

Essay: Cell lineage Tracing in the Haematopoietic System

Anne van Wijngaarden - S3494969 -
Supervisor: Prof. G. De Haan

Abstract

The classical model of haematopoiesis has already been proposed in the sixties and consists of a stepwise differentiation into the mature myeloid and lymphoid cells. However, due to recent developments in the field of lineage tracing technologies, other models of haematopoiesis have been proposed. Current cell tracing methods and their contribution include barcoding, CRISPR/Cas9 based technologies and sequencing methods. These methods, their advantages and limitations are described as well as what they have taught us in the haematopoietic system. It is proposed that contrary to a stepwise differentiation, that haematopoiesis likely occurs as a continuum out of a heterogeneous pool of HSCs instead of a stepwise differentiation process. Furthermore, it is discussed what method will be the most optimal to fully untangle human haematopoiesis and what future applications and innovations might be expected.

Introduction

Cellular heterogeneity is an emerging feature in many research fields. Its importance has been particularly emphasised in cancer in which can generally be explained by genomic mutations (Kikutake, 2018). However, in healthy tissue and especially in development it is an increasingly important topic, given the fact that all cells of an organism arise from one zygote. This stresses current studies into not only exploring genomic alterations, but transcriptomics and epigenomics as well to elucidate how all these cells differentiate from a single founder cell to their mature state. Therefore, lineage tracing studies have been developed and improved over the years to elucidate these cellular differentiation pathways. This study will focus on the haematopoietic system in which initially it was thought that all cells arise from a homogeneous population of haematopoietic stem cells (HSCs), but recent advances in lineage tracing all point out a high degree of heterogeneity in the progenitor population. This has profound impact on the model of haematopoiesis. Not only does it suggest that there are defined subsets of HSCs, it could further imply that there is no 'unwritten' HSC and all cells are biased to produce their mature cell state. Very subtle changes could thus impact haematopoiesis in health but also in disease. In addition, transplantation is an indispensable method to treat several types of diseases. Currently, patients are treated with HSCs isolated based on cell surface markers and this population is thus composed of a mixture of progenitor cells, even though the patient might require a specific cell type.

Heterogeneity also raises the question how many HSCs are generated and what fraction actually contributes to steady-state haematopoiesis. To solve these questions, this essay will give an overview of what lineage techniques have been developed and what insights they have taught us in this field.

Haematopoiesis

The highly variable landscape of blood cells is derived from a few HSCs which makes it an interesting field of study. Furthermore, it is a very useful system to study *in vivo* since it allows multi-timepoint sampling without having to terminate animals or humans at each timepoint. In addition, the suspension cells are relatively easy to manipulate and can be well studied using Flow Cytometry (Ye, 2017). HSCs ascertain the production and differentiation of all blood and immune cells in a highly complex process. HSCs arise in the haemogenic endothelium, also known as the ventral dorsal aorta (VDA) upon which blood development relies on self-renewal and differentiation of existing clones. Haematopoiesis was classically described as a bifurcating tree in which HSCs self-renew and generate multipotent progenitors (MPP) which commit to Common Lymphoid or Common Myeloid progenitors (CLP, CMP). CLPs subsequently give rise to further committed progenitors which will finally produce lymphoid cells, while CMPs produce progenitors that will result in monocytes and granulocytes and progenitors that will produce megakaryocytes (Mk) and erythrocytes (Ery). Dendritic cells (DC) are, in this model, derived from both CLPs and CMPs (Jagannathan-Bogdan, 2013). However, there is still much subject to discussion due to emerging cell lineage tracing studies which underscore presence and importance of high heterogeneity, even in apparent HSCs (Copley, 2012). An overview of the classical haematopoietic model is shown in Figure 1.

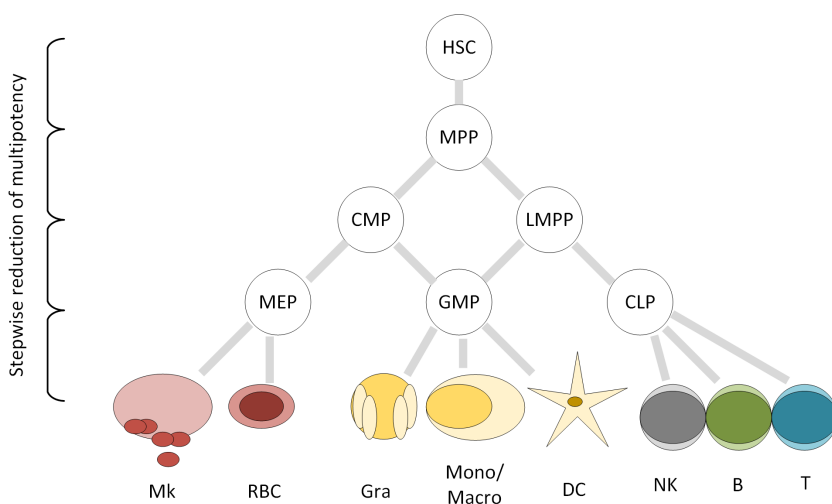


Figure 1: Classical model of haematopoiesis shown as a stepwise differentiation process with one homogenous HSC population at the apex of the tree differentiating in distinct populations.

Lineage Tracing

Cell lineage tracing is a classical tool to study the origin of cells in a heterogeneous population. This provides useful information about from which cell a (sub)population is derived and what changes it has undergone to its current state (Woodworth, 2017; Kretzschmar, 2012). Delineating cell trajectories can not only drive understanding of pathological pathways, but is a fundamental interest of developmental biology as well. The number of lineage tracing methods has exponentially grown and allows cell tracing on larger scales with higher resolution than was previously possible. In very early classical approaches a single founder cell was labeled to trace its progeny which was pioneered in developmental biology by Charles O. Whitman who observed distinct cellular fates of individual cells in leech development and thus realised that cells arise from pre-existing cells, rather than spontaneous generation (Conklin, 1905). Time-Lapse microscopy studies in *C. elegans* further developed cell tracing (Sulston, 1983; Brenner, 1974). Microscopy is noninvasive, but therefore requires an intact animal and cannot be used to directly mark cells. Vital dyes were developed and could be used to directly mark cells. However, such dyes are not in accordance with the essential lineage tracing requirements:

- A clear starting population
- A marker which remains exclusively in the original cells and will not diffuse to neighbouring cells
- Markers should not be toxic
- The stability of markers should be maintained (Woodworth, 2017)

New lineage tracing methods had to be developed and used genetic markers. The markers used were GFP and β -Galactosidase into cells (Chaffle, 1994; Itasaki, 1999; Doetsch, 1999). In fact, the haematopoietic system was one of the earliest applications of lineage tracing by genetic markers (Lemischka, 1986). Even though genetic markers do not diffuse to neighbouring cells, a potential problems include low efficiency of gene introduction and the fact that retroviruses only integrate in dividing cells. Furthermore, spontaneous cell fusion can occur, which can still lead to the transfer of lineage markers providing low-resolution and incomplete assessment of clonality (Kretzschmar, 2012). One method to determine lineage contribution is single cell transplantation. Even though it provides insight into the cell fate of single transplanted cells, it is very time consuming and costly and in addition requires an enormous amounts of mice (Hamilton, 2018) to be representative of the entire HSC population.

