
Gene D	**5**	**10**	**1**

Rank Genes →

Ranked counts

	Cell 1	Cell 2	Cell 3
Gene A	4	4	2
Gene B	**2**	**2**	**3**
Gene C	3	3	1
Gene D	**1**	**1**	**4**

Average ranks of gene set → (B+D/2)

Scores

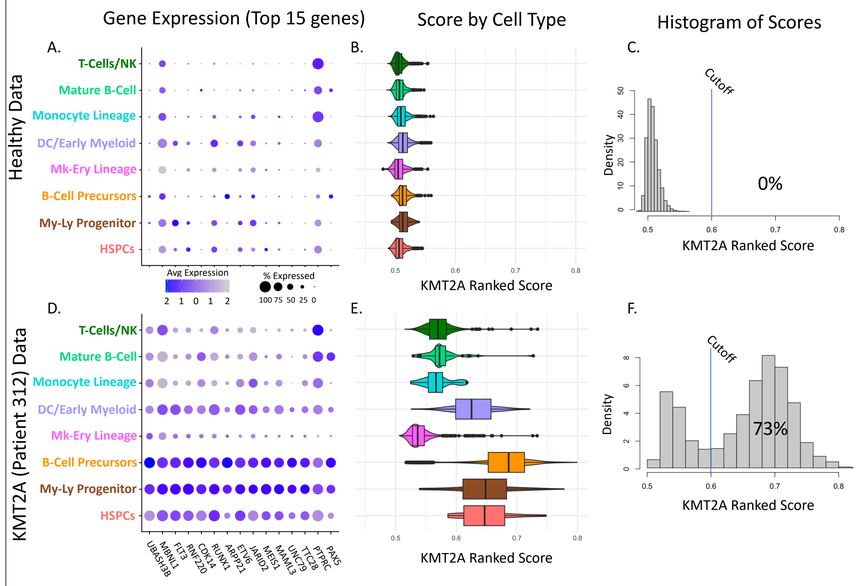
	Cell 1	Cell 2	Cell 3
Rank score (B and D)	$\frac{3.5}{\beta_1}$	$\frac{3.5}{\beta_2}$	$\frac{1.5}{\beta_3}$

Normalization factor (β_i) = max-min possible rank

Ranked-based scoring metric that provides gene set enrichment quantification at single cell resolution.[3]

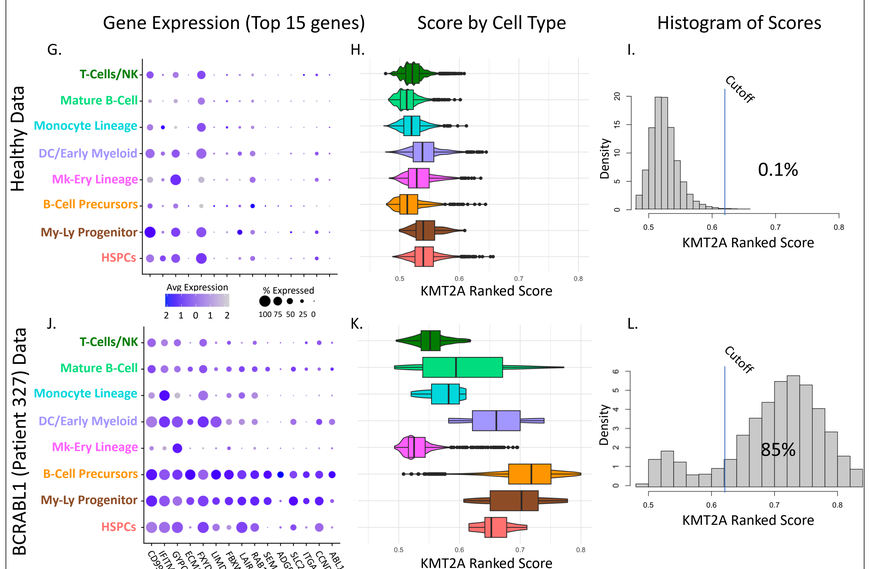
Leukemic cells can be identified by known transcriptomic signatures

KMT2A-R Signature



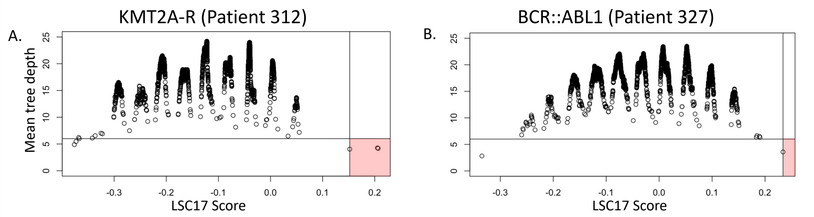
Established KMT2A-R gene signature[8] (253 genes) applied to healthy (A-C) and KMT2A-R (D-F) data.

