



### SINGLE-CELL TECHNOLOGY MEETS HEMATOLOGY

# New insights into hematopoietic differentiation landscapes from single-cell RNA sequencing

Sam Watcham, Iwo Kucinski, and Berthold Gottgens

Department of Haematology and Wellcome–Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Cambridge, United Kingdom

**Single-cell transcriptomics has recently emerged as a powerful tool to analyze cellular heterogeneity, discover new cell types, and infer putative differentiation routes. The technique has been rapidly embraced by the hematopoiesis research community, and like other technologies before, single-cell molecular profiling is widely expected to make important contributions to our understanding of the hematopoietic hierarchy. Much of this new interpretation relies on inference of the transcriptomic landscape as a representation of existing cellular states and associated transitions among them. Here we review how**

**this model allows, under certain assumptions, charting of time-resolved differentiation trajectories with unparalleled resolution and how the landscape of multipotent cells may be rather devoid of discrete structures, challenging our preconceptions about stem and progenitor cell types and their organization. Finally, we highlight how promising technological advances may convert static differentiation landscapes into a dynamic cell flux model and thus provide a more holistic understanding of normal hematopoiesis and blood disorders. (*Blood*. 2019;133(13):1415-1426)**

## Introduction

Hematopoiesis research spanning >150 years has been significantly driven by technological breakthroughs. Microscopy-based observations in the 19th century established that blood is composed of 2 bone marrow–derived cell lineages: myeloid and lymphoid, perhaps sharing a common stem cell origin.<sup>1</sup> It was not until the 1950s when bone marrow transplantation rescue of lethally irradiated mice<sup>2–4</sup> confirmed this hypothesis. Subsequently, *in vitro* hematopoietic colony assays provided functional evidence for intermediate stages between hematopoietic stem cells (HSCs) and terminally differentiated cells,<sup>5</sup> ranging from multipotent (MPP) to unipotent progenitor cells. These findings arose from the shadow cast by the destructive effects of radiation on the blood system after the first use of nuclear weapons in the 1940s,<sup>6</sup> with the first successful human bone marrow transplantation reported in 1959.<sup>7,8</sup> This approach remains the only curative therapy for a number of hematopoietic malignancies to this date.<sup>9</sup> Although these practical applications were developed early on, our biological understanding of hematopoiesis lagged behind until isolation of specific cell populations became possible.

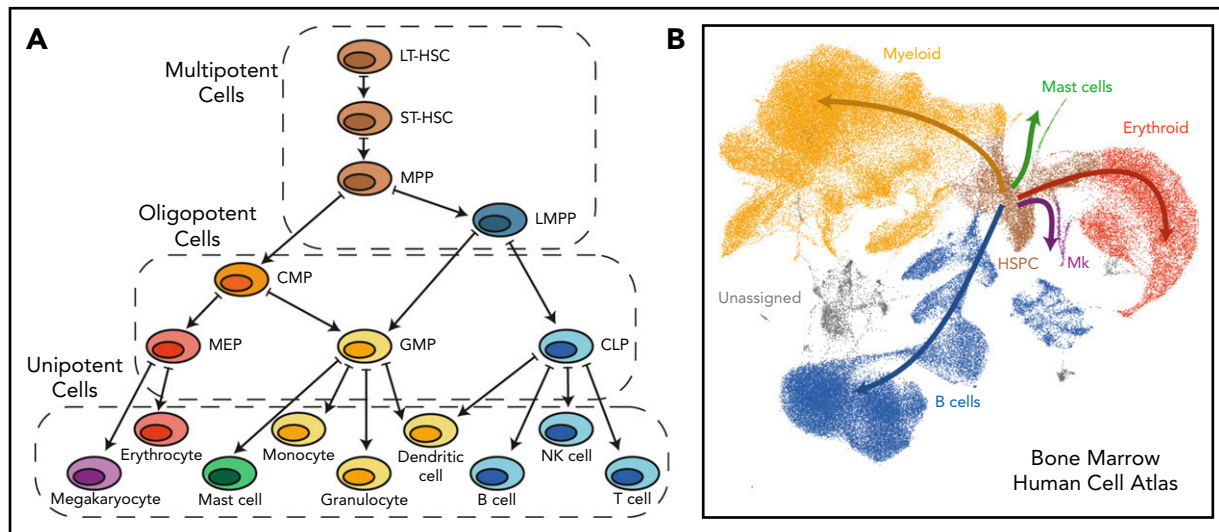
A critical advance came from the related field of immunology, allowing the sorting of individual cells<sup>10</sup> and generation of monoclonal antibodies to detect surface markers.<sup>11</sup> At this stage, a key achievement of the hematopoietic community had begun to take form, with the establishment of the differentiation tree. By the end of the 20th century, the hematopoietic tree was rooted in long-term HSCs (LT-HSCs), followed by short-term HSCs (ST-HSCs) and MPPs, partitioned according to their ability to repopulate blood in transplantation assays over diminishing

periods of time.<sup>12–16</sup> These cells were proposed to differentiate through a set of bifurcations that produced distinct progenitor cell populations with decreasing lineage potential and self-renewal activity (Figure 1A). In the past 2 decades, this model has been subjected to constant extensions and refinements, largely because of new evidence highlighting cellular heterogeneity obtained from single-cell assays. At the same time, cell barcoding approaches have mediated clonal tracking of native hematopoiesis<sup>17–19</sup> and stressed the importance of gaining insight into the unperturbed tissue state. The resulting evolution of the hematopoietic tree has been discussed in detail elsewhere.<sup>6,15,20,21</sup>

We are currently witnessing another single-cell revolution, in which vast transcriptomic data sets are transforming our understanding of hematopoiesis. As a result, the idea of cellular transitions between discrete progenitor states as they differentiate has become difficult to accommodate.<sup>20</sup> Instead, multiple studies have proposed the idea of continuous differentiation landscapes, with little or no discrete differentiation stages and smooth transitions across the cell states. In this context, cells within a heterogeneous pool of HSPCs differentiate along a multitude of potential trajectories that contain poorly defined branch points, which determine the fate of a particular cell. In this review, we aim to highlight recent biological insights gained into the nature of these landscapes using single-cell RNA sequencing (scRNA-seq) and downstream computational tools.

## scRNA-seq: opportunities and limitations

Although single-cell quantification of gene expression for small numbers of genes was achieved in the early 1990s,<sup>22</sup> it is because



**Figure 1. Comparison of a hematopoietic tree diagram with a single-cell transcriptomic landscape.** (A) Schematic showing one of the classic views of the hematopoietic cell hierarchy. Dashed boxes show 3 compartments encompassing cells of different potency: multipotent cells on top, bipotent/oligopotent cells in the middle, and terminally differentiated (unipotent) cells at the bottom. (B) A dimensionality reduction projection (UMAP algorithm) of single-cell transcriptomes from the bone marrow mononuclear cell fraction. Arrows indicate main directions of differentiation, inferred from analysis of typical marker genes. Gray indicates unassigned cells, in which identity based on markers is unclear (data set downloaded from Human Cell Atlas data portal and processed by I.K.). CMP, common myeloid progenitor; CLP, common lymphoid progenitor; GMP, granulocyte-monocyte progenitor; HSPC, hematopoietic stem and progenitor cell; LMPP, lymphoid-primed MPP; MEP, megakaryocyte-erythroid progenitor; Mk, megakaryocyte.

of breakthroughs in parallelization in the past few years that single-cell transcriptomics is now chasing its conceptual predecessors flow and mass cytometries in terms of throughput.<sup>23</sup> However, unlike flow or mass cytometry measurements, which are typically restricted to at most a few dozen predefined markers, scRNA-seq can measure expression of up to  $10^4$  genes simultaneously in each cell, thus offering unprecedented detail for the definition of cellular states. Two key parameters of any scRNA-seq experiment are the number of cells assayed, which determines the probability of being able to capture rare cell populations, and detection sensitivity, which dictates the number of genes available for cell-to-cell comparisons. These 2 criteria are leveraged into 2 classes of scRNA-seq platforms. High-throughput methods (eg, droplet-based methods like Drop-Seq, InDrops, 10X genomics)<sup>24-26</sup> provide in excess of  $10^4$  cells per experiment but detect only 1000 to 3000 expressed genes per cell. In contrast, low-throughput methods (eg, plate-based methods like Smart-Seq2, CelSeq2, mcSCR-Seq, RamDA-Seq)<sup>27-30</sup> are typically used to profile up to  $\sim 10^3$  cells but detect  $>5000$  genes per cell. The details of specific methods and their technical considerations have been extensively reviewed.<sup>23,31,32</sup>

