Immune cell development

https://doi.org/10.1038/s41590-024-01793-1

Mapping human hematopoiesis

Kathrin M. Bernt

Understanding normal hematopoiesis is critical to understanding disease. Technological advances are driving insight into human hematopoiesis at unprecedented resolution. Integrating '-omics' datasets with machine learning has yielded a high-resolution map of primary human bone marrow hematopoietic progenitor cells that supports the study of immune cell development, as well as the origins of disease.

Single-cell technologies are in the process of revolutionizing the understanding of biology and disease, revealing how cells are wired and communicate with each other. In this issue of *Nature Immunology*, Zhang et al.[1] apply a combination of CITE-Seq (cellular indexing of transcriptomes and epitopes by sequencing) and the Infinity Flow tool for flow cytometry[2], integrated through the use of a machine learning computational platform (scTriangulate)[3], to refine the understanding of hematopoietic differentiation. In the process, the authors provide insights and principles that go well beyond hematopoiesis.

CITE-Seq has gained popularity since it was first described[4]. This technology can overcome some of the drawbacks of both flow cytometry (limited number of antigens that can be probed on a single cell) and single-cell RNA sequencing (shallow sequencing may miss critical surface antigens with a long half-life, and misses regulation at the level or protein stability or modifications). The combination of cytometric data and transcriptomic data has been successful at resolving subpopulations missed by either technique alone. Curated and titrated antibody panels are commercially available and widely used. Investigators rely on these panels, as individual titration of antibodies is time-consuming and, given the single-cell sequencing readout, typically expensive. Such panels also have the advantage of facilitating cross-comparison between datasets from different labs. However, there are clear downsides to the use of large commercial panels. Irrelevant antibodies that are present in a panel add noise and a risk for false-positive results. At the same time, a given antibody concentration that performs well for one tissue may not do so for another with lower (but real) antigen expression. These drawbacks limit the full potential of commercial CITE-Seq panels. Zhang et al. devised an ingenious system of hashing pools of antibodies to titrate a large set of antibodies to their tissue of interest in an efficient and (relatively) cost-effective manner[1]. The resulting insights are powerful and broadly applicable.

Commercial titration for hematological and immunological panels is frequently performed on peripheral blood cells, a process that is risky due to the following two issues. First, many antigens that are critical for differentiating hematopoietic and immune cells are present in the peripheral blood in only minute quantities; therefore, non-specific binding to these antigens is not assessed optimally. Second, some

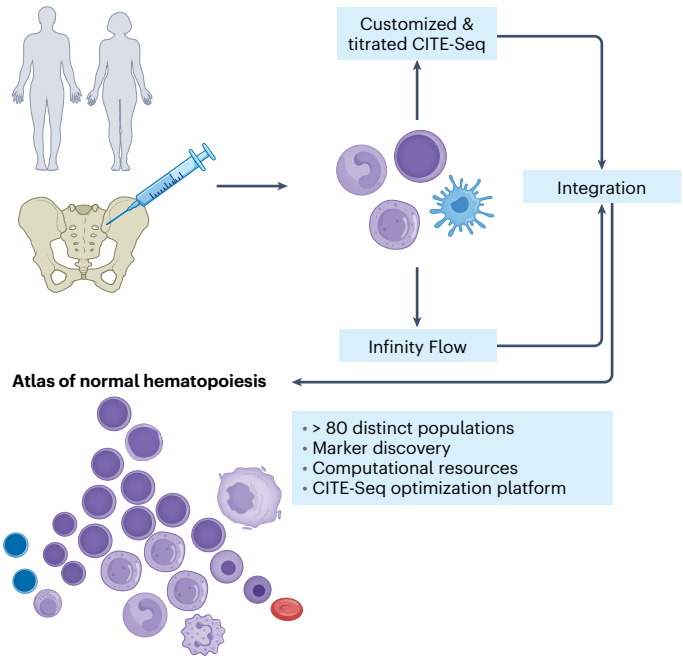

Fig. 1 | An atlas of normal human hematopoiesis. A combination of customized CITE-Seq and Infinity Flow software was used to generate a high-resolution atlas of human adult normal hematopoiesis. The multi-modal analysis identified over 80 distinct populations, including newly resolved populations and cell states.

antigens are expressed at different levels during specific stages of differentiation; therefore, titration that works for one developmental stage might fail for another. Zhang et al. evaluated and optimized the concentration of several commonly used antibodies directed against important antigens that define lineage, developmental state or activation state or serve as drug targets[1]. Of 266 antibodies present in the popular BioLegend TotalSeq-A antibody cocktail (titrated on peripheral blood), 126 exhibited non-dynamic, non-specific or no staining of bone marrow cells. Only 64 of 266 antibodies titrated on peripheral blood cells performed well when applied to the entire hematopoietic hierarchy and were retained at the same concentration as in the original panel; the remainder required optimization for best performance. Example antibodies that required titration include those directed against the frequently relied upon stem and progenitor cell markers CD34, CD38, CD90, c-Kit and CD123, the important lymphoid antigens CD2, CD4 and CD127, or the immunotherapy target PD-L1, to name just a few. It is likely that the underlying principle – customization and titration of CITE-Seq panels to the tissue of study for best performance – applies to nearly all biological contexts addressed with CITE-Seq. In addition to providing an optimized and well-functioning panel for hematopoiesis and immunological research, this Resource provides a blueprint for optimizing a CITE-Seq panel to a specific biological context in a thorough and cost-effective manner[1].