BCR::ABL1 Signature



Established BCR::ABL1 signature[1] (42 genes) applied to healthy (G-I) and BCR::ABL1 (J-L) data.

LSC17 score and outlier detection together classify putative LSCs



C. KMT2A-R Patients

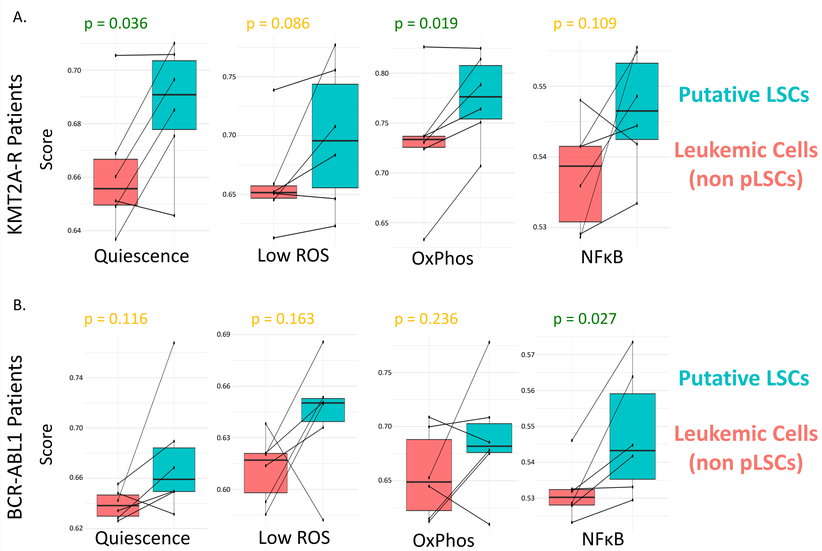
Patient	Total Cells	Leukemic (%)	Leukemic (#)	pLSCs
312	8714	73%	6363	3
361	5592	60%	3366	2
385	7362	69%	5048	3
359	19430	28%	5539	4
326	7878	84%	6601	1
366	8827	71%	6270	3

D. BCR::ABL1 Patients

Patient	Total Cells	Leukemic (%)	Leukemic (#)	pLSCs
327	7784	85%	6591	1
360	12658	63%	7985	2
370	21129	48%	10154	3
348	9093	48%	4401	0
377	8493	56%	4780	1
347	12656	65%	8219	0
346	15976	60%	9562	1
328	8572	79%	6804	0
322	10672	88%	9437	1
374	9385	91%	8538	0

A-B. Isolation tree algorithm tests how many divisions it takes to isolate a point from the population on average. This outlier classification allows for a standard between the datasets for recognizing putative LSCs (pLSCs) from total leukemic cells based on the LSC17 expression distribution.
C-D. Leukemic cell and pLSC quantification across datasets.

Orthogonal gene signatures show expected pLSC phenotypes



A-B. For each dataset (two points connected by line), we scored the putative LSC (left point) and the remainder of the leukemic cells (right point) against additional LSC metrics. Each point is a mean of scores (i.e. the mean of putative LSC scores in patient X). Paired t-test performed.

CONCLUSIONS & DISCUSSION

- Here we leverage B-ALL scRNA-seq patient data to look for rare leukemic stem cells.
- We classified all sequenced cells as leukemic or non-leukemic using subtype-specific genes.
- Using LSC17 transcriptome score with outlier detection, we identify putative LSCs (pLSCs).
- These pLSCs show phenotypes consistent with AML LSCs. However, the decreased statistical significance in the BCR::ABL1 comparisons, as well as the lack of pLSC identification in some BCR::ABL1 datasets (no LSC17 outliers), suggests subtype-specific differences in B-ALL.
- While *in vivo* transplantation is required to definitively prove a cell acts as an LSC, these bioinformatic findings using many datasets may be useful to guide future research.

References

[1]Bhojwani et al (2007). *Pediatric Drugs*, 9(3), 149–156.
[2]Chen, et al (2022). *Disease Markers*, *2022*, 1–20.
[3]Foroutan et al (2018). *BMC Bioinformatics*, *19*(1).
[4]Jiang, & Rothe (2017). *Faculty Opinions – Post-Publication Peer Review of the Biomedical Literature*
[5]Lagadinou et al (2013). *Cell Stem Cell*, *12*(3), 329–341.
[6]Lee et al (2023). *Frontiers in Immunology*, 14
[7]*Scientific Image and Illustration Software*. BioRender. (n.d.)
[8]Zeng et al (2023). *Precise Single-Cell Transcriptomic Mapping of Normal and Leukemic Cell States Reveals Unconventional Lineage Priming in Acute Myeloid Leukemia*

Contacts

davidrm3@g.ucla.edu
jchia@mednet.ucla.edu