For organ systems with a fast turnover, such as hematopoiesis, scRNA-seq offers a snapshot of cells and their expression states at a particular time point. Although temporal information is missing, much excitement has been generated by the idea that the distribution of single-cell expression profiles in the high-dimensional space can be considered as a transcriptional landscape, which encodes information on cellular transitions, enabling ordering of cells along pseudotime, from immature progenitors toward the various hematopoietic lineages.<sup>33,34</sup> The inference of such putative differentiation trajectories requires certain assumptions and imposes some limitations, mainly differentiation is a continuous process (detection of jumps between cell states is difficult); cells differentiate asynchronously and are captured at multiple points along their differentiation routes;

cells differentiate using defined (nonrandom) trajectories; cells move in one direction, so additional knowledge to determine start and finish is required; and cell-state information is complete. The last point is clearly problematic, because we cannot see hidden variables like epigenetics or cell location with current scRNA-seq protocols. Furthermore, some analytic methods require stricter assumptions, such as a tree-like structure of the data, where cells undergo a series of bifurcations during their differentiation, or absence of oscillations between cell states (eg, cell cycle), which may disregard potentially relevant biological information. For a detailed discussion of these assumptions, we refer readers to Weinreb et al.<sup>35</sup>

The past 6 years have seen an explosion of scRNA-seq studies within the hematopoiesis field, starting from profiling of previously well-defined hematopoietic populations<sup>36-39</sup> and shifting toward less biased selection of cells,<sup>26,40-43</sup> finally arriving at whole-tissue studies surveying  $>10^5$  cells. We have summarized some of the most useful resources in Table 1. The current culmination comes from a large collaborative effort: the Human Cell Atlas,<sup>42</sup> which aims to create reference maps for cells across  $>50$  tissues in the human body. This includes a recently released data set with  $>500\,000$  human bone marrow and cord blood cells, providing the most complete scRNA-seq map yet of the human hematopoietic compartments.<sup>44</sup> We present a general view of this data set in Figure 1B. Of note, we have excluded from the analysis natural killer and T lymphocytes, which account for  $\sim 50\%$  of cells and appear disconnected from the HSPC cluster. This separation is expected for T cells, which differentiate in the thymus but may also indicate that natural killer progenitors are rare or even absent in the bone marrow mononuclear cell fraction. Because these large-scale data sets await their full exploitation, we will focus on the main insights into hematopoietic differentiation delivered by scRNA-seq so far.

## scRNA-seq reshapes the bottom and middle tiers of the hematopoietic tree

scRNA-seq has quickly proved to be a powerful tool for distinguishing discrete cell states. Studies concerning lymphoid lineages have been reviewed previously<sup>45</sup>; therefore, we will discuss a couple of examples from the myeloid lineage. Single-cell profiling allowed discovery of several DC subtypes, which respond differently to lipopolysaccharides.<sup>46</sup> A recent study laid out a new taxonomy for human DCs and monocytes in peripheral blood,<sup>47</sup> with a new DC population responsible for T-cell activation previously misclassified as plasmacytoid DCs and a new conventional DC progenitor population (CD100<sup>hi</sup>, CD34<sup>int</sup>), functionally distinct from the CD34<sup>hi</sup> HSCs. Yanez et al<sup>48</sup> showed that Ly6C<sup>hi</sup> monocytes can be divided into neutrophil-like and monocyte-derived DC precursors arising through alternative differentiation routes, in agreement with another report.<sup>49</sup>

Recent studies have also challenged our understanding of the oligopotent progenitor compartments. Profiling of the (lineage<sup>-</sup>, Sca1<sup>-</sup>, c-Kit<sup>+</sup>) population containing the CMP/granulocyte-monocyte progenitor/megakaryocyte-erythroid progenitor populations (Figure 1A) revealed an unanticipated heterogeneity, with at least 18 distinct subtypes,<sup>36</sup> with various degrees of lineage priming (ie, expression of lineage-specific genes), reflecting their progressive commitment. Importantly, the key surface markers FcgR and CD34, typically used to distinguish the CMP/granulocyte-monocyte progenitor/megakaryocyte-erythroid progenitor populations, turned out to be poor predictors of cell identity and lineage bias. Consistently, Olsson et al<sup>39</sup> showed that only a small proportion of cells sorted from conventional myeloid progenitor populations displayed gene expression profiles consistent with multilineage progenitor activity. This study also elegantly demonstrates how the wealth of information provided by scRNA-seq can be exploited to infer gene regulatory information, refine cell isolation strategies, and enrich for specific bipotent progenitors by focusing on the counteracting regulators Gfi1 and Irf8.

These findings are in line with the accumulating evidence from single-cell colony assays,<sup>36,39,50-53</sup> single-cell transplantations,<sup>36,54,55</sup> and cellular barcoding<sup>19,56</sup> experiments performed both in human and mouse, which collectively suggest that most cell fate decisions are taken earlier than expected from the classical hematopoietic tree model. This is also consistent with the original observations that many cells within progenitor gates (eg, CMPs or lymphoid-primed MPPs) are already lineage restricted.<sup>50,57</sup> Therefore, the emerging picture is that unipotent cells dominate the middle tier of hematopoietic progenitors, albeit with a small contribution of bipotent/oligopotent progenitors. Whether these rare cell populations are true functional intermediates or perhaps outlier cells spilling over from less mature compartments remains to be seen.

## Revisiting the HSC and MPP cell compartments by scRNA-seq

The upper tier of the hematopoietic tree classically contained 3 multipotent subpopulations with decreasing repopulation potentials: LT-HSCs, ST-HSCs (also known as MPP1), and MPPs, all 3 capable of generating both myeloid and lymphoid cells.<sup>12,54,58</sup> Before the advent of scRNA-seq, immunophenotyping showed

that MPPs have at least 4 subpopulations (MPP1-4) with distinct cell-cycle characteristics,<sup>12</sup> molecular features,<sup>37,59-62</sup> and functional biases.<sup>58,61,62</sup> Subsequent molecular analysis at the single-cell level, first by single-cell quantitative polymerase chain reaction<sup>62</sup> and later by scRNA-seq,<sup>19,37,63</sup> showed that the HSC/MPP compartment may be characterized by a continuous landscape. Furthermore, transcriptional priming toward erythroid/megakaryocytic (MPP2), myeloid (MPP3), or lymphoid lineages (MPP4) suggested that many key fate decisions have already been initiated by earlier upstream progenitors (MPP1). Rodriguez-Fraticelli et al<sup>19</sup> confirmed these findings *in vivo* by employing high-throughput scRNA-seq and lineage barcoding to quantify the degree of lineage priming and lineage contributions of each of the MPP subpopulations. Importantly, the compartment structure adapts to changing conditions; the MPP subpopulations shift toward more myeloid bias with increased animal age<sup>37</sup> or after external stress.<sup>62</sup>

The true LT-HSC state characterized by the most durable self-renewal has been pursued for decades.<sup>12,54,58,64-68</sup> Comparison of various HSC isolation strategies<sup>38</sup> allowed the definition of molecular overlap and surface marker overlap scores, which describe gene expression patterns correlating with stem cell self-renewal capacity. A major component of the molecular overlap score are genes negatively regulating the cell cycle, consistent with the long-standing idea that dormancy marks stem cells with the highest self-renewal.<sup>12</sup> The transition along the dormancy axis also seems to be gradual, with the dormant end of the spectrum characterized by deep quiescence (ie, slow return to the cell cycle), low biosynthetic activity, and high retinoic acid signaling,<sup>69</sup> as well as quiescence and low respiration being conserved among the most immature human HSPCs.<sup>70</sup>

The other aspect of HSC heterogeneity is lineage output bias, which may manifest in lineage priming. Although Kowalczyk et al<sup>37</sup> found no evidence for transcriptional lineage priming in HSCs regardless of age, Grover et al<sup>71</sup> identified a subset of HSCs expressing megakaryocytic genes, including the *Vwf* marker, in old mice. The reason for this disparity is not clear, but it may be due to cell isolation techniques, animal breeding or ageing, or the bioinformatic tools used. Nevertheless, the presence of megakaryocyte-biased HSCs has been confirmed through transplantation experiments<sup>19,71-73</sup> and tracking of native hematopoiesis,<sup>19</sup> indicating that committing to a megakaryocytic fate may be one of the earliest fate decisions. Interestingly, these megakaryocyte-biased HSCs can still behave as multipotent stem cells after transplantation, highlighting that cell potential does not necessarily reflect cell behavior under native conditions. HSC skewing toward myeloid and lymphoid lineages evident from transplantation experiments<sup>67,74,75</sup> remains unexplained within the transcriptional landscape; however, data reported by Mann et al<sup>76</sup> suggest that a myeloid-biased HSC subpopulation becomes detectable in aged mice after inflammation, consistent with previous functional data.<sup>77,78</sup>

## How to navigate the hematopoietic differentiation landscape

Because the HSC landscape seems to be at least in part continuous, traditional immunophenotyping approaches aiming to dissect distinct populations would now seem, to some extent, counterintuitive. Instead, broad and unbiased approaches

