Deng *et al. Journal of Translational Medicine* (2025) 23:270
https://doi.org/10.1186/s12967-025-06318-4

Journal of Translational Medicine

REVIEW Open Access

Single-cell lineage tracing techniques in hematology: unraveling the cellular narrative

Lu-Han Deng[1], Mu-Zi Li[1], Xiao-Jun Huang[1] and Xiang-Yu Zhao[1*]

Abstract
Lineage tracing is a valuable technique that has greatly facilitated the exploration of cell origins and behavior. With the continuous development of single-cell sequencing technology, lineage tracing technology based on the single-cell level has become an important method to study biological development. Single-cell Lineage tracing technology plays an important role in the hematological system. It can help to answer many important questions, such as the heterogeneity of hematopoietic stem cell function and structure, and the heterogeneity of malignant tumor cells in the hematological system. Many studies have been conducted to explore the field of hematology by applying this technology. This review focuses on the superiority of the emerging single-cell lineage tracing technologies of Integration barcodes, CRISPR barcoding, and base editors, and summarizes their applications in the hematology system. These studies have suggested the vast potential in unraveling complex cellular behaviors and lineage dynamics in both normal and pathological contexts.

Keywords Lineage tracing, Single-cell lineage tracing, Barcodes, CRISPR/Cas9, Hematology

Background

Lineage tracing is a technique that can track all descendants of a single progenitor cell and reveal their fate trajectories. Lineage tracing can provide biological characteristics of progenitor cell progeny, such as cell expression profile characteristics, differentiation status, etc. It provides a powerful tool for cell fate mapping and is recognized as the primary method for revealing the complex connections between progenitor cells and their progeny. Originally prominent in classical developmental biology, lineage tracing is now expanding to stem cell research and modeling of cellular diversity in cancer [1, 2]. It provides a unique perspective for studying cellular behavior within intact tissues or organisms. Single-cell sequencing technology allows us to resolve cellular states at the single-cell resolution level. Lineage tracing, the gold standard for cellular trajectory inference, combined with single-cell sequencing, enables single-cell lineage tracing (SCLT) to map cell lineage connectivity at single-cell resolution and is the best tool for exploring the heterogeneity of cellular differentiation [3]. Hematopoietic stem cells (HSCs) are a continuous source of hematopoietic and immune cells, and their clonal behavior and subclonal diversity are closely related to blood disorders, cancer treatments, and the aging process, and thus are important for advancing regenerative medicine and precision therapies. The development of SCLT technology provides a great opportunity to gain insights into how these stem cells can contribute to the hematopoietic process in both

*Correspondence:
Xiang-Yu Zhao
zhao_xy@bjmu.edu.cn
[1]Peking University People's Hospital, Peking University Institute of Hematology, Beijing Key Laboratory of Hematopoietic Stem Cell Transplantation, National Clinical Research Center for Hematologic Disease, No. 11 Xizhimen South Street, Beijing 100044, China

healthy and diseased states. This review briefly summarizes the evolution and development of SCLT technology, focuses on its application in hematology, and discusses its implications for the field of hematology.

The single-cell lineage tracing techniques

History and application of traditional lineage tracing

In the 19th century, lineage tracing was revolutionized by Charles O. Whitman and colleagues, who pioneered direct observation using light microscopy to track cell fates in invertebrate embryos [4]. Direct observation offers advantages such as speed, ease of use, and non-invasiveness. This method has also been successfully applied to zebrafish embryos due to their transparency [5]. However, continuous observation of intact organisms requires transparent embryos with a limited number of cells, which complicates the study of complex tissues or post-implantation mammalian embryos. When direct tracing isn't possible, alternative labeling techniques are necessary, such as using dyes and enzymes [6–8], cross-species transplantation [9], inducible activation of reporter gene expression through recombinase-mediated mechanisms [10, 11], and insertion of exogenous DNA sequences [12, 13]. Advances in molecular genetics have enabled more precise cell labeling through direct transfection or viral infection. Fluorescent proteins, often used as "molecular dyes," are common for tracking cell behavior. However, traditional lineage tracing methods have limited resolution. For example, the control of fluorescent proteins by site-specific recombinases [14], such as *Cre*, is typically regulated by cell-type-specific promoters, which restricts the fluorescent signal to a group of cells, reducing the precision of tracking individual cells.