Lineage tracing techniques can be divided in prospective and retrospective tracing. In prospective methods, a lineage tracing mark is applied to a single founder cell which is tracked over time to trace its progeny. By contrast, retrospective tracing follows the cell backwards to read endogenous marks which have accumulated over time and is only possible recently due to genome sequencing of single cells. Compared with retrospective lineage tracing, prospective tracing requires less intervention to read the result of lineage tracing, but greater intervention at the onset of development. This potentially alters cellular behaviour which makes retrospective lineage tracing the preferable method (Hsu, 2015; Wu, 2019; Kretzschmar, 2012). In addition, if cells remain in a quiescent state, population expansion might not be sufficient to be detectable. The next section will describe advances in lineage tracing techniques, their application and what these techniques have taught us in haematopoiesis. An overview of prospective tracing methods can be found in Table 1.

Barcoding

Viral barcoding uses a viral plasmid library which consists of vectors containing a random sequence tag or barcode. These barcodes will be chromosomally integrated and hereby introduce an identifiable, unique and heritable mark into the genome of progenitor. Progenitor cells are allowed to develop into their progeny. Different cell types can then be isolated and their genome will be assessed for its barcode by next generation sequencing (Naik, 2013; Merino, 2019; Bystrykh, 2010; Lu, 2011; Schepers, 2008). The method is frequently used, but not all kinships with cell types are assessed. Barcodes could be present in cells which are not harvested and this could impact measurement of engraftment and contribution of progenitors (Thielecke, 2017).

Transposon tagging

Barcoding is already being widely used, but transplanted barcoded cells might not fully recapitulate native stem cell activity. To overcome this, Sun and Rodriguez-Fraticelli used random transposon (Tn) insertions to endogenously mark HSCs in animals. Native lineage relationships were shown by using doxycycline (Dox) inducible Sleeping Beauty (SB) transposase. Dox administration

activates SB expression allowing Tn mobilisations which will then randomly integrate in the genome. Every cell will thus carry a distinct insertion site (Sun, 2014; Rodriguez-Fraticelli, 2018).

Genetic Recombination

To leverage expression of recombinase enzymes in a cell- or tissue specific manner to allow activation of a conditional reporter gene, cell lineage tracing by genetic recombination was developed. Cre-LoxP and FLP-FRT were the first systems to be developed (Carlone, 2016). In the Cre-LoxP system, mice are engineered to express Cre recombinase under control of a chosen promoter and thereby limiting its expression to a specific cell type. Mice are then crossed with a second line in which a reporter transgene is preceded by a loxP-flanked transcriptional STOP sequence. If cells express Cre recombinase, it will excise this sequence allowing expression of the transgene. Cre recombinase can be activated in presence of tamoxifen or anti-progestin which allows the system to be used to determine lineage relationships (Chen, 2018). Initially, only one transgene could be checked, but multicolour reporter lines have gained popularity, including Mosaic Analysis with Double Markers (MADM) (Espinosa, 2014). MADM and other multicolour mouse line systems like Brainbow (Abdeladim, 2019) and Confetti (Amitai-Lange, 2015) allow single cell resolution and precise examination of progenitor division patterns. Yet the limited number of colours remains a restricting factor. Orthogonal systems include Nigrinox (Liu, 2018) and Dre-lox (He, 2018). A big improvement was performed by Pei in Polylox (Pei, 2017,2019). Polylox enables endogenous barcoding based on Cre-LoxP recombination, but reaches a practical diversity of barcodes. Polylox is composed of ten loxP sites which are spaced apart in alternating orientation and can also be spatiotemporally controlled.

Table 1 - Prospective lineage tracing

Technique	Short Description	Prospective/ Retrospective	Main application	Pro	Con
Viral Barcoding	Barcode libraries are virally integrated, providing a heritable mark which can be retrieved and sequenced	Prospective	Mouse, Human (IS)	Easy to apply barcodes, easily coupled with scRNAseq	Lacks spatial information, Limited targetable tissue, ex vivo culture might alter cell behaviour
Genetic Recombination	Mice are engineered to express a Cre Recombinase under a chosen promoter to express a reporter in specific tissue.	Prospective	Mouse	Relatively easy to provide cells with barcodes, allows tissue-specific expression	Only available in mouse, limited by number of colours
Transposon Studies	Dox-Inducible Sleeping Beauty Transposase allows Tn mobilisation and random integrations, providing unique insertion sites	Prospective	Mouse	Relatively easy retrieval of transposon, high resolution	Only available in mouse, lacks spatial information

CRISPR/Cas9 based techniques

The discovery of CRISPR/Cas9 had a huge impact in scientific research in general and is currently applied as lineage tracing method as well. Several similar techniques arose which are slightly different but all rely on the CRISPR/Cas9 system. CRISPR/Cas9 is a bacterial endonuclease which can generate DNA double strand breaks (DSBs) at specific sequences (Marakova, 2011). The break will only be corrected properly in presence of a template for homology-directed repair. If not, DSBs will be repaired by an error-prone process which results in errors at the target site. In CRISPR/Cas9 based techniques, these errors are used as genetic, heritable scars to serve as an internal marker (McKenna, 2016).

GESTALT was the pioneering CRISPR/Cas9 lineage tracing technique and integrates a Cas9 target sequence (± 300 bp) into the host genome (McKenna, 2016). Cas9-mediated scarring is initiated and the barcode can be recovered from tissue by extracting DNA or RNA. Upon processing and aligning sequencing data, an overview of the scarring pattern is obtained. However, the method is restricted to early development and could not identify cell types. Single-cell GESTALT (scGESTALT) combines GESTALT lineage tracing with scRNAseq Raj, 2018). In addition it allows multiple scarring timepoints using a heat-shock inducible Cas9 and is thus not only restricted to embryonic phase.

ScarTrace (Alemany, 2018) and LINNAEUS (Spanjaard, 2018) are based on similar principles. In ScarTrace, sgRNAs are directed to the GFP transgene of a zebrafish line. Scars and the transcriptome are analysed of gDNA and mRNA of which the first is preferred due to potential silencing of GFP (Alemany, 2018).