**Table 1. Selection of scRNA-seq data sets and resources relevant to hematopoiesis**

Reference	Cell n	Cell populations	Platform	Summary	Online resource
<b>HSCs</b>					
71	>130	BM: LT-HSCs (LSK, CD150 <sup>+</sup> , CD48 <sup>-</sup> )*	C1 + SMARTer	Profiling of young and old mice reveals expansion of platelet-biased HSCs	No
37	>1 100	BM: LT-HSCs (LSK, CD150 <sup>+</sup> , CD48 <sup>-</sup> ), ST-HSCs (LSK, CD150 <sup>-</sup> , CD48 <sup>-</sup> ), MPPs (LSK, CD150 <sup>-</sup> , CD48 <sup>+</sup> )*	SMART-Seq	Analysis of most immature stem and progenitor populations from old and young mice	Yes <sup>122</sup>
38	>90	BM: LT-HSCs (LSK, CD34 <sup>-</sup> , Flt3 <sup>-</sup> , CD48 <sup>-</sup> , CD150 <sup>-</sup> )*	SMART-Seq2	Analysis of the LT-HSCs reveals subpopulations with highest repopulation potential; surface marker data allow immunophenotyping	No
69	>310	BM: LT-HSCs (LSK, CD34 <sup>-</sup> , CD48 <sup>-</sup> , CD150 <sup>+</sup> , CD135 <sup>-</sup> ), LRC/non-LRC*	C1 + SMARTer	Comparison of active and dormant HSC populations	No
<b>Narrow gate data sets (HSPCs)</b>					
36	>2700	BM: Lin <sup>-</sup> , Kit <sup>+</sup> , Sca1 <sup>-</sup> cells*	MARS-Seq	Dissection of heterogeneity and lineage bias within CMPs, MEPs, GMPs; surface marker data allow immunophenotyping	No
63	>1 600	BM: HSPCs including: LT-HSCs, LMPPs, MPPs, IMEPs, CMPs, GMPs*	SMART-Seq2	Reconstruction of 3 differentiation trajectories (erythroid, granulocytic-monocytic, lymphoid); surface marker data allow immunophenotyping	Yes <sup>123</sup>
92	>2800	BM: LSK, LMPPs, CLPs + unipotent progenitors (ie, B cells, DCs, NK cells, Neu, and erythrocytes)*	CEL-Seq2	Highlights trajectories toward B cells, neutrophils/monocytes, DCs, and erythrocytes	Yes <sup>124</sup>
39	>380	BM: LSK, CMPs, GMPs, (Lin <sup>-</sup> Kit <sup>+</sup> , CD34 <sup>+</sup> )*	C1 + SMARTer	Dissection of heterogeneity and lineage bias in the intermediate progenitors	Yes <sup>125</sup>
19	>4 900	BM: LT-HSCs, ST-HSCs, MPP2s, MPP3s, MPP4s*	inDrops	Highlights lineage priming in the multipotent progenitor compartment and direct HSC-megakaryocyte differentiation trajectory	Yes <sup>126</sup>
70	>1 400	BM: HSPCs (CD34 <sup>+</sup> , Lin <sup>-</sup> )†	Quartz-Seq	Transcriptomics/functional data suggest cloud HSC state in human followed by unipotent progenitors; surface marker data allow immunophenotyping	Yes <sup>127</sup>
<b>Broad gate data sets (HSPCs + differentiated cells)</b>					
47	>2 300	PB: range of DC and monocyte populations*	SMART-Seq2	Analysis proposes a new taxonomy for human DCs and monocytes	Yes <sup>128</sup>
43	>400 000 BM: >38 000	51 mouse tissues, including total BM and c-Kit <sup>+</sup> fraction*	Microwell-Seq	Large-scale overview of the bone marrow and its progenitor compartment; also includes polymorphonuclear cells	Yes <sup>129</sup>

Raw and processed data can be obtained from the European Bioinformatics Institute Single Cell Expression Atlas<sup>130</sup> or Gene Expression Omnibus<sup>131</sup> databases.

AML, acute myeloid leukemia; BM, bone marrow; BMMc, bone marrow mononuclear cell; CB, cord blood; CBMC, cord blood mononuclear cell; DC, dendritic cell; EPO, erythropoietin; LK, Lin<sup>-</sup>, Kit<sup>+</sup>; LMPP, lymphoid-primed MPP; LRC, label-retaining cell; LSK, Lin<sup>-</sup>, Kit<sup>+</sup>, Sca1<sup>+</sup>; NK, natural killer; PB, peripheral blood; PBMC, peripheral blood mononuclear cell; WT, wild type.

\*Mouse.

†Human.

Table 1. (continued)

Reference	Cell n	Cell populations	Platform	Summary	Online resource
41	>7 300	BM: c-Kit <sup>+</sup> *	inDrops	Highlights basophil, megakaryocyte, and erythrocyte differentiation branches and effects of EPO stimulation	Yes <sup>130</sup>
40	>58 000	BM: LSK, LK*	10X	Multiple differentiation trajectories and identification of basophil/mast cell progenitors; comparison with c-Kit-defective hematopoietic system	Yes <sup>31</sup>
26	>180 000	PBMCs (WT) + BMMC (AML patients + controls)†	10X	Overview of the PBMCs (WT patients) and BMMCs before and after transplantation	No
97	>15 400	PBMCs: CD3 <sup>+</sup> T cells, CD11b <sup>+</sup> myeloid cells, CD19 <sup>+</sup> B cell‡	10X + REAP-Seq	Simultaneous analysis of the whole transcriptome and protein levels for 45 surface markers	No
42, 44	>270 000	BMMCs†	10X	Largest data set of human bone marrow yet	Yes <sup>132</sup>
10X Genomics	>8 300	PBMCs†	10X	PBMCs from a from a healthy human donor	No
116	>3 600	PBMCs†	Seq-Well	Resolves major cell types and highlights heterogeneity in the monocyte population	No
96	>8 000	CBMCs†	10X + CITE-Seq	Simultaneous analysis of the whole transcriptome and protein levels for 13 surface markers	No
117	>21 300	CB: CD34 <sup>+</sup> †	Drop-Seq	Analysis reconstructs trajectories toward 4 distinct cell fates in the most immature compartment	Yes <sup>133</sup>
42	>270 000	CBMCs†	10X	The largest data set of human cord blood yet	No
118	>8 000	BM: unfractionated + LSK, B cells, T cells/NK cells, granulocyte and monocyte fractions*	SMART-Seq2 + 10X	>100 000 cells profiled from 20 mouse organs	Yes <sup>134</sup>
<b>Gene perturbations and diseases</b>					
109	~70 000	BM: DCs*	10X	CRISPR perturbations provide insight into regulators of DCs	Yes <sup>35</sup>
112	>10 000	BM: myeloid cells (CD11c <sup>+</sup> ) cells and HSPCs (LSK)*	MARS-Seq	CRISPR perturbations provide insight into regulators of myeloid cells in vitro and in vivo	No
115	>2 000	BM: Lin <sup>-</sup> , CD34 <sup>+</sup> , CD38 <sup>-</sup> HSPCs†	SMART-Seq2	Simultaneous analysis of leukemic and normal cells from patients with chronic myeloid leukemia; modification to SMART-Seq2 protocol allows mutation detection	No
119	>970	BM: CD34 <sup>+</sup> †	C1 + SMARTer	Use of scRNA-seq to detect aneuploidy in patients	No

Raw and processed data can be obtained from the European Bioinformatics Institute Single Cell Expression Atlas<sup>20</sup> or Gene Expression Omnibus<sup>21</sup> databases.

AML, acute myeloid leukemia; BM, bone marrow; BMMC, bone marrow mononuclear cell; CB, cord blood; CBMC, cord blood mononuclear cell; DC, dendritic cell; EPO, erythropoietin; LK, Lin<sup>-</sup>, Kit<sup>+</sup>; LMPP, lymphoid-primed MPP; LRC, label-retaining cell; LSK, Lin<sup>-</sup>, Kit<sup>+</sup>, Sca1<sup>+</sup>; NK, natural killer; PB, peripheral blood; PBMC, peripheral blood mononuclear cell; WT, wild type.

\*Mouse.

†Human.

tracking possible routes through the landscape are required. Recently, there has been a surge of analytical methods for differentiation trajectory inference. They commonly rely on measuring the distance between cells in a high-dimensional gene expression space, under the assumption that cells with a small distance between them are related, for example, representing stages of the same differentiation trajectory. The aim is to find a measure of distance that reflects the structure of the data, rather than just relying on a shortest possible path (Figure 2A). To perform this task, most methods use dimensionality reduction techniques to learn the data structure<sup>79-82</sup> and simplify computation of cell-to-cell distances (Figure 2A). However, excessive dimensionality reduction will inevitably oversimplify the data (Figure 2B). Hence, there is a clear distinction between learned data (10-100 dimensional representation with minimal information loss) and visualizable data (2-3 dimensions interpretable by humans but at the expense of potentially discarding important biological information; Figure 2C). This means that although learned data are used for trajectory inference, it is not advisable to interpret complex differentiation pathways simply from a 2-dimensional representation of a data set (tSNE, UMAP).