Table 1 Comparison of various tracing methods

Methods	Mechanism	Application and advantages	Shortcoming
Multicolor label	Generate a diverse palette of colors that allow for the discrimination of different cells upon activation of the *Cre* recombinase	Neuronal connectivity patterns, stem cell proliferation dynamics, organ homeostasis, genetic regulation at the single-cell level within living organisms	Low resolution; the complexity of determining the optimal time and dosage to start the labeling process
Integration barcodes	Retroviral plasmid library	More tracing information quantity and accuracy. Analyze the clonal relationships between subpopulations of primitive hematopoietic hierarchy	Not suitable for quiescent or non-proliferating cell population; lack of reliable maintenance of lineage markers over time
Polylox barcodes	An artificial DNA recombination locus enables endogenous barcoding based on the *Cre-loxP* recombination system	Higher specificity Label single progenitor cells in vivo	Not suitable for human primary cells
CRISPR barcodes	Induce InDels and mutations to record more mitotic divisions	Revealed the trajectory of cell births and provided insight into the balance of symmetric and asymmetric cell division	Not suitable for human primary cells
Natural barcodes	Somatic cells spontaneously acquire mutations during development and aging processes	Used in human primary cells	Sequencing methods are not mature enough

Development and advantages of Single-cell lineage tracing

In the traditional approach to lineage tracing, all target cells are marked with the same label, leading to a lack of detailed information about the individual biological behaviors of each cell. Breakthroughs in single-cell sequencing have provided a way to break through the bottleneck, but to comprehensively track cell lines in complex multicellular systems, individual cells, and their progeny must be meticulously and sequentially labeled throughout numerous cell divisions. This labeling process must be performed without disturbing normal developmental processes. In addition, the selected cellular markers should accumulate irreversibly over time, a critical aspect that facilitates the subsequent reconstruction of the lineage tree and an essential component of a comprehensive lineage-tracing effort [2]. To achieve appropriate labeling and achieve higher-resolution genealogical tracing, the non-invasive study of gene fate mapping [1] has gained widespread use (Table 1; Fig. 1).

Multicolor labeling systems

Researchers have implemented strategies involving multiple *flox* sites and fluorescent proteins in specific combinations. Exemplified by systems like Brainbow and Confetti, these techniques create a diverse palette of colors that enable the discrimination of different cells upon activation of the *Cre* recombinase. Brainbow, in particular, has been a valuable source of biological insights, contributing to the understanding of various domains such as neuronal connectivity patterns [15–19], stem cell proliferation dynamics, organ homeostasis, and genetic regulation at the single-cell level within living organisms [20–22]. However, achieving single-cell resolution with these multicolor labeling systems has proven challenging, primarily due to complexities in determining the optimal timing and dosage for initiating the labeling process [1].