LINNAEUS was also designed to apply scars in the RFP transgene in the zebrafish line ZebraBow-M. Cas9 and sgRNA for RFP are injected into one-cell-stage embryos to mark cells with genetic scars and loss of RFP fluorescence serves as a direct confirmation of efficient scarring. It uses scRNAseq analysis as barcode and transcriptome readout (Spanjaard, 2018).

Homing CRISPR was the first CRISPR/Cas9 based method to be applied to mouse model MARC1 which contains heritable homing guide RNAs (hgRNAs). A Cas9:hgRNA complex targets the DNA locus of the gRNA itself and allows prolonged scarring, creating higher diversity (Kalhor, 2018). A similar approach is used to track human cells in mice: Mammalian Synthetic Cellular Recorder Integrating Biological Events (mSCRIBE). It consists of a self-targeting guide RNA (stgRNA) which also directs Cas9 to its own locus (Perli, 2016).

Further techniques include COLBERT (Control Of Lineages by Barcode-Enabled Recombinant Transcription), and tags a cell population with a barcode gRNA which is regulated by a specific promoter. Cells are transfected with a transcriptional activator variant of Cas9 carried by a plasmid in addition to a 'Recall' plasmid that encodes the lineage barcode of interest. The technique allows tracing of specific cell lineages, but is currently only used *in vitro* (Al'Khafaji, 2018).

Finally, very recently developed, is CRISPR/Cas9 using nickase Cas9 (nCas9) fused with cytidine deaminase to target the long interspersed nuclear element (LINE1) in the genome. It does, contrary to previously mentioned methods, not require introduction of a barcode and instead uses endogenous repeat elements (Hwang, 2019). An overview of CRISPR/Cas9 based lineage tracing methods is shown in Table 2.

Table 2 - CRISPR/Cas9 based lineage tracing

Technique	Short Description	Prospective/ Retrospective	Application	Pro	Con
CRISPR/ Cas9: GESTALT	Barcodes are applied which are scarred with Cas9-gRNA, scRNAseq allows cell lineage tracing	Prospective	Zebrafish	Allows second round of scarring to track events happening later in life	Currently only applied in zebrafish, delivery at the right developmental stage is a technical challenge
CRISPR/ Cas9: ScarTrace	Cas9 creates DSBs in GFP of Transgene zebrafish	Prospective	Zebrafish	Transcriptomes can identify identification and scar identifies cell's history	Currently only applied in zebrafish, delivery at the right developmental stage is a technical challenge, some sequences are more commonly scarred providing false positive ancestry
CRISPR/ Cas9: LINNAEUS	Cas9 is targeted to RFP of Zebrabow-M	Prospective	Zebrafish	Easier system than ScarTrace and GESTALT	Currently only applied in zebrafish, delivery at the right developmental stage is a technical challenge
CRISPR/ Cas9: Cas9-deamination	nCas9-deaminase targets LINE-1 in the genome and mutation pattern is targeted by sgRNA	Prospective	In Vitro	Does not require exogenous barcodes, higher diversity	Deamination could be toxic, current tests only test marginal number of divisions and have some errors in tree construction
CRISPR/ Cas9: hgRNA	Mutations are introduced at target locus of gRNA itself.	Prospective	Mouse	Creates higher diversity of CRISPR barcodes, Can be combined with in situ readout	hgRNAs have limited duration of evolvability which still limits diversity

Retrospective Lineage Tracing

Somatic mutations accumulate throughout life and are inherited by descendant cells, which allows reconstruction of clonal structures in both development and disease. Somatic mutation burden increases linearly with age and can be used as clonal markers enabling quantification of the number and activity of human blood stem cells during haematopoiesis. Sequencing can be done to detect several alterations: CNVs are generally accessible genetic elements and requires relatively shallow sequencing. Lineage tracing can also be performed on SNVs, LINE1 elements and Microsatellites (MS), which are less present and require increased sequencing depth (Ju, 2018). HSPCs are captured, expanded and subsequently subjected to whole genome sequencing. This allows formation of a general phylogenetic tree upon which targeted deep sequencing is performed during the 'recapture' phase. Deep sequencing provides an organ's clonal architecture and allows reconstruction of somatic cell dynamics. Single cell sequencing has dramatically improved

lineage tracing methods since it can detect cells which would normally be missed in bulk sequencing (Osorio, 2018; Perié, 2016). One limitation is that cells need to be destructed for analysis and subsequent cell fate cannot be followed.

This can now be overcome in a new SIS-sequencing (SISseq) technique. SISseq provides a method in which single cells are allowed to divide and progeny cells are analysed and assayed in SISter conditions; some for cell fate, other by RNA-seq, DNA-seq or epigenome-seq-sequencing (Tian, 2018).

Transcriptome analysis by scRNAseq has the advantage over genome analysis that it can provide information about a cell's state and dynamic gene expression patterns. A basic workflow consists of single cell isolation, lysis and reverse transcription for cDNA generation, which can be subsequently amplified and sequenced. scRNAseq transcriptome analysis provides valuable information about transitional states, not only at start or end, but also in intermediate phases. This is very useful in haematopoietic studies in which it becomes more apparent that instead of a stable population of HSCs, the cells reside in a continuum and therefore this transitory state is an important aspect to conduct research on. scRNAseq by itself could be used for lineage tracing, but the technique only a snapshot of a cell's state only and is therefore preferentially combined with other tracing techniques. RNA velocity can be used as a high-dimensional vector that predicts future state of individual cells. RNA velocity predicts a cell's future state by quantification of spliced and unspliced transcripts (McKenna, 2019; Kester, 2018)

Another sequencing method is Transposon Intertion Profiling by sequencing (TIPseq). This method uses vectorette PCR to amplify species-specific LINE1 insertion sites followed by paired-end Illumina sequencing. LINE1 is with 17% the most abundant mobile DNA in the human genome. Most of it exists in a fixed state, but approximately 500 insertions of LINE1 are highly variable. LINE1 activity leads to transposable element insertions that are a source of variation in genomes, but characterisation is highly challenging because of their repetitive nature (Steranka, 2019).

In addition, instead of genomic mutations, somatic mutations in mitochondrial DNA (mtDNA) can be tracked and analysed by single-cell RNA sequencing or assay for transposase accessible chromatin (ATAC) sequencing. It is hypothesised that mtDNA sequence variation potentially provides a natural barcode from which to derive clonal relationships. mtDNA has the advantage that the 16.6kb-long genome is sufficiently small for cost-effective sequencing, yet suffices as a target for genetic diversity. mtDNA mutation rate is estimated to be much higher than for genomic DNA and a major advantage is that mtDNAseq is immediately combined with transcriptome sequencing since the mitochondrial genome is caught as an unwanted by-product in ATAC-seq (Ludwig, 2019). mtDNAseq enlarges available tracing methods by the possibility of combining clonal tracking data with transcriptomic and epigenetic data.

mtDNA, TIPseq and SISseq have been developed too recently to currently provide insights in the fraction of HSCs contributing to steady-state haematopoiesis and will therefore not be used in the following section. mtDNA studies have only focused on elucidating differences in HSCs and pre-leukemic HSCs (pHSCs), but future studies could point out valuable information due to its great advantages (Ludwig, 2019; Tian, 2018). TIPseq was only applied to DC development due to the ability to generate them in vitro. The study showed that even though cell fate of all cells from a population was highly heterogeneous, fate of sister cells was often similar (Steranka, 2019).