Early methods for trajectory inference focused on the idea of pseudotime, where the calculated distances are used to order cells along a putative trajectory according to their distance from a predefined starting point.<sup>33,34,83-87</sup> This ordering allows recovery of the pseudodynamic gene expression along a trajectory. However, the need to assign cells to unidirectional trajectories restricts its applicability. Attempts to improve upon this by implementing methods that can discover putative branch points in a data set may provide a solution, but their ability to produce bifurcations consistent with functional data in an unsupervised manner has been limited.<sup>87-89</sup> Alternative methods are aiming to identify the likely connection between clustered cells (StemID,<sup>90</sup> PAGA<sup>91</sup>) or stepwise cell classification working back toward the stem cells (FateID<sup>92</sup>). Population balance analysis takes a physically motivated approach,<sup>35</sup> attempting to constrain the problem of trajectory inference into a set of differential equations describing the flux of cells through the transcriptional landscape. These methods and others like them<sup>93</sup> successfully recapitulate trajectories through the unipotent hematopoietic states but can also elucidate routes of differentiation among the heterogeneous HSPCs, as highlighted in the following section.

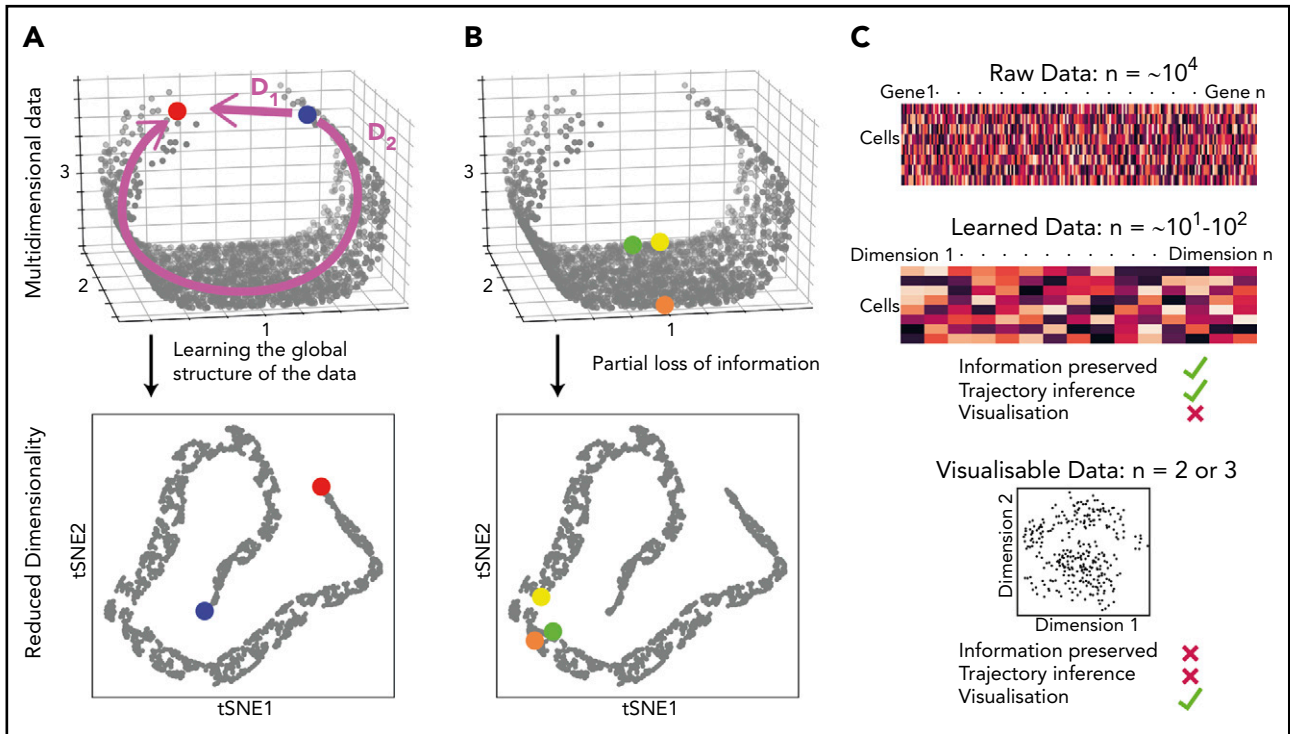
## An evolving view of HSPC hierarchy

An early analysis of the multipotent stem/progenitor compartment examined 1600 cells spanning 10 classically defined HSPC populations.<sup>63</sup> Although the surface marker information served as reference points for this and future studies, scRNA-seq analysis identified 3 broad trajectories in the differentiation landscape: lymphoid, erythroid, and granulocytic/monocytic, together with their gene expression signatures. A more complete view was provided by high-throughput scRNA-seq platforms, which allowed dense cell sampling of large populations (typically >4000 cells), overcoming the restriction of narrow sorting gates. Consequently, trajectories toward megakaryocytes, erythrocytes, monocytes/DCs, lymphoid cells, neutrophils, and rare populations of basophils, mast cells, and eosinophils are now beginning to be defined.<sup>40,41</sup> The emerging cellular hierarchy is largely consistent

with that inferred from recent *in situ* barcoding studies,<sup>18,19</sup> with the exception of the erythroid lineage. This branch seems to be coupled with the megakaryocytic fate in scRNA-seq experiments,<sup>41,91</sup> whereas barcoding data indicate closer linkage with monocytic/granulocytic lineages. Additional experiments focusing on native hematopoiesis will be required to resolve this issue.

The identified trajectories in scRNA-seq data contain a wealth of information. Using gene expression correlations, it is possible to extract putative regulators of fate decision and drivers or inhibitors of differentiation pathways, which will greatly facilitate future studies.<sup>41,63</sup> In-depth analysis of a single trajectory can also help explain the dynamics of differentiation. Tusi et al<sup>41</sup> delineated stages of early erythroid differentiation within the transcriptional landscape, including traditionally defined erythroid colony-forming unit and erythroid burst-forming unit populations, and described cell-cycle remodeling at putative cell amplification stages during differentiation. This provided a platform on which to analyze the effects of erythropoietin stimulation on the progenitor population structure and their cell-cycle profiles. Global analysis of the progenitor landscape revealed a surprising pattern of multiple types of mature cells arising via >1 trajectory. Analysis of c-Kit<sup>+</sup> cells, for example, suggested that monocytes have 2 origins: 1 coupled with the DC branch and 1 with the neutrophil branch.<sup>41</sup> Similarly, megakaryocytes are predicted to arise directly from megakaryocyte-biased HSCs as well as through a MPP intermediate.<sup>19</sup> Finally, the yet unpublished PAGA analysis of the Lin<sup>-</sup> c-Kit<sup>+</sup> population suggests that basophils can originate from the neutrophil/monocyte branch or from the megakaryocyte/erythroid branch,<sup>91</sup> with the latter supported by recent computational results.<sup>41</sup> Although the complexity of megakaryocyte and monocyte trajectories is supported by experimental data,<sup>19,48,71</sup> firmly establishing the existence of alternative basophil trajectories would still benefit from further experimental validation.

Although mice have proved to be an excellent model system, there is accumulating evidence suggesting that human progenitors may be organized differently. A 2016 study used clonal assays and xenograft transplantations to show that in human adult bone marrow, unlike the fetal liver, previously described oligopotent progenitors (MPPs, CMPs, multilymphoid progenitors) are predominantly unipotent and arise directly from the multipotent stem cell compartment.<sup>55</sup> Subsequently, a detailed scRNA-seq study analyzed the most immature HSPC compartments (Lin<sup>-</sup>, CD34<sup>+</sup>, CD38<sup>-</sup>, and Lin<sup>-</sup>, CD34<sup>+</sup>, CD38<sup>+</sup>) and revealed that although the unipotent progenitor populations described indeed form discrete subpopulations, the stem cell compartment seems to form a rather continuous structure.<sup>70</sup> From the lack of obvious priming in the immature populations (with the exception of a minor lymphoid/myeloid vs megakaryocyte/erythroid bias), the authors concluded that HSCs exist in a fluid cloud state, which gives rise directly to committed progenitors without much cell hierarchy in between. Furthermore, even the least primed populations can give rise to single lineages *in vitro*, indicating that true multipotent cells constitute only a small fraction of the conventional stem cell population. These findings are at odds with evidence of the HSPC hierarchy from the murine system, where the upper tiers, although difficult to resolve molecularly, have functionally multipotent output.