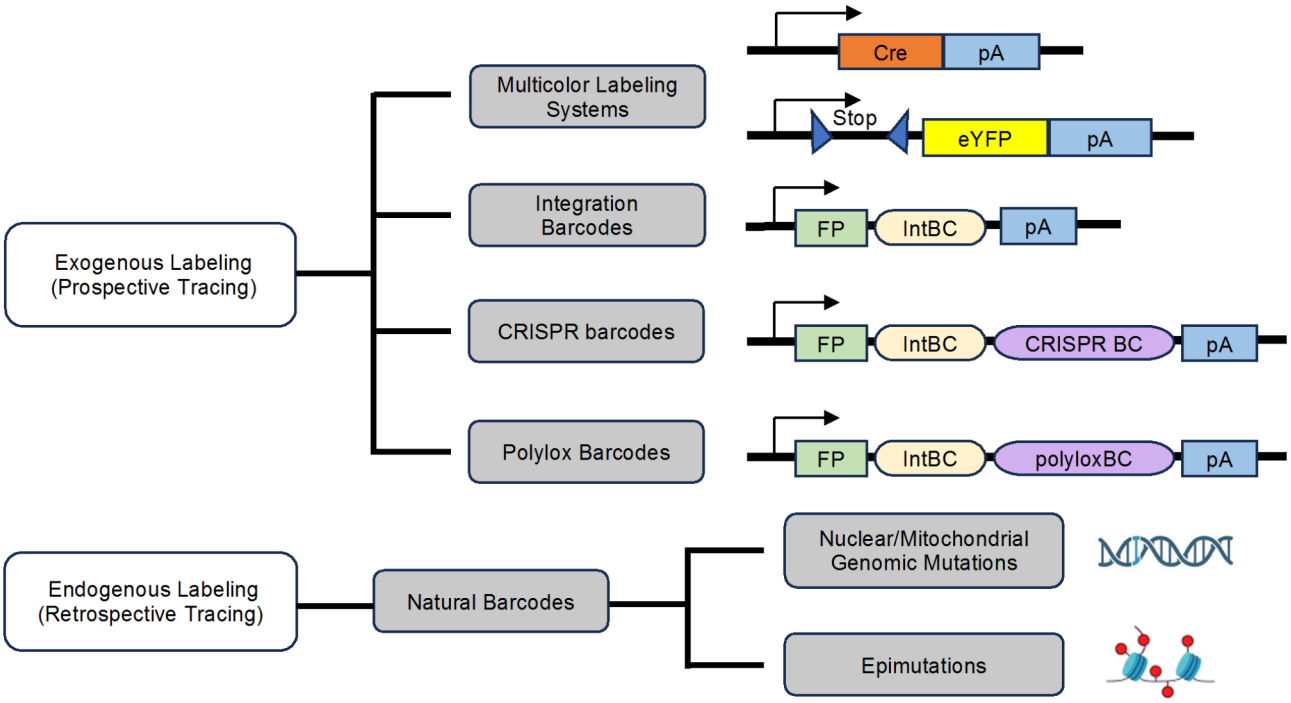

Fig. 1 The classification of single-cell lineage tracing

Additionally, the limited number of different fluorescently labeled proteins also constrains its resolution. Substantially greater quantities of cellular markers have been generated through techniques involving viral integration sites [23], viral barcodes [24], and strategies based on transposons [25] and CRISPR/Cas9 genome editing [2].

Integration barcodes

To further improve the resolution of lineage tracing, DNA barcoding techniques have been introduced, utilizing DNA fragments with extensive sequence variations to label individual cells. The discovery of retroviral vectors has provided a means to introduce new genetic material into HSCs [26–28], offering an approach for analyzing hematopoietic cell clones. Retroviral libraries containing DNA barcodes enable the simultaneous labeling of thousands of cells, thereby enhancing the quantity and precision of lineage tracing information [29–31].

In recipients of HSC transplantation mediated by retroviral transduction, the differentiation of distinct hematopoietic stem cell-derived clones can be achieved by considering retroviral integration sites as unique clonal markers [32, 33]. A retroviral plasmid library was developed, comprising vectors incorporating a variable, random sequence tag, or "barcode." Upon stable chromosomal integration, this barcode imparts a unique, discernible, and inheritable identifier into the genome, enabling the long-term tracking of clonal descendants originating from the host cell [24].

Through this technique, distinct subpopulations within the primitive hematopoietic hierarchy could be selectively isolated through fluorescence-activated cell sorting from recipients subjected to transplants involving barcoded HSCs. Subsequent analysis facilitated the examination of clonal relationships between these cellular compartments. Compared to previous studies involving single-cell transplants, the volume of information obtainable from a single transplant recipient was substantially enhanced. Moreover, such data served as a valuable resource for validating or augmenting established theoretical frameworks and quantitative models elucidating the clonal dynamics inherent to HSCs [34, 35].

