

Hematopoietic lineage diversification, simplified.

Roy Drissen and Claus Nerlov.

MRC Molecular Hematology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DS, United Kingdom

Hematopoiesis is a complex process that requires a high degree of transcriptional diversification during lineage commitment and differentiation. de Graaf et al. (2016) have now generated a comprehensive gene expression data set that allows cell type specific genes as well as associated transcription factor expression patterns to be readily identified.

Hematopoiesis involves the generation of over 10 distinct cellular lineages from a small pool of multi-potent hematopoietic stem cells (HSCs) (Orkin and Zon, 2008). The process of lineage specification from multi-potent HSCs occurs via an intricate series of increasingly lineage-restricted progenitor cells, culminating in the formation of fully committed uni-lineage progenitors, which then undergo terminal differentiation. The mature, differentiated hematopoietic cell types are transcriptionally highly divergent, reflecting their very distinct physiological functions, ranging from oxygen transport and hemostasis to bacterial phagocytosis and antibody production. Understanding how the hematopoietic lineages are formed, and how their individual differentiation programs are executed, therefore requires detailed knowledge of the transcriptional programming, not only of the mature cell types themselves, but also of the large number of intermediate multi-, bi-, and mono-potent progenitors involved in their generation.

The primary experimental organism for hematopoiesis research is the mouse, and numerous murine studies have described important aspects of hematopoietic lineage specification, including the identification and gene expression profiling of key progenitor intermediates. However, such studies have generally focused on the specific aspect of hematopoiesis under investigation, and currently available large datasets (such as ImmGen, HemaExplorer and Gene Expression Commons) often omit some important cell types, such as eosinophils, basophils, mast cells and dendritic cells, that are rare or more difficult to isolate and purify. We therefore currently lack a comprehensive data set that incorporates all the hematopoietic lineages, as well as the full spectrum of intermediate progenitors.

A recent report by de Graaf et al. in Stem Cell Reports (de Graaf et al., 2016) provides an important step in this direction by using the same platform to profile murine lymphoid (B-, T-, and NK cells), myeloid (neutrophils, monocytes/macrophages, basophils, mast cells and eosinophils), erythroid, megakaryocytic and dendritic cells, as well as the stem- and progenitor cells that generate them. This allowed the identification of genes that are highly specific for each of these lineages, providing both important functional information and valuable tools for the gene expression-based identification and analysis of individual cell types. The authors demonstrate this by generating a transgenic fluorescent reporter driven by the promoter of *Mkx*, a gene identified as selectively expressed in eosinophils, and showing that it selectively labels a subset of peritoneal eosinophils, but no other cell type, potentially identifying previously unrecognized

heterogeneity of differentiated eosinophils. Further exploration of the dataset should allow the systematic exploration of such heterogeneity of other cell types.

Comparison to gene expression profiles of human hematopoietic cell types and progenitor cells (Novershtern et al., 2011) showed that a high proportion of the lineage-specific genes identified in the mouse showed similar lineage-specificity in the human system. While this underscores the usefulness of the mouse as a model for human hematopoiesis and its associated disorders notable differences were also observed, and in particular the granulocyte/macrophage lineages did not show a high degree of specificity in cross-species correspondence. The availability of comparable murine and human datasets should be particularly useful when comparing genetic phenotypes across these species, and when generating models of human genetic disorders.

The authors also compare the gene profiles of multi-potent and lineage-restricted progenitor populations using minimum spanning tree analysis. This analysis confirms that terminal differentiation is accompanied by increasing transcriptional distances between the cellular lineages, and also identifies a number of transcription factors and surface markers with high expression variance across the data set. Notably, the identified cellular relationships could be largely replicated using either of these subsets of differentially expressed genes supporting the relevance of the identified transcription factors in lineage specification. All cellular profiles obtained have been added to the Haemosphere portal, and using the associated analysis platform those transcription factors that differ between individual lineages or differentiation stages can be readily identified and tested as candidate regulators of both hematopoietic lineage diversification and terminal differentiation.