**Figure 2. Distances and data dimensionality.** (A) A set of single cells expressing 3 genes arranged along a curved shape has been simulated. There are 2 measures of distance between the blue and red cells. Whereas  $D_1$  represents the shortest possible distance between the 2 cells,  $D_2$  is the distance between the cells through the structure of the data (manifold). The two arms of the curved shape may represent continuous transition processes (eg, cell differentiation); therefore, distance  $D_2$  is the important distance measure. A dimensionality reduction technique (here tSNE) should capture such features. (B) Excessive reduction in dimensionality causes important information to be lost. In this case, a 2-dimensional representation of the data incorrectly suggests that the green cell is farther from the yellow cell than the orange cell, because information has been lost about axis 2. (C) To infer cellular trajectories from scRNA-seq data, dimensionality reduction is used to learn the structure of the data (learned data), which captures the important distances between cells in a suitable number of dimensions, typically 10 to 100. Trajectory inference can then be attempted from this learned data. For visualization, the dimensionality of the data needs to be reduced to either 2 or 3, but this will inevitably result in the loss of some of the important biological information, rendering data unsuitable for trajectory inference.

Although differences between the human and mouse hematopoietic landscapes are bound to exist, drawing conclusions about the underlying reasons remains challenging, because current scRNA-seq data analyses produce a rather flat HSPC landscape, with limited ability to resolve distinct cellular states.<sup>19,37,38,41,63</sup> Some aspects of population structure may be drowned in the noise caused by processes such as cell cycle or metabolism. Furthermore, current scRNA-seq data represent an incomplete view of cellular states as a result of the failure to detect lowly expressed genes and the absence of entire domains of information, such as protein levels or epigenetic status. Importantly, current mouse and human data sets have been analyzed using different bioinformatic methods, and careful cross analysis is required before drawing any strong conclusions.

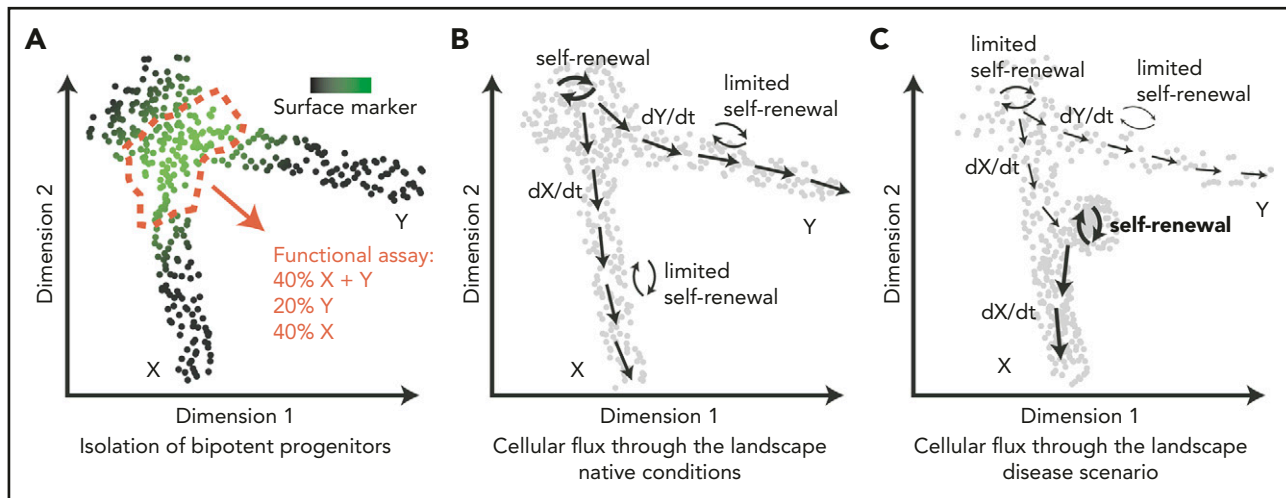
## Outlook: extending the paradigm

In light of the accumulated scRNA-seq data, the concept of clearly demarcated stem and progenitor cell types becomes questionable. Barring the caveat of hidden variables (proteins, epigenetics, cell location), the landscape encompassing the most primitive HSPCs appears mostly continuous and flat. It is difficult to argue for a reason why the progenitor hierarchy needs to be discrete, other than simplicity or aesthetics. Indeed, even the very surface markers used for progenitor isolation commonly exhibit continuous, rather than discrete, levels. The notion of progenitor types has been historically dictated by technical

limitations: ability to observe cells using only a small number of markers and limited number of parallel functional assays. By contrast, in a transcriptomic landscape, each cell is positioned using information from several thousand genes.

Of note, a continuous and flat shape does not imply lack of embedded features or information. On the contrary, it is evident that positions/territories within this space are functionally relevant and associated with key functional qualities, such as durability of self-renewal<sup>38,69</sup> or cell output,<sup>19,36,41,94</sup> measured using a variety of techniques, including cell barcoding,<sup>19</sup> HSC transplantation assays,<sup>36,38,69</sup> or in vitro clonal assays.<sup>41</sup> However, because of this continuous nature, a typical flow cytometry gating approach provides an arbitrary section through the progenitor landscape, providing a mixture of cells and covering a range of functional outputs (Figure 3A). Although undoubtedly useful, isolation of specific populations offers a fundamentally restricted insight into the organization and dynamics of stem and progenitor cells.

Transcriptomic data offer a more complex and likely more faithful representation, without the need for subjective categories. Moreover, if necessary for experimental purposes, discretization is assisted by rich expression information. Although scRNA-seq data are static, the method encodes information on cell states as they exist in vivo, and at deep coverage, it has the potential to capture molecular states representative of cellular transitions. This means that for each location within the landscape



**Figure 3. From transcriptomic landscapes to tissue function.** The 3 diagrams depict 2-dimensional landscapes with single cells as points, where a pool of progenitors gives rise to 2 differentiated populations X and Y. (A) A classical approach to identify a specific bipotent progenitor population, an immunophenotypic gate drawn based on expression of a surface marker (green) sections through the landscape and yields a heterogeneous population of cells upon functional analysis. (B) Normal hematopoiesis; dynamic information about cellular fluxes (arrows) indicates directions of differentiation throughout the landscape and degree of self-renewal. (C) An example of how cell-flux analysis reveals mechanisms of a disease. In this case, the stem cell pool is exhausted (low self-renewal), compromising production of Y cells, whereas there is increased production of X cells because of acquired increased self-renewal of progenitors X.

(ie, cellular state), it may become possible to infer transition directions and probabilities associated with them under native conditions. Quantitative description of this cell flux through the multidimensional space will constitute a major advance (Figure 3B). We highlight 4 directions that will facilitate this process and combined will advance our understanding of hematopoiesis.

### Toward complete cell-state information

The single-cell landscape constitutes an essential framework but is currently limited to a subset of messenger RNA information and lacks potential heterogeneity hidden at protein, epigenetic, or tissue location level. Further development of scRNA-seq technology will offer denser sampling<sup>95</sup> and increased coverage/accuracy of transcriptional profiling<sup>30</sup> and, combined with new techniques, will enable simultaneous detection of proteins (CITE-Seq or REAP-Seq)<sup>96,97</sup> or chromatin status.<sup>98-101</sup> Imaging-based transcriptomics is also being developed to complement the data with spatial information.<sup>102,103</sup> Altogether, this near-complete information will provide precise locations of cells in a multidimensional feature space and tie together information at molecular, cellular, and tissue levels.

### Advanced analytic tools

As increasingly complex data sets accumulate, analysis becomes more challenging. This includes trajectory inference, identification of branching points, and extraction of gene regulatory information. Quantitative description of cell fluxes through the progenitor space is still in its infancy, but newly developed numerical frameworks already attempt to approximate cell transitions from snapshot data (described in "How to navigate the haematopoietic differentiation landscape").<sup>35,91,104</sup> Nonetheless, parallel experimental information remains critical to provide directionality and real-time information for cell differentiation processes.

### Real-time cell flux through the progenitor space

Rather than discretizing HSPCs by immunophenotyping, a more promising approach would be to link cell positions within the landscape with experimentally derived cell output (Figure 3B), under both native and transplantation conditions. Thus, each position would encode information on cellular flow in particular directions, quantifying differentiation and self-renewal. New technologies allowing simultaneous *in vivo/in vitro* barcoding and single-cell transcriptional profiling (where a barcode can be assigned to cells within scRNA-seq data) are emerging, using either transposon tagging<sup>105</sup> or CRISPR scarring.<sup>106,107</sup> This should enable integration of real-time cell ancestry information with transcriptomic landscapes and potentially allow for the identification of transcriptomic signals generated only transiently during real-time differentiation. For finer time scales and insight into cell cycle-related effects, pulse-chase experiments may become important.<sup>108</sup>

### Molecular mechanisms driving cell transitions

To manipulate hematopoiesis, we need to understand the underlying molecular mechanisms. Correlations inferred from expression data provide ample hypotheses for regulatory mechanisms, but experimental testing is essential. Gene regulatory networks and chromatin state can now be efficiently interrogated with recently developed techniques combining CRISPR screening with scRNA-seq.<sup>109-114</sup> Targeting multiple genes and observing the effects globally will help us understand how RNA, proteins, epigenetics, and extrinsic signals establish the shape of the differentiation landscape and drive cellular fluxes.