Despite the advantages offered by viral barcoding methodologies, the use of retroviruses faces two significant challenges. Firstly, retroviruses are limited to labeling actively dividing cells, restricting their applicability to quiescent or non-proliferating cell populations. Secondly, retroviral vectors are prone to spontaneous silencing, potentially hindering the consistent maintenance of lineage markers over time. Additionally, it's crucial to recognize that spontaneous cell fusion phenomena occur in various tissue contexts [36, 37]. These fusion events can enable the transfer of lineage markers between cells of different lineages [38, 39], inadvertently leading to inaccurate conclusions about transdifferentiating processes [40].

Polylox barcodes and CRISPR barcodes

To address these challenges, novel genetic labels such as polylox barcodes and CRISPR barcodes have been developed. Polylox represents an artificial DNA recombination locus that enables endogenous barcoding using the *Cre-loxP* recombination system [41, 42], providing versatile applications and the ability to label single progenitor cells in vivo. The low probabilities of generating identical barcodes among multiple cells enhance the specificity of the labeling [2, 43, 44]. Additionally, cumulative CRISPR/Cas9 insertions and deletions (InDels) serve as genetic landmarks for reconstructing lineage hierarchies [45–47].

Given the extremely low rate of naturally occurring somatic mutations [48, 49], a common strategy in model organisms involves engineering a genetic cassette designed to mutate at a high rate [2, 43, 44, 50–58]. This cassette acts as a barcode, recording the mitotic division history of cells. Despite significant progress in cell barcoding [47, 59], the recording capacity of each barcode averages about three mutations, tracking at most three mitotic divisions. Since there are often tens of mitotic divisions for each cell in a fully developed complex organism during development, the reconstructed cell phylogenies from existing barcoding methods contain limited information [2, 53, 55, 60].

A recent breakthrough in lineage tracing involves the development of base editors, introducing informative sites to document cell division events. Indels and mutations induced by CRISPR-based genome editing offer faster mutation rates, allowing for the recording of more mitotic divisions and the construction of more detailed cell lineage trees [61]. Researchers applied this technique to Drosophila melanogaster, obtaining an average of more than 20 mutations on a three-kilobase-pair barcoding sequence in early-adult cells. Leveraging these barcoding mutations, high-quality cell phylogenetic trees were generated, each comprising several thousand internal nodes with 84–93% median bootstrap support. The resulting cell phylogenies facilitated a population genetic analysis, estimating the longitudinal dynamics of the number of actively dividing parental cells in each organ throughout development. The Np dynamics revealed the trajectory of cell births and provided insights into the balance of symmetric and asymmetric cell division [61].

Natural barcodes

While external markers mentioned above can distinguish and trace each cell and its descendants [55, 62, 63], enabling researchers to track the developmental pathways and differentiation processes of individual cells, these exogenous markers cannot be employed in the routine human body. However, somatic cells spontaneously acquire mutations during the development and aging process [45, 64–66]. These spontaneous mutations can be utilized for retrospective lineage tracing in human samples. Natural barcodes include nuclear genomic mutations and mitochondrial mutations [45]. In addition, epigenetic mutations can also serve as clues for lineage tracing [67, 68].

The tracing approach using natural barcodes is also called retrospective tracing because the inference of clonal relationships is obtained by backtracking based on mutation results. An important advantage of this endogenous marker method is that it is safe does not interfere with the natural differentiation process, and can be used to analyze the genealogy of human cells. However, the low mutation rate of the nuclear genome leads to the need for costly deep sequencing of the whole genome or exome of single cells [45]. Mitochondrial mutations have a relatively high mutation rate and higher copy number than nuclear mutations. However, the horizontal transfer of mitochondria between cells affects the accuracy of genealogical relationship analysis, and enrichment of mitochondrial DNA (mtDNA) is a major challenge [45]. Although previous research has demonstrated the potential of mitochondrial DNA (mtDNA) mutations as natural cell barcodes, existing methods can only detect a limited subset of mtDNA mutations, hindering the ability to resolve fine-grained subclonal relationships and hierarchical structures (69–73). However, the technology of using mtDNA as a barcode for lineage tracing has continued to progress. Previously, a research team used the modified 10× platform-based scATAC-seq technology for deep sequencing of mtDNA to construct lineage relationship [74]. Recently, the team has improved the technique by incorporating probe hybridization experiments and so on to further increase mtDNA enrichment and enhance the accuracy of lineage tracing. Generally, both nuclear and mitochondrial genomic mutations exhibit low mutation rates, making lineage tracing over short periods challenging [75]. Therefore, these methods are more suitable for long-term lineage tracing or when we have a longer observation window for mutation accumulation.