However, while progenitors of the megakaryocytic/erythroid and lymphoid lineages readily organized into branching differentiation hierarchies, the observed relationships between granulocyte/macrophage progenitors and the mature myeloid lineages that derive from them were less clear. One possible explanation is that the current FACS-based definitions of myeloid progenitors do not identify homogeneous cell populations, and instead an underlying heterogeneity exists. This is highlighted by recent studies using single cell transcriptomics to identify myeloid progenitor subpopulations (Drissen et al., 2016; Paul et al., 2015) that indicate that progenitors for eosinophils and basophils/mast cells may be distinct from those of neutrophils and monocytes/macrophages. It is therefore likely that accurate modeling of the relationships of granulocyte/macrophages and their upstream progenitors will require further progenitor sub-fractionation. The profiling data also lend support to the notion that megakaryocyte development may proceed, at least in part, independently of a bi-potent megakaryocytic/erythroid progenitor (MEP) as committed megakaryocyte progenitors were observed to be more closely related to HSCs than to MEPs. This is in line with recent reports of platelet-biased HSCs (Sanjuan-Pla et al., 2013), of megakaryocyte-restricted repopulating progenitors (Yamamoto et al., 2013), and of the activation of megakaryocyte-primed HSC-like progenitors during inflammation (Haas et al., 2015).

It therefore seems clear that the comprehensive data set generated by de Graaf et al. can give rise to both cellular and molecular hypotheses for subsequent experimental testing. In particular, the data highlight the need to better define the cellular pathways of

granulocyte/macrophage specification in both mice and humans, and the future importance of integrating population- and single cell gene profiling data. Finally, the availability of parallel, comprehensive profiling of both the human and murine hematopoietic system on a single analytical platform will be a valuable tool for translational hematology research.

References:

- de Graaf, C.A., Choi, J., Baldwin, T.M., Bolden, J.E., Fairfax, K., Robinson, A.J., Biben, C., Morgan, C., Ramsay, K., Ng, A.P., *et al.* (2016). Haemopedia: a gene expression atlas of murine haematopoietic cells. *Stem Cell Reports NA*, NA.
- Drissen, R., Buza-Vidas, N., Woll, P., Thongjuea, S., Gambardella, A., Giustacchini, A., Mancini, E., Zriwil, A., Lutteropp, M., Grover, A., *et al.* (2016). Distinct myeloid progenitor-differentiation pathways identified through single-cell RNA sequencing. *Nat Immunol* 17, 666-676.
- Haas, S., Hansson, J., Klimmeck, D., Loeffler, D., Velten, L., Uckelmann, H., Wurzer, S., Prendergast, A.M., Schnell, A., Hexel, K., *et al.* (2015). Inflammation-Induced Emergency Megakaryopoiesis Driven by Hematopoietic Stem Cell-like Megakaryocyte Progenitors. *Cell Stem Cell* 17, 422-434.
- Novershtern, N., Subramanian, A., Lawton, L.N., Mak, R.H., Haining, W.N., McConkey, M.E., Habib, N., Yosef, N., Chang, C.Y., Shay, T., *et al.* (2011). Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* 144, 296-309.
- Orkin, S.H., and Zon, L.I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* 132, 631-644.
- Paul, F., Arkin, Y., Giladi, A., Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Winter, D., Lara-Astiaso, D., Gury, M., Weiner, A., *et al.* (2015). Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell* 163, 1663-1677.
- Sanjuan-Pla, A., Macaulay, I.C., Jensen, C.T., Woll, P.S., Luis, T.C., Mead, A., Moore, S., Carella, C., Matsuoka, S., Jones, T.B., *et al.* (2013). Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature* 502, 232-236.
- Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K.L., Ema, H., and Nakauchi, H. (2013). Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell* 154, 1112-1126.