Altogether, these approaches will provide a reference framework with computational modeling capacity, a promising starting point for understanding abnormal hematopoiesis. Even at this early stage, single-cell transcriptomics provides valuable insight. As reported by Dahlin et al,<sup>40</sup> global analysis of



the progenitor compartment revealed that c-Kit–defective signaling reshapes the top of the hematopoietic hierarchy and blocks the mast cell fate at an early stage in mice. Combining scRNA-seq with detection of mutations in single cells opens the possibility of analyzing samples directly from leukemia patients. Giustachini et al<sup>15</sup> applied this approach to identify cell sub-populations persisting after prolonged chemotherapy and related to blast crisis in chronic myeloid leukemia patients. A global comparison of single cell landscapes between leukemic and wild-type states will reveal new cellular states or changes in the cellular fluxes associated with, for example, changes in self-renewal or enhanced/reduced differentiation in particular lineages (Figure 3C). Combining our knowledge on cell flux with regulatory mechanisms will enable a more informed development of future therapies. Thus, we expect that scRNA-seq analysis will soon shed new light on leukemia pathogenesis and become closely relevant in the clinical setting.

## Acknowledgments

Work in the Gottgens Laboratory is supported by grants from Wellcome, Bloodwise, Cancer Research UK; National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases (DK106766); and core support grants by Wellcome to the Wellcome–

Medical Research Council Cambridge Stem Cell Institute. S.W. is the recipient of a Medical Research Council Studentship.

## Authorship

Contribution: S.W., I.K., and B.G. wrote the paper together; S.W. and I.K. generated the diagrams; and all authors approved the final manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: S.W., 0000-0002-8399-7875; I.K., 0000-0002-9385-0359; B.G., 0000-0001-6302-5705.

Correspondence: Berthold Gottgens, Cambridge University Department of Haematology, Cambridge Institute for Medical Research & Wellcome and MRC Cambridge Stem Cell Institute Hills Rd, Cambridge CB2 0XY, UK; e-mail: bg200@cam.ac.uk; and Iwo Kucinski, Cambridge University Department of Haematology, Cambridge Institute for Medical Research & Wellcome and MRC Cambridge Stem Cell Institute Hills Rd, Cambridge CB2 0XY, UK; e-mail: idk25@cam.ac.uk.

## Footnote

Submitted 15 August 2018; accepted 22 October 2018. Prepublished online as *Blood* First Edition paper, 6 February 2019; DOI 10.1182/blood-2018-08-835355.

## REFERENCES

- Boisset JC, Robin C. On the origin of hematopoietic stem cells: progress and controversy. *Stem Cell Res (Amst)*. 2012;8(1):1-13.
- Jacobson LO, Simmons EL, Marks EK, Eldredge JH. Recovery from radiation injury. *Science*. 1951;113(2940):510-511.
- Lorenz E, Uphoff D, Reid TR, Shelton E. Modification of irradiation injury in mice and guinea pigs by bone marrow injections. *J Natl Cancer Inst*. 1951;12(1):197-201.
- Urso P, Congdon CC. The effect of the amount of isologous bone marrow injected on the recovery of hematopoietic organs, survival and body weight after lethal irradiation injury in mice. *Blood*. 1957;12(3):251-260.
- Iscove NN, Till JE, McCulloch EA. The proliferative states of mouse granulopoietic progenitor cells. *Proc Soc Exp Biol Med*. 1970;134(1):33-36.
- Eaves CJ. Hematopoietic stem cells: concepts, definitions, and the new reality. *Blood*. 2015;125(17):2605-2613.
- Thomas ED, Lochte HL Jr, Cannon JH, Sahler OD, Ferrebee JW. Supralethal whole body irradiation and isologous marrow transplantation in man. *J Clin Invest*. 1959;38:1709-1716.
- Thomas ED. A history of haemopoietic cell transplantation. *Br J Haematol*. 1999;105(2):330-339.
- Passweg JR, Baldomero H, Bader P, et al. Hematopoietic stem cell transplantation in Europe 2014: more than 40 000 transplants annually. *Bone Marrow Transplant*. 2016;51(6):786-792.
- Julius MH, Masuda T, Herzenberg LA. Demonstration that antigen-binding cells are precursors of antibody-producing cells after purification with a fluorescence-activated cell sorter. *Proc Natl Acad Sci USA*. 1972;69(7):1934-1938.
- Köhler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*. 1975;256(5517):495-497.
- Wilson A, Laurenti E, Oser G, et al. Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell*. 2008;135(6):1118-1129.
- Orkin SH, Zon LI. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell*. 2008;132(4):631-644.
- Doulatov S, Notta F, Laurenti E, Dick JE. Hematopoiesis: a human perspective. *Cell Stem Cell*. 2012;10(2):120-136.
- Nimmo RA, May GE, Enver T. Primed and ready: understanding lineage commitment through single cell analysis. *Trends Cell Biol*. 2015;25(8):459-467.
- Morrison SJ, Weissman IL. The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. *Immunity*. 1994;1(8):661-673.
- Sun J, Ramos A, Chapman B, et al. Clonal dynamics of native haematopoiesis. *Nature*. 2014;514(7522):322-327.
- Pei W, Feyerabend TB, Rössler J, et al. Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature*. 2017;548(7668):456-460.
- Rodriguez-Fraticelli AE, Wolock SL, Weinreb CS, et al. Clonal analysis of lineage fate in native haematopoiesis. *Nature*. 2018;553(7687):212-216.
- Laurenti E, Göttgens B. From haematopoietic stem cells to complex differentiation landscapes. *Nature*. 2018;553(7689):418-426.
- Ema H, Morita Y, Suda T. Heterogeneity and hierarchy of hematopoietic stem cells. *Exp Hematol*. 2014;42(2):74-82.e2.
- Hu M, Krause D, Greaves M, et al. Multilineage gene expression precedes commitment in the hemopoietic system. *Genes Dev*. 1997;11(6):774-785.
- Svensson V, Vento-Tormo R, Teichmann SA. Exponential scaling of single-cell RNA-seq in the past decade. *Nat Protoc*. 2018;13(4):599-604.
- Macosko EZ, Basu A, Satija R, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*. 2015;161(5):1202-1214.
- Klein AM, Mazutis L, Kartuna I, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*. 2015;161(5):1187-1201.
- Zheng GXY, Terry JM, Belgrader P, et al. Massively parallel digital transcriptional profiling of single cells. *Nat Commun*. 2017;8:14049.
- Picelli S, Faridani OR, Björklund ÅK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc*. 2014;9(1):171-181.
- Hashimshony T, Senderovich N, Avital G, et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol*. 2016;17:77.
- Bagnoli JW, Ziegenhain C, Janjic A, et al. Sensitive and powerful single-cell RNA sequencing using mcSCR-seq. *Nat Commun*. 2018;9(1):2937.
- Hayashi T, Ozaki H, Sasagawa Y, Umeda M, Danno H, Nikaido I. Single-cell full-length