DNA methylation epimutations, however, present a potential solution to the challenge of low mutation rates in somatic mutations, due to their higher mutation frequency, easier access to DNA methylation data, and genetic stability. The mutation rate of DNA methylation (10^{-3} per CpG site per division) is significantly higher than that of nuclear (10^{-9} per nucleotide per division) and mitochondrial mutations (10^{-7} per nucleotide per division) [68]. Several studies have attempted to use epimutations for lineage tracing, and their feasibility has been preliminarily validated [67, 76]. Recently, Chen et al. developed MethylTree [68], a tool that reconstructs cell lineage from single-cell DNA methylation data with near-perfect accuracy, offering a promising approach

for non-invasive, high-resolution lineage tracing in the human body.

Overall, the ongoing advancements in natural barcodes-based lineage tracing technology will open new possibilities for investigating the lineage relationships of human progenitor cells.

Advancements in single-cell sequencing are crucial to support SCLT

Labeling cells is only one aspect; sequencing is equally crucial. Detecting the numerous barcodes that mark thousands of progenies poses a challenging task. To address this, transcribing barcode information into the mRNA pool and retrieving these barcodes through high-throughput sequencing methods, such as single-cell RNA sequencing (scRNA-seq), proves more practical. scRNA-seq allows for the measurement of gene expression in individual cells, offering insights into cell states and molecular dynamics. Its advantages over conventional single-cell sequencing include precise gene expression analysis in each cell, accurate discrimination of cell populations, inter-cellular classification comparisons, and the ability to detect the expression patterns of rare cells. Researchers leverage scRNA-seq to create molecular cell atlases, identify new cell types, and characterize cell lineage segregation [77–80].

The current trend in lineage tracing aims for finer resolution (from distinguishing a group of cells to a specific cell) and broader scale (from tracing a few cells to thousands of cells), encompassing both progenitor cells and their offspring. The development of single-cell lineage tracing, which combines single-cell omics (such as scRNA-seq) with lineage tracing methods, allows for the simultaneous detection of cell states and clonal relationships. This integration provides a more comprehensive understanding of lineage segregation and cell fate transitions. scRNA-seq provides high-resolution gene expression analysis but suffers from technical noise and high dropout rates, particularly for lowly expressed genes. Its limited genetic coverage prevents capturing all genomic variations, and short-read sequencing may lead to lost barcode information for lineage tracing, complicating accurate lineage reconstruction. Nevertheless, advancements in single-cell sequencing are expected to address these issues over time.

There are different types of single-cell lineage tracing techniques, and the figure briefly shows several different types. Abbreviations: IntBC, integration barcode; polyloxBC, polylox barcode; CRISPR BC, CRISPR barcode; FP, fluorescent protein; pA, polyA tail. Figure created with BioRender.com.

Single-Cell Lineage Tracing in Hematology

SCLT, as a new technology capable of tracing cell lineage differentiation relationships at single-cell resolution, can help resolve many aspects of cell differentiation, development, migration, and homeostasis, and thus can help answer many questions in the field of hematology.

Single-cell lineage tracing techniques in hematopoietic development

Many studies have applied SCLT technology to explore hematopoietic development.