Table 3 - Retrospective Lineage Tracing

Technique	Short Description	Prospective/ Retrospective	Application	Pro	Con
Sequencing: Somatic Mutations	Somatic mutations increase with age and provides a natural barcode to trace origin	Retrospective	Human	Application to human blood samples and does not require any intervention	High costs and requires improved bioinformatic processing, high error rates in amplification, lacks spatial information
Sequencing: mtDNA	mtDNA is a by-product of ATAC-seq and provides a more sensitive natural barcode	Retrospective	Human	Small genome for cost effective sequencing and genome is automatically captured using scATACseq, mtGenomes have a high copy number, applicable to human tissue	Horizontal Transfer of mitochondria between cells, unable to account for phenotypic effects of mtDNA mutations
Sequencing: SISseq	SISter cells are allowed minimal divisions and each SISter is subjected to a different assay (cell fate, RNA-seq)	Retrospective	In vitro DCs	Connects cell fate with gene programs which allows early gene expression signature,	Not available in vivo context yet
Sequencing: TIPseq	Uses vectorette PCR to amplify species-specific LINE-1 inception sites	Retrospective	Human (cancer biopsies)	Relies on naturally occurring LINE1 positions	Restriction digestion requires large amount of high quality DNA, not applicable to scDNA, does not differentiate between insertion types
Sequencing: scRNAseq	Reveal gene expression profiles of both steady and transitioning cell state	Retrospective	Human samples, in vitro cultures	Can distinguish between expression profiles of both steady and transitioning cells, detects infrequently-represented transcripts which would be missed in bulk-analyses	Gene expression constantly changes and one 'snapshot' is not sufficient to trace cells.

Lineage tracing; What have we learned?

As described in the previous section, many lineage tracing methods have been developed both in prospective as well as retrospective contexts. All of these have their advantages, limitations and applicable models, which are summarised above. Here we focus on what insights these techniques have brought us in the field of haematopoiesis (Woodworth, 2018; Grun, 2016; Kretzschmar, 2012; Hsu, 2016).

The classical haematopoietic model has been largely discussed due to constant new discoveries provided by lineage tracing studies. Some studies hypothesise a quiescent population of HSCs, whereas others assume that quite a large population of HSCs actively contribute to haematopoiesis by self-renewal and stepwise progression to committed progenitors. Lineage tracing reveals not only heterogeneity at the apex of the haematopoietic tree, but new developmental branches as well.

Lineage tracing reveals additional branches to the classical haematopoietic tree

In the classical model (Fig.1), it was assumed that branches arise from equal divisions and that cells need to pass certain increasingly differentiated progenitor stages to reach their mature state. However, lineage tracing studies show not only additional branches sprouting at the apex of the haematopoietic tree, but also priming of cells in the HSC population.

In scRNAseq studies, it was shown that Mk cells can arise directly from HSCs, instead of progressing through the MEP-stage. A distinct HSC population depicted a high expression of the Von Willebrand Factor (Vwf) which is associated with Mks (Sanjuan, 2013; Wilson, 2015). In transposon studies, it was shown that at least a fraction of LT-HSCs behaved as Mk-progenitors again confirming early lineage priming and an additional Mk branch (Rodriguez-Fraticelli, 2018).

In addition, barcoding experiments have shown that in LMPPs, the majority of cells gave rise to DCs only, which led the authors to hypothesise that there must be a direct branch leading to these cells (Naik, 2013; Perié, 2016). To further confirm early priming, LMPPs were barcoded and combined with splitting transplantation. Sister cells often showed conserved fate and were thus hypothesised to be programmed for a certain cell fate high in the haematopoietic tree.

In another study in which further research was performed on cell fate of LMPP studies, it was found that LMPPs could only generate neutrophils via GMPs, whereas CMPs could only generate eosinophils and basophils through a common EoBP (Gorgens, 2013). Furthermore, IS barcoding experiments in macaques on CD34+ cells showed an additional separate branch to NKs since the cells did not share any barcodes with either lymphocytes and myeloid cells. Proposed additional branches are shown in Figure 2.

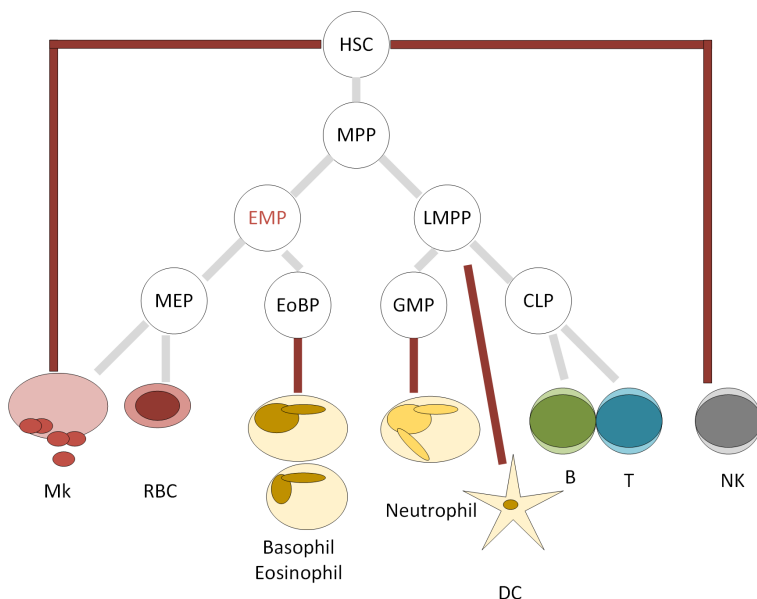


Figure 2: Additional branches added to the haematopoietic tree due to advances in lineage tracing

Lineage tracing reveals heterogeneity and lineage priming in HSC population

In addition to additional branches in the haematopoietic tree, lineage tracing studies showed high HSC heterogeneity. Initially it was thought that the HSC population was quite homogeneous and consisted only of quiescent LT-HSCs and proliferating and differentiating ST-HSCs. This model was already extended by addition of multiple MPP populations (MPP1,2,3,4). Transcriptome comparison of FACS sorted cells, which were GFP-labeled and transplanted, showed that even these populations have different potencies. Whereas MPP2 cells are suggested multipotent, MPP3 and MPP4 already show a myeloid lineage bias, again adding to heterogeneity. Yet even this model is not sufficient to depict all heterogeneity. Barcoding experiments showed that some HSCs do actually give rise to all cell types analysed, but the majority of HSCs are biased and committed already in HSC state. The authors did show that there is early branching in the classical haematopoietic tree to CMPs and CLPs, but interestingly, demonstrated that when individual CMPs are transplanted, they yield either only erythrocytes or only myeloid cells. Further analysis showed that within lymphoid biased cells, a B- or T-cell bias is present as well.