- total RNA sequencing uncovers dynamics of recursive splicing and enhancer RNAs. *Nat Commun.* 2018;9(1):619.
31. Svensson V, Natarajan KN, Ly LH, et al. Power analysis of single-cell RNA-sequencing experiments. *Nat Methods.* 2017;14(4):381-387.
  32. Ziegenhain C, Vieth B, Parekh S, et al. Comparative analysis of single-cell RNA sequencing methods. *Mol Cell.* 2017;65(4):631-643.e4.
  33. Trapnell C, Cacchiarelli D, Grimsby J, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol.* 2014;32(4):381-386.
  34. Bendall SC, Davis KL, Amir AD, et al. Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. *Cell.* 2014;157(3):714-725.
  35. Weinreb C, Wolock S, Tusi BK, Socolovsky M, Klein AM. Fundamental limits on dynamic inference from single-cell snapshots. *Proc Natl Acad Sci USA.* 2018;115(10):E2467-E2476.
  36. Paul F, Arkin Y, Giladi A, et al. Transcriptional heterogeneity and lineage commitment in myeloid progenitors [published correction appears in *Cell.* 2016;164(1-2):325]. *Cell.* 2015;163(7):1663-1677.
  37. Kowalczyk MS, Tirosh I, Heckl D, et al. Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. *Genome Res.* 2015;25(12):1860-1872.
  38. Wilson NK, Kent DG, Buettner F, et al. Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. *Cell Stem Cell.* 2015;16(6):712-724.
  39. Olsson A, Venkatasubramanian M, Chaudhri VK, et al. Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature.* 2016;537(7622):698-702.
  40. Dahlin JS, Hamey FK, Pijuan-Sala B, et al. A single-cell hematopoietic landscape resolves 8 lineage trajectories and defects in Kit mutant mice. *Blood.* 2018;131(21):e1-e11.
  41. Tusi BK, Wolock SL, Weinreb C, et al. Population snapshots predict early haematopoietic and erythroid hierarchies. *Nature.* 2018;555(7694):54-60.
  42. Regev A, Teichmann SA, Lander ES, et al; Human Cell Atlas Meeting Participants. The Human Cell Atlas. *Elife.* 2017;6:e27041.
  43. Han X, Wang R, Zhou Y, et al. Mapping the mouse cell atlas by microwell-Seq [published correction appears in *Cell.* 2018;173(5):1307]. *Cell.* 2018;172(5):1091-1107.e17.
  44. Hay SB, Ferchen K, Chetal K, Grimes HL, Salomonis N. The Human Cell Atlas bone marrow single-cell interactive web portal. *Exp Hematol.* 2018;68:51-61.
  45. Papalexi E, Satija R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol.* 2018;18(1):35-45.
  46. Jaitin DA, Kenigsberg E, Keren-Shaul H, et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science.* 2014;343(6172):776-779.
  47. Villani AC, Satija R, Reynolds G, et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science.* 2017;356(6335):eaah4573.
  48. Yáñez A, Coetzee SG, Olsson A, et al. Granulocyte-monocyte progenitors and monocyte-dendritic cell progenitors independently produce functionally distinct monocytes. *Immunity.* 2017;47(5):890-902.e4.
  49. Menezes S, Melandri D, Anselmi G, et al. The heterogeneity of Ly6C<sup>hi</sup> monocytes controls their differentiation into iNOS<sup>+</sup> macrophages or monocyte-derived dendritic cells. *Immunity.* 2016;45(6):1205-1218.
  50. Adolfsson J, Månsson R, Buza-Vidas N, et al. Identification of Flt3<sup>+</sup> lympho-myeloid stem cells lacking erythro-megakaryocytic potential: a revised road map for adult blood lineage commitment. *Cell.* 2005;121(2):295-306.
  51. Görgens A, Radtke S, Möllmann M, et al. Revision of the human hematopoietic tree: granulocyte subtypes derive from distinct hematopoietic lineages. *Cell Reports.* 2013;3(5):1539-1552.
  52. Pronk CJH, Rossi DJ, Månsson R, et al. Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. *Cell Stem Cell.* 2007;1(4):428-442.
  53. Alberti-Servera L, von Muenchow L, Tzapogas P, et al. Single-cell RNA sequencing reveals developmental heterogeneity among early lymphoid progenitors. *EMBO J.* 2017;36(24):3619-3633.
  54. Yamamoto R, Morita Y, Ooehara J, et al. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell.* 2013;154(5):1112-1126.
  55. Notta F, Zandi S, Takayama N, et al. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science.* 2016;351(6269):aab2116.
  56. Naik SH, Perié L, Swart E, et al. Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature.* 2013;496(7444):229-232.
  57. Akashi K, Traver D, Miyamoto T, Weissman IL. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature.* 2000;404(6774):193-197.
  58. Oguro H, Ding L, Morrison SJ. SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell.* 2013;13(1):102-116.
  59. Bock C, Beerman I, Lien WH, et al. DNA methylation dynamics during in vivo differentiation of blood and skin stem cells. *Mol Cell.* 2012;47(4):633-647.
  60. Gazit R, Garrison BS, Rao TN, et al; Immunological Genome Project Consortium. Transcriptome analysis identifies regulators of hematopoietic stem and progenitor cells. *Stem Cell Reports.* 2013;1(3):266-280.
  61. Cabezas-Wallscheid N, Klimmeck D, Hansson J, et al. Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. *Cell Stem Cell.* 2014;15(4):507-522.
  62. Pietras EM, Reynaud D, Kang YA, et al. Functionally distinct subsets of lineage-biased multipotent progenitors control blood production in normal and regenerative conditions. *Cell Stem Cell.* 2015;17(1):35-46.
  63. Nestorowa S, Hamey FK, Pijuan Sala B, et al. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood.* 2016;128(8):e20-e31.
  64. Kiel MJ, Yilmaz ÖH, Iwashita T, Yilmaz OH, Terhorst C, Morrison SJ. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell.* 2005;121(7):1109-1121.
  65. Kent DG, Copley MR, Benz C, et al. Prospective isolation and molecular characterization of hematopoietic stem cells with durable self-renewal potential. *Blood.* 2009;113(25):6342-6350.
  66. Muller-Sieburg CE, Sieburg HB. The GOD of hematopoietic stem cells: a clonal diversity model of the stem cell compartment. *Cell Cycle.* 2006;5(4):394-398.
  67. Dykstra B, Kent D, Bowie M, et al. Long-term propagation of distinct hematopoietic differentiation programs in vivo. *Cell Stem Cell.* 2007;1(2):218-229.
  68. Upadhaya S, Reizis B, Sawai CM. New genetic tools for the in vivo study of hematopoietic stem cell function. *Exp Hematol.* 2018;61:26-35.
  69. Cabezas-Wallscheid N, Buettner F, Sommerkamp P, et al. Vitamin A-retinoic acid signaling regulates hematopoietic stem cell dormancy. *Cell.* 2017;169(5):807-823.e19.
  70. Velten L, Haas SF, Raffel S, et al. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol.* 2017;19(4):271-281.
  71. Grover A, Sanjuan-Pla A, Thongjuea S, et al. Single-cell RNA sequencing reveals molecular and functional platelet bias of aged haematopoietic stem cells. *Nat Commun.* 2016;7:11075.
  72. Sanjuan-Pla A, Macaulay IC, Jensen CT, et al. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature.* 2013;502(7470):232-236.
  73. Carrelha J, Meng Y, Kettle LM, et al. Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature.* 2018;554(7690):106-111.
  74. Müller-Sieburg CE, Cho RH, Thoman M, Adkins B, Sieburg HB. Deterministic regulation of hematopoietic stem cell self-renewal and differentiation. *Blood.* 2002;100(4):1302-1309.
  75. Benz C, Copley MR, Kent DG, et al. Hematopoietic stem cell subtypes expand differentially during development and