Another surprising insight is the degree of interpersonal variability in protein expression for several antigens that the field has treated as reliable markers of lineage. Only 40 of the 132 markers performed consistently in all donors. Again, many antigens that the field has routinely relied on for the identification of lineage, differentiation and activation states are affected. Some of their variability is fairly minor, whereas some, such as CD45RA, exhibit substantial inter-donor variability. Given the relatively low number of samples in this study, inter-individual variability is probably underestimated. It is also likely that it applies to other biological contexts. Ultimately, this insight also highlights the power of technologies such as CITE-Seq and Infinity Flow for the reliable annotation of cell populations and cell states as additional markers and multi-modal annotation increase accuracy and reproducibility.

Zhang et al. deliver the best-resolved surface-marker-defined atlas of adult human hematopoiesis thus far, composed of over 80 distinct and molecularly defined cell populations[1] (Fig. 1). This work expands substantially on prior published single-cell maps of human hematopoiesis[5–7]. The carefully optimized CITE-Seq approach combined with multi-parameter Infinity Flow resolved several new populations and cell states, particularly in neutrophil and erythroid development. For example, a combination of CD326 and CD235a was able to resolve defined stages of erythroid development more accurately than the current commonly used antigens. This approach also nominated several markers, such as CD133 as a reliable marker for megakaryocyte-erythroid progenitors (MEPs), or C5L2 and TSPAN33 as markers for granulocytic commitment. Careful validation backs up these identified populations and markers. Conceptually, the identification of discrete cell populations by multiple complementary markers and methods supports a model wherein hematopoietic differentiation proceeds through a series of discrete meta-stable states, rather than being a continuum that is broken up artificially into populations on the basis of the expression of a handful of markers[8]. This detailed map of bone marrow populations sets a new reference standard for human hematopoiesis.

Zhang et al. also include an example of the type of application to hematopoietic malignancies that the refined CITE-Seq panel and detailed normal hematopoietic map support[1]. Although stem-cell-like transcriptional states in primitive leukemia stem cells (p-LSCs) have long been recognized as being associated with resistance to chemotherapy and poor outcomes, a second, more differentiated monocytic leukemia stem cell (m-LSC) population that can mediate resistance to venetoclax has been described[9]. However, precise separation of the two LSC populations by surface markers has been challenging. Projection of single-cell profiles from p-LSC and m-LSC populations from that study[9] onto the refined hematopoietic map resolved additional subpopulations within the two LSC populations, with some overlap between the two. Comparison of p-LSCs and m-LSCs nominated additional surface markers for distinguishing the two populations that were validated in three samples. The improved ability to identify surface markers that correlate with biologically relevant, transcriptionally defined cell states could be useful for tracking LSC populations in a clinical setting.

The atlas has two important limitations. Although the authors have taken care to include participants of diverse sex and racial/ethnic origin, the number of participants is relatively small. Expanding the atlas to additional participants while ensuring equitable representation and diversity will be critical. In addition, the atlas focuses on adults only, yet the hematopoietic and immune systems are highly plastic and undergo major changes during embryonic and fetal development, infancy and childhood[10,11], as well as in old age[12]. Although it does not address age-related changes, this carefully annotated and validated atlas will inform ongoing efforts to map developmental changes to the hematopoietic and immune systems across the lifespan.

In summary, Zhang et al. have provided the most detailed, molecularly defined transcriptomic atlas of adult bone marrow progenitors thus far[1]. The normal reference atlas, refined and optimized antibody panels and computational pipelines constitute very useful tools for future research on hematopoietic development and disease states. More importantly, the principles outlined here are likely to be applicable to almost any biological scenario, and set a new standard for any researcher seeking to deploy CITE-Seq as a tool.

Kathrin M. Bernt [1,2]
[1]Division of Pediatric Oncology, Children's Hospital of Philadelphia, Philadelphia, PA, USA. [2]Department of Pediatrics, Perelman School of Medicine at the University of Pennsylvania and Abramson Cancer Center, Philadelphia, PA, USA.
e-mail: berntk@chop.edu

Published online: 21 March 2024

References

1. Zhang, X. *Nat. Immunol.* https://doi.org/10.1038/10.1038/s41590-024-01782-4 (2024).
2. Becht, E. et al. *Sci. Adv.* **7**, eabg0505 (2021).
3. Li, G. et al. *Nat. Commun.* **14**, 406 (2023).
4. Stoeckius, M. et al. *Nat. Methods* **14**, 865–868 (2017).
5. Oetjen, K. A. et al. *JCI Insight* **3**, e124928 (2018).
6. Pellin, D. et al. *Nat. Commun.* **10**, 2395 (2019).
7. Ranzoni, A. M. et al. *Cell Stem Cell* **28**, 472–487 e477 (2021).
8. Velten, L. et al. *Nat. Cell Biol.* **19**, 271–281 (2017).
9. Pei, S. et al. *Cancer Discov.* **13**, 2032–2049 (2023).
10. Jardine, L. et al. *Nature* **598**, 327–331 (2021).
11. Chen, C. et al. *Blood* **139**, 2198–2211 (2022).
12. Ainciburu, M. et al. *eLife* **12**, e79363 (2023).

Competing interests

K.M.B. declares that they have consulted for Agios and Novartis and have received research funding from Syndax Pharmaceuticals. These contributions are unrelated to the subject material discussed in this article.