Transposon studies provided further evidence of biased progenitor cells and HSC heterogeneity. In one study Sun found, by screening for granulocyte transposon tags, that the majority of cells had unique tags, suggesting an extreme polyclonal nature of granulopoiesis. In addition, Rodriguez-Fraticelli used the same method, but revealed Myeloid progenitors (MyPs, consisting of CMP, GMP and MEPs), actually had unilineage outcomes, which is in accordance with single cell expression profiling studies and further validates the hypothesis that progenitor populations actually consist of a mixture of lineage restricted cells. When they aligned patterns of MPPs and mature progeny, there was a striking overlap, suggesting MPPs are indeed already committed (Rodriguez-Fraticelli, 2018).

Somatic mutations allow construction of a developmental lineage tree which can even reveal pre-natal mutation rates. Authors show a paucity of shared mutations and therefore a highly polyclonal hierarchy of haematopoietic cells. They were able to construct a lineage tree which revealed an asymmetric contribution of developmental branches to haematopoiesis. This asymmetry was more prevalent compared to MSCs. Within the branches, there was biased contribution to the pools of HSCs and MPPs (Osorio, 2018).

All these studies opt for a haematopoietic model which captures this level of heterogeneity. Velten proposed a model consisting of a Continuum of Low Primed Undifferentiated Haematopoietic Stem and Progenitor Cells (CLOUD-HSCPs) by combining flow cytometric, transcriptomic and functional data at single cell level. He labeled cells with surface markers and used FACS to characterise human HPSCs (absence of lineage markers, expression of CD34). HPSCs were subjected to RNA-Seq or cultured *ex vivo* to quantify Mk, Ery and Mye lineage potential. All data was combined and interestingly, authors found a lack of hierarchical structures in the progenitor compartment due to gradual differences between cells. This was incompatible with the classical model which led them to propose the CLOUD-HSPC model. When progenitors were studied further in depth, it was shown that these populations actually consisted of cell types with unilineage-specific expression profiles and functional unipotency. In addition, this unilineage-specific expression is linked to the restriction of lineage potential in early stages *in vitro* and *vivo* (Velten, 2017; Scala, 2019).

It was then interesting to question what processes underly HSC commitment and authors characterised expression sets associated with specific primed cellular states. Interestingly, even in stem cells, they could already find gene expression from earliest priming modules of all lineages and thus priming exists already in most primitive HSCs.

A similar model was also proposed by combining ATAC-seq with scRNAseq on HSPC subsets to study changes in transcription factor expression and link these to changes in chromatin accessibility. The study shows that haematopoietic differentiation takes place in a broad pool of allowable HSC states (Yabe, 2018).

Lineage tracing to determine HSC number and contribution to haematopoiesis

The current haematopoietic model consists of a broad pool of heterogeneous HSPC, which leads to the question of how many HSCs are present and what fraction does actually contribute to haematopoiesis?

Limiting dilution has been the standard to quantify HSC number. Previous methods to analyse number of HSCs were dependent FACS analysis by labeling cells and determining percentage of HSC in blood or BM samples, but is complicated by the constantly changing panel of cell-surface markers to distinguish HSCs (Ashley, 2017; Georgolopoulos, 2019). Other approaches use the changing ratio with age of maternal/paternal X-chromosome phenotypes from blood cells of females. In a comparable way, shortening of telomeres is used to quantify number and HSC activity. Both studies estimate replication rate of HSCs is ± 1 per year and based on this calculation, it was proposed that ± 1275 HSC derived clones actively contribute to human haematopoiesis. These findings were in line with clonal tracking studies which relied on vector integration sites to mark and monitor dynamics of haematopoietic reconstitution (Caitlin, 2011; Werner, 2015).

Several barcoding experiments have attempted to define stem cell number and their contribution to haematopoiesis. One study revealed that actually the majority of HSC clones contributes to haematopoiesis (Bystrykh, 2013). In another barcoding study, the authors showed that HSCs do not equally contribute to blood cells, validating the current model of haematopoiesis, but the study lacks a real quantification of HSC contribution (Lu, 2011). In a barcoding study of Umbilical Cord Blood (UCB) cells, one of the aims was to determine the number of HSCs actively producing clonal offspring. Human UCB cells (CD34+) were barcoded and transplanted into mice and allowed engraftment and differentiation. Cells were harvested and subjected to deep sequencing. Authors found already high heterogeneity in contribution to specific lineage between donors between UCB donor cells. When studying the number of HSCs contributing to long-term progeny production, they estimated that approximately 0.007% of CD34+ cells took part in this process and that this number declined over time (Belderbos, 2019). However, this number could vary up to 10-fold and interestingly, a higher number of clones did not associate with faster engraftment. Transposon studies (Sun, Rodriguez-Fraticelli) were confirmed by Biasco, who provided a tracking model of Haematopoietic reconstitution after HSPC transplantation in humans by integration Site (IS) barcoding (Biasco, 2016). This allowed tracking of cells up to four years to study clonal composition. Biasco showed that only 15 out of 51 HSCs were detected in at least one other lineage, whereas 9 out of 10 MPPs were found back in mature cells and thus play a more profound role in haematopoiesis than HSCs.

PolyLox barcoding by Pei attempted to quantify HSC number as well. He extrapolated clone sizes to whole adult population which would consist of ± 15.000 HSCs. Furthermore, using this method, they revealed contrary to some barcoding studies that HSCs have a very small contribution to steady-state haematopoiesis (Pei, 2017, 2019). This is further supported in a study in which $>90\%$ of HPSCs were ablated which did not affect the rate of steady-state haematopoiesis. In addition, LT-HSCs recovered from $<1\%$ to a maximum of 10% whereas progenitor cells quickly and substantially recovered after depletion (Schoedel, 2016).

This number was lower than the number proposed by somatic mutation analysis by Lee-Six (Lee-Six, 2018), who used a hybridisation bait-set to identify mutations in colonies derived from HSCs. Targeted sequencing on these colonies allows formation of a phylogenetic tree revealing a paucity of recent branch points. This suggests that phylogeny is dominated by events occurring in stem cells. To really calculate how many HSCs contribute to haematopoiesis, they used a Bayesian Computational Framework in which many haematopoietic simulations were generated varying the number of stem cells. Each simulation was compared to in vivo data and the highest credibility was found in the range of 44.000-215.000 cells. However, this method highly relies on bioinformatic modelling and might not truly represent in vivo data (Lee-Six, 2018).