- display distinct lymphopoietic programs. *Cell Stem Cell*. 2012;10(3):273-283.
76. Mann M, Mehta A, De Boer C, et al. Heterogeneous responses of hematopoietic stem cells to inflammatory stimuli are altered with age. *Cell Rep*. 2018;25(11):2992-3005.e5.
  77. Pietras EM, Mirantes-Barbeito C, Fong S, et al. Chronic interleukin-1 exposure drives haematopoietic stem cells towards precocious myeloid differentiation at the expense of self-renewal. *Nat Cell Biol*. 2016;18(6):607-618.
  78. Pang WW, Price EA, Sahoo D, et al. Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proc Natl Acad Sci USA*. 2011;108(50):20012-20017.
  79. Van Der Maaten LJP, Hinton GE. Visualizing high-dimensional data using t-sne. *J Mach Learn Res*. 2008;9:2579-2605.
  80. Haghverdi L, Buettner F, Theis FJ. Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics*. 2015;31(18):2989-2998.
  81. Weinreb C, Wolock S, Klein AM. SPRING: a kinetic interface for visualizing high dimensional single-cell expression data. *Bioinformatics*. 2018;34(7):1246-1248.
  82. McInnes L, Healy J, Melville J. UMAP: uniform manifold approximation and projection for dimension reduction. <https://arxiv.org/pdf/1802.03426.pdf>. Accessed 14 February 2019.
  83. Ocone A, Haghverdi L, Mueller NS, Theis FJ. Reconstructing gene regulatory dynamics from high-dimensional single-cell snapshot data. *Bioinformatics*. 2015;31(12):i89-i96.
  84. Marco E, Karp RL, Guo G, et al. Bifurcation analysis of single-cell gene expression data reveals epigenetic landscape. *Proc Natl Acad Sci USA*. 2014;111(52):E5643-E5650.
  85. Setty M, Tadmor MD, Reich-Zeliger S, et al. Wishbone identifies bifurcating developmental trajectories from single-cell data. *Nat Biotechnol*. 2016;34(6):637-645.
  86. Shin J, Berg DA, Zhu Y, et al. Single-cell RNA-Seq with waterfall reveals molecular cascades underlying adult neurogenesis. *Cell Stem Cell*. 2015;17(3):360-372.
  87. Haghverdi L, Büttner M, Wolf FA, Buettner F, Theis FJ. Diffusion pseudotime robustly reconstructs lineage branching. *Nat Methods*. 2016;13(10):845-848.
  88. Welch JD, Hartemink AJ, Prins JF. SLICER: inferring branched, nonlinear cellular trajectories from single cell RNA-seq data. *Genome Biol*. 2016;17(1):106.
  89. Giecord G, Marco E, Garcia SP, Trippa L, Yuan GC. Robust lineage reconstruction from high-dimensional single-cell data. *Nucleic Acids Res*. 2016;44(14):e122.
  90. Grün D, Muraro MJ, Boisset JC, et al. De novo prediction of stem cell identity using single-cell transcriptome data. *Cell Stem Cell*. 2016;19(2):266-277.
  91. Wolf FA, Hamey F, Plass M, et al. Paga: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol*. In press.
  92. Herman JS, Sagar, Grün D. FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat Methods*. 2018;15(5):379-386.
  93. Lummertz da Rocha E, Rowe RG, Lundin V, et al. Reconstruction of complex single-cell trajectories using CellRouter. *Nat Commun*. 2018;9(1):892.
  94. Khoramian Tusi B, Socolovsky M. High-throughput single-cell fate potential assay of murine hematopoietic progenitors in vitro. *Exp Hematol*. 2018;60:21-29.e3.
  95. Rosenberg AB, Roco CM, Muscat RA, et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science*. 2018;360(6385):176-182.
  96. Stoeckius M, Hafemeister C, Stephenson W, et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods*. 2017;14(9):865-868.
  97. Peterson VM, Zhang KX, Kumar N, et al. Multiplexed quantification of proteins and transcripts in single cells. *Nat Biotechnol*. 2017;35(10):936-939.
  98. Buenrostro JD, Corces MR, Lareau CA, et al. Integrated single-cell analysis maps the continuous regulatory landscape of human hematopoietic differentiation. *Cell*. 2018;173(6):1535-1548.e16.
  99. Lay F, Kelly T, Jones P. Nucleosome occupancy and methylome sequencing (NOME-seq). *Methods Mol Biol*. 2018;1708:267-284.
  100. Cusanovich DA, Daza R, Adey A, et al. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science*. 2015;348(6237):910-914.
  101. Clark SJ, Argelaguet R, Kapourani CA, et al. scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat Commun*. 2018;9(1):781.
  102. Crosetto N, Bienko M, van Oudenaarden A. Spatially resolved transcriptomics and beyond. *Nat Rev Genet*. 2015;16(1):57-66.
  103. Wang X, Allen WE, Wright MA, et al. Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science*. 2018;361(6400):eaat5691.
  104. Schiebinger G, Shu J, Tabaka M, et al. Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell*. 2019;176(4):928-943.
  105. Wagner DE, Weinreb C, Collins ZM, Briggs JA, Megason SG, Klein AM. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science*. 2018;360(6392):981-987.
  106. Raj B, Wagner DE, McKenna A, et al. Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. *Nat Biotechnol*. 2018;36(5):442-450.
  107. Spanjaard B, Hu B, Mitic N, et al. Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. *Nat Biotechnol*. 2018;36(5):469-473.
  108. Akinduro O, Weber TS, Ang H, et al. Proliferation dynamics of acute myeloid leukaemia and haematopoietic progenitors competing for bone marrow space. *Nat Commun*. 2018;9(1):519.
  109. Dixit A, Parnas O, Li B, et al. Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell*. 2016;167(7):1853-1866.e17.
  110. Adamson B, Norman TM, Jost M, et al. A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell*. 2016;167(7):1867-1882.e21.
  111. Datlinger P, Rendeiro AF, Schmid C, et al. Pooled CRISPR screening with single-cell transcriptome readout. *Nat Methods*. 2017;14(3):297-301.
  112. Jaitin DA, Weiner A, Yofe I, et al. Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-Seq. *Cell*. 2016;167(7):1883-1896.e15.
  113. Giladi A, Paul F, Herzog Y, et al. Single-cell characterization of haematopoietic progenitors and their trajectories in homeostasis and perturbed haematopoiesis. *Nat Cell Biol*. 2018;20(7):836-846.
  114. Xie S, Duan J, Li B, Zhou P, Hon GC. Multiplexed engineering and analysis of combinatorial enhancer activity in single cells. *Mol Cell*. 2017;66(2):285-299.e5.
  115. Giustacchini A, Thongjuea S, Barkas N, et al. Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. *Nat Med*. 2017;23(6):692-702.
  116. Gierahn TM, Wadsworth MH II, Hughes TK, et al. Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. *Nat Methods*. 2017;14(4):395-398.
  117. Zheng S, Papalexis E, Butler A, Stephenson W, Satija R. Molecular transitions in early progenitors during human cord blood hematopoiesis. *Mol Syst Biol*. 2018;14(3):e8041.
  118. Tabula Muris Consortium. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature*. 2018;562(7727):367-372.
  119. Zhao X, Gao S, Wu Z, et al. Single-cell RNA-seq reveals a distinct transcriptome signature of aneuploid hematopoietic cells. *Blood*. 2017;130(25):2762-2773.
  120. European Bioinformatics Institute. Single Cell Expression Atlas. <https://www.ebi.ac.uk/gxa/sc/home>. Accessed 14 February 2019.
  121. National Center for Biotechnology Information. Gene Expression Omnibus. <https://www.ncbi.nlm.nih.gov/geo>. Accessed 14 February 2019.
  122. Single Cell Portal, Aviv Regev's laboratory, Broad Institute, Cambridge. Available at: [https://portals.broadinstitute.org/single\\_cell/study/single-cell-ma-seq-reveals-changes-in-cell-cycle-and-differentiation-programs-upon-aging-of-hematopoietic-stem-cells#study-visualize](https://portals.broadinstitute.org/single_cell/study/single-cell-ma-seq-reveals-changes-in-cell-cycle-and-differentiation-programs-upon-aging-of-hematopoietic-stem-cells#study-visualize). Accessed 15 February 2019.

123. Gene and protein expression in adult haematopoiesis, Berthold Gottgens' laboratory, Wellcome-MRC Cambridge Stem Cell Institute, Available at: [http://blood.stemcells.cam.ac.uk/single\\_cell\\_atlas.html](http://blood.stemcells.cam.ac.uk/single_cell_atlas.html). Accessed 15 February 2019.
124. Fate Bias Inference in Lymphoid Progenitors, Dominic Grün's laboratory, Max Planck Institute of Immunobiology and Epigenetics, Freiburg. Available at: <http://hematopoietic-progenitors.ie-freiburg.mpg.de>. Accessed 15 February 2019.
125. AltAnalyze, Nathan Salomonis' laboratory, Cincinnati Children's Hospital Medical Center, University of Cincinnati. Available at: <http://www.altanalyze.org/hematopoietic.html>. Accessed 15 February 2019.
126. Klein Lab tools, Allon Klein's laboratory, Department of Systems Biology, Harvard Medical School. Available at: [https://kleintools.hms.harvard.edu/tools/springViewer\\_1\\_6\\_dev.html?client\\_datasets/ARF2017\\_combined\\_nocycle/ARF2017\\_combined\\_nocycle](https://kleintools.hms.harvard.edu/tools/springViewer_1_6_dev.html?client_datasets/ARF2017_combined_nocycle/ARF2017_combined_nocycle). Accessed 15 February 2019.
127. indeXplorer, Lars M. Steinmetz's laboratory, EMBL Heidelberg. Available at: <http://steinmetzlab.embl.de/shiny/indexplorer/?launch=yes>. Accessed 15 February 2019.
128. Single Cell Portal, Aviv Regev's laboratory, Broad Institute, Cambridge. Available at: [https://portals.broadinstitute.org/single\\_cell/study/atlas-of-human-blood-dendritic-cells-and-monocytes](https://portals.broadinstitute.org/single_cell/study/atlas-of-human-blood-dendritic-cells-and-monocytes). Accessed 15 February 2019.
129. Mouse Cell Atlas, Ming Chen and Guoji Guo's Laboratories, Center for Stem Cell and Regenerative Medicine, Hangzhou. Available at: <http://bis.zju.edu.cn/MCA>. Accessed 15 February 2019.
130. Klein Lab tools, Allon Klein's laboratory, Department of Systems Biology, Harvard Medical School. Available at: [https://kleintools.hms.harvard.edu/paper\\_websites/tusi\\_et\\_al/](https://kleintools.hms.harvard.edu/paper_websites/tusi_et_al/). Accessed 15 February 2019.
131. Single-cell RNA-seq of Haematopoietic Stem and Progenitor Cells, Berthold Gottgens' laboratory, Wellcome-MRC Cambridge Stem Cell Institute. Available at: <https://gottgens-lab.stemcells.cam.ac.uk/adultHSPC10X/>. Accessed 15 February 2019.
132. AltAnalyze, Nathan Salomonis' laboratory, Cincinnati Children's Hospital Medical Center, University of Cincinnati. Available at: <http://www.altanalyze.org/ICGS/HCA/splash.php>. Accessed 15 February 2019.
133. Molecular transitions in early progenitors during human cord blood haematopoiesis, Rahul Satija's laboratory, New York Genome Center. Available at: <http://satijalab.org/cd34/>. Accessed 15 February 2019.
134. Muris Consortium, Chan Zuckerberg Biohub. Available at: <https://tabula-muris.ds.czbiohub.org/>. Accessed 15 February 2019.
135. Single Cell Portal, Aviv Regev's laboratory, Broad Institute, Cambridge. Available at: [https://portals.broadinstitute.org/single\\_cell/study/perturb-seq](https://portals.broadinstitute.org/single_cell/study/perturb-seq). Accessed 15 February 2019.