A color-labeling study, the confetti mouse model is used and even though the paper mainly focuses on alterations in clonal complexity upon ageing, it does confirm Lee-Six' estimates that 50.000-200.000 contribute to blood formation at any given time (Ganuza, 2019).

Discussion

This essay has focused on describing which lineage tracing methods have been developed over the years and what they have taught us in the field of haematopoiesis. HSC heterogeneity has become very important as well as distinct biased progenitor states serving as main contributors of haematopoiesis. Yet a fully detailed model is still lacking. In addition, there are studies which reason against this revised model. Sawai and colleagues label the *pdzk1ip1* gene which is profoundly expressed in HSCs and is reduced upon differentiation. They do show that *Pdzk1ip1* labeled HSCs contribute to all haematopoietic lineages and thereby provide evidence that HSCs and not MPPs contribute to steady-state haematopoiesis. However, this method does not precisely delineate cellular pathways leading to mature cells and the amount of information obtained from tracing one gene is in stark contrast with all the data acquired from transcriptomes and epigenomes which provide much earlier markers of lineage bias. In addition, *pdzk1ip1* is only expressed in mice and their model is thus still far from implications in human haematopoiesis (Sawai, 2016). Chapple et al had a similar approach by tracing *Krt18* and *Fgd5* in a recombinase assay (*Krt18-CreER*, *Fgd5-CreER*) (Chapple, 2018). Both genes are known to be enriched in HSCs with similar expression patterns as *Pdzk1ip1* and *Hoxb5*. Transgenic, tamoxifen induced mice were followed up for 1 year to determine HSC contribution and showed that HSCs have a robust contribution to haematopoiesis. Interestingly, the study shows a higher contribution to the myeloid lineage and platelet production in particular. This is consistent with previous reports, but does not fully rule out the possibility of lineage biased progenitor cells (Sawai, 2016; Busch, 2015). Even though these studies do suggest HSC contribution to steady-state haematopoiesis, their studies are just based on a few genes associated with HSC state and do not provide lineage relationships. Many (FACS) studies rely on set haematopoietic subsets but do not address detailed genomic and transcriptomic information. Therefore, these studies do not rule out the model of a heterogeneous progenitor population.

A model in which biased progenitors are the main contributors to haematopoiesis are despite some conflicting studies broadly accepted. There are currently some estimates in number of HSCs contributing to steady-state haematopoiesis, but with the exception studies by Lee-Six' and Ganuza rely on data derived from mice and zebrafish studies. The optimal study described in this review would therefore be somatic mutation sequencing, since it allows studying unperturbed haematopoiesis in humans. Even though the mutational landscape provides a detailed description of cellular origin, somatic mutations are quite rare and appropriate sequencing depth is required for sufficient resolution (Woordworth, 2018). In addition, many errors can occur during amplification, relies on single-cell sequencing causing reduced throughput and is more expensive. Last, it becomes more evident that even though somatic mutations may provide a detailed lineage tree, it is not sufficient to study initial events in haematopoiesis. It could be that there are even earlier cellular events prior to transcriptional heterogeneity that could predict a cell's destination. In other words: How is the fate of lineage-biased HSCs predetermined? To elucidate this, more research should be done in epigenomes of early haematopoietic cells. Therefore mtATAC-seq would be a valuable tool (Ludwig, 2019). In addition, combinatory studies are currently in development and will provide high resolution lineage profiling in multiple dimensions. smFISH would provide spatial information, which is currently lost in many studies. However, with advances in multidimensional single cell studies, bioinformatic modelling of these large amounts of data should be improved as well. Many groups design their own bioinformatic tools to process sequencing data, which could lead to inconsistencies in output. A central bioinformatic tool should be used to circumvent this problem (Cannoodt, 2016).

If the underlying mechanisms in haematopoiesis are elucidated, it will aid to provide improved therapy in patients requiring transplantation. If specific cells are required, these could be more precisely transplanted for better engraftment. In addition, sequencing studies could help create mutational profiles of development and disease, which could potentially be used as early diagnostic markers in cancer. Some dominant clones could outgrow healthy cells and by studying their mutational profile give better information about pathology and mechanisms of resistance. In addition it would be interesting to study how HSCs react differ in function during acute and chronic inflammation and differ between healthy homeostasis and emergency haematopoiesis. Overall, lineage tracing techniques have provided us with great insights in haematopoiesis, but do require increased sequencing depth and improved bioinformatic analysis for optimal resolution.

References

1. Abdeladim, L., Matho, K. S., Clavreul, S., Mahou, P., Sintes, J. M., Solinas, X., et al. (2019). Multicolor multiscale brain imaging with chromatic multiphoton serial microscopy. *Nature Communications*, 10(1), 1662-019-09552-9.
2. Alemany, A., Florescu, M., Baron, C. S., Peterson-Maduro, J., & van Oudenaarden, A. (2018). Whole-organism clone tracing using single-cell sequencing. *Nature*, 556(7699), 108-112.
3. Al'Khafaji, A. M., Deatherage, D., & Brock, A. (2018). Control of lineage-specific gene expression by functionalized gRNA barcodes. *ACS Synthetic Biology*, 7(10), 2468-2474.
4. Amitai-Lange, A., Berkowitz, E., Altshuler, A., Dbayat, N., Nasser, W., Suss-Toby, E., et al. (2015). A method for lineage tracing of corneal cells using multi-color fluorescent reporter mice. *Journal of Visualized Experiments : JoVE*, (106):e53370. doi(106), e53370.
5. Belderbos, M. E., Jacobs, S., Koster, T. K., Ausema, A., Weersing, E., Zwart, E., et al. (2019). Donor-to-donor heterogeneity in the clonal dynamics of transplanted HumanCord blood stem cells in murine xenografts. *Biology of Blood and Marrow Transplantation : Journal of the American Society for Blood and Marrow Transplantation*,
6. Biasco, L., Pellin, D., Scala, S., Dionisio, F., Basso-Ricci, L., Leonardelli, L., et al. (2016). In vivo tracking of human hematopoiesis reveals patterns of clonal dynamics during early and steady-state reconstitution phases. *Cell Stem Cell*, 19(1), 107-119.
7. Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics*, 77(1), 71-94.
8. Busch, K., Klapproth, K., Barile, M., Flossdorf, M., Holland-Letz, T., Schlenner, S. M., et al. (2015). Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature*, 518(7540), 542-546.
9. Cannoodt, R., Saelens, W., & Saeys, Y. (2016). Computational methods for trajectory inference from single-cell transcriptomics. *European Journal of Immunology*, 46(11), 2496-2506.
10. Carlone, D. L. (2016). Identifying adult stem cells using cre-mediated lineage tracing. *Current Protocols in Stem Cell Biology*, 36, 5A.2.1-18.
11. Catlin, S. N., Busque, L., Gale, R. E., Guttorp, P., & Abkowitz, J. L. (2011). The replication rate of human hematopoietic stem cells in vivo. *Blood*, 117(17), 4460-4466.
12. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., & Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science (New York, N.Y.)*, 263(5148), 802-805.
13. Chapple, R. H., Tseng, Y. J., Hu, T., Kitano, A., Takeichi, M., Hoegenauer, K. A., et al. (2018). Lineage tracing of murine adult hematopoietic stem cells reveals active contribution to steady-state hematopoiesis. *Blood Advances*, 2(11), 1220-1228.
14. Chen, Z., Ren, X., Xu, X., Zhang, X., Hui, Y., Liu, Z., et al. (2018). Genetic engineering of human embryonic stem cells for precise cell fate tracing during human lineage development. *Stem Cell Reports*, 11(5), 1257-1271.
15. Conklin, E. G. (1905). The organization and cell-lineage of the ascidian egg (Vol. 13).
16. Copley, M. R., Beer, P. A., & Eaves, C. J. (2012). Hematopoietic stem cell heterogeneity takes center stage. *Cell Stem Cell*, 10(6), 690-697.
17. Doetsch, F., Caille, I., Lim, D. A., Garcia-Verdugo, J. M., & Alvarez-Buylla, A. (1999). Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. *Cell*, 97(6), 703-716.
18. Espinosa, J. S., Tea, J. S., & Luo, L. (2014). Mosaic analysis with double markers (MADM) in mice. *Cold Spring Harbor Protocols*, 2014(2), 182-189.
19. Ganuza, M., Hall, T., Finkelstein, D., Wang, Y. D., Chabot, A., Kang, G., et al. (2019). The global clonal complexity of the murine blood system declines throughout life and after serial transplantation. *Blood*, 133(18), 1927-1942.
20. Georgolopoulos, G., Iwata, M., Psatha, N., Yiangou, M., & Vierstra, J. (2019). Unbiased phenotypic identification of functionally distinct hematopoietic progenitors. *Journal of Biological Research (Thessalonike, Greece)*, 26, 4-019-0097-7. eCollection 2019 Dec.
21. Georgolopoulos, G., Iwata, M., Psatha, N., Yiangou, M., & Vierstra, J. (2019). Unbiased phenotypic identification of functionally distinct hematopoietic progenitors. *Journal of Biological Research (Thessalonike, Greece)*, 26, 4-019-0097-7. eCollection 2019 Dec.
22. Gerrits, A., Dykstra, B., Kalmykova, O. J., Klauke, K., Verovskaya, E., Broekhuis, M. J., et al. (2010). Cellular barcoding tool for clonal analysis in the hematopoietic system. *Blood*, 115(13), 2610-2618.
23. Gorgens, A., Radtke, S., Mollmann, M., Cross, M., Durig, J., Horn, P. A., et al. (2013). Revision of the human hematopoietic tree: Granulocyte subtypes derive from distinct hematopoietic lineages. *Cell Reports*, 3(5), 1539-1552.
24. Grun, D., Muraro, M. J., Boisset, J. C., Wiebrands, K., Lyubimova, A., Dharmadhikari, G., et al. (2016). De novo prediction of stem cell identity using single-cell transcriptome data. *Cell Stem Cell*, 19(2), 266-277.
25. Hamilton, N., Sabroe, I., & Renshaw, S. A. (2018). A method for transplantation of human HSCs into zebrafish, to replace human-

- ised murine transplantation models. *F1000research*, 7, 594.
26. He, L., Li, Y., Huang, X., Li, Y., Pu, W., Tian, X., et al. (2018). Genetic lineage tracing of resident stem cells by DealT. *Nature Protocols*, 13(10), 2217-2246.
 27. Hsu, Y. C. (2015). Theory and practice of lineage tracing. *Stem Cells (Dayton, Ohio)*, 33(11), 3197-3204.
 28. Hwang, B., Lee, W., Yum, S. Y., Jeon, Y., Cho, N., Jang, G., et al. (2019). Lineage tracing using a Cas9-deaminase barcoding system targeting endogenous L1 elements. *Nature Communications*, 10(1), 1234-019-09203-z.
 29. Itasaki, N., Bel-Vialar, S., & Krumlauf, R. (1999). 'Shocking' developments in chick embryology: Electroporation and in ovo gene expression. *Nature Cell Biology*, 1(8), E203-7.
 30. Jagannathan-Bogdan, M., & Zon, L. I. (2013). Hematopoiesis. *Development (Cambridge, England)*, 140(12), 2463-2467.
 31. Ju, Y. S., Martincorena, I., Gerstung, M., Petljak, M., Alexandrov, L. B., Rahbari, R., et al. (2017). Somatic mutations reveal asymmetric cellular dynamics in the early human embryo. *Nature*, 543(7647), 714-718.
 32. Kalhor, R., Kalhor, K., Mejia, L., Leeper, K., Graveline, A., Mali, P., et al. (2018). Developmental barcoding of whole mouse via homing CRISPR. *Science (New York, N.Y.)*, 361(6405), 10.1126/science.aat9804. Epub 2018 Aug 9.
 33. Kester, L., & van Oudenaarden, A. (2018). Single-cell transcriptomics meets lineage tracing. *Cell Stem Cell*, 23(2), 166-179.
 34. Kikutake, C., Yoshihara, M., Sato, T., Saito, D., & Suyama, M. (2018). Pan-cancer analysis of intratumor heterogeneity associated with patient prognosis using multidimensional measures. *Oncotarget*, 9(102), 37689-37699.
 35. Kretzschmar, K., & Watt, F. M. (2012). Lineage tracing. *Cell*, 148(1-2), 33-45.
 36. Lee-Six, H., Obro, N. F., Shepherd, M. S., Grossmann, S., Dawson, K., Belmonte, M., et al. (2018). Population dynamics of normal human blood inferred from somatic mutations. *Nature*, 561(7724), 473-478.
 37. Lemischka, I. R., Raulet, D. H., & Mulligan, R. C. (1986). Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell*, 45(6), 917-927.
 38. Liu, K., Yu, W., Tang, M., Tang, J., Liu, X., Liu, Q., et al. (2018). A dual genetic tracing system identifies diverse and dynamic origins of cardiac valve mesenchyme. *Development (Cambridge, England)*, 145(18), 10.1242/dev.167775.
 39. Lu, R., Neff, N. F., Quake, S. R., & Weissman, I. L. (2011). Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. *Nature Biotechnology*, 29(10), 928-933.
 40. Lu, R., Neff, N. F., Quake, S. R., & Weissman, I. L. (2011). Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. *Nature Biotechnology*, 29(10), 928-933.
 41. Ludwig, L. S., Lareau, C. A., Ulirsch, J. C., Christian, E., Muus, C., Li, L. H., et al. (2019). Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. *Cell*, 176(6), 1325-1339.e22.
 42. Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., et al. (2011). Evolution and classification of the CRISPR-cas systems. *Nature Reviews Microbiology*, 9(6), 467-477.
 43. McKenna, A., Findlay, G. M., Gagnon, J. A., Horwitz, M. S., Schier, A. F., & Shendure, J. (2016). Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science (New York, N.Y.)*, 353(6298), aaf7907.
 44. McKenna, A., & Gagnon, J. A. (2019). Recording development with single cell dynamic lineage tracing. *Development (Cambridge, England)*, 146(12), 10.1242/dev.169730.
 45. Merino, D., Weber, T. S., Serrano, A., Vaillant, F., Liu, K., Pal, B., et al. (2019). Barcoding reveals complex clonal behavior in patient-derived xenografts of metastatic triple negative breast cancer. *Nature Communications*, 10(1), 766-019-08595-2.
 46. Naik, S. H., Perie, L., Swart, E., Gerlach, C., van Rooij, N., de Boer, R. J., et al. (2013). Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature*, 496(7444), 229-232.
 47. Ng, A. P., & Alexander, W. S. (2017). Haematopoietic stem cells: Past, present and future. *Cell Death Discovery*, 3, 17002.
 48. Osorio, F. G., Rosendahl Huber, A., Oka, R., Verheul, M., Patel, S. H., Hasaart, K., et al. (2018). Somatic mutations reveal lineage relationships and age-related mutagenesis in human hematopoiesis. *Cell Reports*, 25(9), 2308-2316.e4.
 49. Pei, W., Feyerabend, T. B., Rossler, J., Wang, X., Postrach, D., Busch, K., et al. (2017). Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature*, 548(7668), 456-460.
 50. Pei, W., Wang, X., Rossler, J., Feyerabend, T. B., Hofer, T., & Rodewald, H. R. (2019). Using cre-recombinase-driven polylox barcoding for in vivo fate mapping in mice. *Nature Protocols*, 14(6), 1820-1840.
 51. Perie, L., & Duffy, K. R. (2016). Retracing the in vivo haematopoietic tree using single-cell methods. *FEBS Letters*, 590(22), 4068-4083.
 52. Perli, S. D., Cui, C. H., & Lu, T. K. (2016). Continuous genetic recording with self-targeting CRISPR-cas in human cells. *Science (New York, N.Y.)*, 353(6304), 10.1126/science.aag0511. Epub 2016 Aug 18.
 53. Raj, B., Wagner, D. E., McKenna, A., Pandey, S., Klein, A. M., Shendure, J., et al. (2018). Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. *Nature Biotechnology*, 36(5), 442-450.

54. Rodriguez-Fraticelli, A. E., Wolock, S. L., Weinreb, C. S., Panero, R., Patel, S. H., Jankovic, M., et al. (2018). Clonal analysis of lineage fate in native haematopoiesis. *Nature*, 553(7687), 212-216.
55. Sanjuan-Pla, A., Macaulay, I. C., Jensen, C. T., Woll, P. S., Luis, T. C., Mead, A., et al. (2013). Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature*, 502(7470), 232-236.
56. Sawai, C. M., Babovic, S., Upadhaya, S., Knapp, D. J. H. F., Lavin, Y., Lau, C. M., et al. (2016). Hematopoietic stem cells are the major source of multilineage hematopoiesis in adult animals. *Immunity*, 45(3), 597-609.
57. Schepers, K., Swart, E., van Heijst, J. W., Gerlach, C., Castrucci, M., Sie, D., et al. (2008). Dissecting T cell lineage relationships by cellular barcoding. *The Journal of Experimental Medicine*, 205(10), 2309-2318.
58. Schoedel, K. B., Morcos, M. N. F., Zerjatke, T., Roeder, I., Grinenko, T., Voehringer, D., et al. (2016). The bulk of the hematopoietic stem cell population is dispensable for murine steady-state and stress hematopoiesis. *Blood*, 128(19), 2285-2296.
59. Spanjaard, B., Hu, B., Mitic, N., Olivares-Chauvet, P., Janjuha, S., Ninov, N., et al. (2018). Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. *Nature Biotechnology*, 36(5), 469-473.
60. Steranka, J. P., Tang, Z., Grivainis, M., Huang, C. R. L., Payer, L. M., Rego, F. O. R., et al. (2019). Transposon insertion profiling by sequencing (TIPseq) for mapping LINE-1 insertions in the human genome. *Mobile DNA*, 10, 8-019-0148-5. eCollection 2019.
61. Sulston, J. E., Schierenberg, E., White, J. G., & Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Developmental biology*, 100(1), 64-119.
62. Sun, J., Ramos, A., Chapman, B., Johnnidis, J. B., Le, L., Ho, Y. J., et al. (2014). Clonal dynamics of native haematopoiesis. *Nature*, 514(7522), 322-327.
63. Thielecke, L., Aranyosy, T., Dahl, A., Tiwari, R., Roeder, I., Geiger, H., et al. (2017). Limitations and challenges of genetic barcode quantification. *Scientific Reports*, 7, 43249.
64. Werner, B., Beier, F., Hummel, S., Balabanov, S., Lassay, L., Orlikowsky, T., et al. (2015). Reconstructing the in vivo dynamics of hematopoietic stem cells from telomere length distributions. *Elife*, 4, 10.7554/eLife.08687.
65. Wilson, N. K., Kent, D. G., Buettner, F., Shehata, M., Macaulay, I. C., Calero-Nieto, F. J., et al. (2015). Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. *Cell Stem Cell*, 16(6), 712-724.
66. Wu, S. S., Lee, J. H., & Koo, B. K. (2019). Lineage tracing: Computational reconstruction goes beyond the limit of imaging. *Molecules and Cells*, 42(2), 104-112.
67. Yabe, I. M., Truitt, L. L., Espinoza, D. A., Wu, C., Koelle, S., Panch, S., et al. (2018). Barcoding of macaque hematopoietic stem and progenitor cells: A robust platform to assess vector genotoxicity. *Molecular Therapy. Methods & Clinical Development*, 11, 143-154.
68. Ye, F., Huang, W., & Guo, G. (2017). Studying hematopoiesis using single-cell technologies. *Journal of Hematology & Oncology*, 10(1), 27-017-0401-7.