Exploring lineage tracing with integration barcodes

A study reported a new technique for simultaneously mapping cell state and cell fate by first expressing "DNA barcodes" in cells and then tracking transcriptome changes over time. A heterogeneous population of progenitor cells was genetically labeled with DNA barcodes, and some cells were subjected to immediate single-cell RNA sequencing (scSeq) analysis, while others were subjected to scSeq analysis after cell division and differentiation. The DNA barcode structure was inserted into the 3'UTR region of a ubiquitously expressed EF1α promoter-driven enhancer green fluorescent protein gene in a DNA barcode library was constructed consisting of a random 28 bp fragment. The heritable DNA barcodes were subsequently inserted into the genome using modified lentiviral transfection, and scSeq sequencing detected the DNA barcodes. The authors applied this technique to study the differentiation process of hematopoietic stem and progenitor cells (HSPCs) in mice and identified two monocyte differentiation pathways: DC-like and Neu-like mononuclear precursor cells differentiated from each other, corresponding to MDPs (Mo-dendritic progenitors) and GMPs (granulocyte- Mo progenitors), respectively [31]. The application of Integration Barcodes-based single-cell lineage tracing technology provides us with a new approach to studying the cell differentiation process and its predictive factors, particularly in the context of the diverse differentiation pathways of hematopoietic stem/progenitor cells.

Insights from Polylox-Based lineage tracing: PolyloxExpress

Another study used PolyloxExpress, an SCLT technology based on Polylox barcodes. This tool builds on the Polylox system by expressing DNA barcodes as RNA barcodes, capturing the transcriptome of individual cells by single-cell sequencing, and using the flow of RNA barcodes to determine the developmental fate of each cell. The authors used Tie2-MeriCreMer mice to drive PolyloxExpress barcode labeling of hematopoietic stem cells and found that hematopoietic stem cells were classified into three main groups according to developmental fate by lineage-tracing experiments: multilineage hematopoietic

stem cells, which give rise to all lineages, and myeloid-erythroid hematopoietic stem cells, which give rise only to myeloid cells and erythroid cells (myeloid-erythroid restricted), and differentiation-inactive. The authors also revealed for the first time the relationship between hematopoietic stem cell division and differentiation in the natural state, and found that slow-dividing dormant hematopoietic stem cells also differentiate into mature cells and participate in the entire hematopoietic process; and that differentiation-inactive stem cells divide symmetrically to maintain self-renewal and the homeostasis of the hematopoietic stem cell pool [81]. These advancements in Polylox technology have laid the foundation for more accurate lineage tracing models in hematopoiesis. These findings underscore the potential of advanced barcoding technologies to enhance our understanding of cellular development and differentiation pathways.

Advancements in CRISPR-Based lineage tracing: CARLIN and DARLIN

In 2020, Sarah et al. first reported the CRISPR array repair lineage tracing (CARLIN) mouse line and corresponding analysis methods that can be used to simultaneously analyze the lineage and transcriptomic information of single cells in vivo. The team used CARLIN to characterize intrinsic biases in the clonal activity of fetal liver hematopoietic stem cells and to identify clonal bottlenecks in the HSC response to injury. However, its application is somewhat limited due to its low barcode diversity [60]. To further enhance barcode diversity and resolution at the single-cell level, a new generation of lineage tracing mouse lines, DARLIN, and a multi-omics analysis technology, Camellia-seq, have been developed. DARLIN, an inducible Cas9 barcoding mouse line that utilizes terminal deoxynucleotidyl transferase (TdT) and 30 CRISPR target sites. TdT is a broad-spectrum DNA polymerase that efficiently inserts additional bases after Cas9-induced DNA double-stranded breaks, thereby increasing the diversity of mutations to adequately label all cells in adult mice. Meanwhile, Camellia-seq can simultaneously measure transcriptome, DNA methylation, chromatin accessibility, and cellular genealogy at the single-cell level. Using the DARLIN with combined transcriptomic and epigenomic single-cell measurements to analyze multiple hematopoietic stem cells from the same lineage, the researchers validated the heterogeneity of HSC again. In addition, they found the existence of a group of HSCS with only one fate choice of macrophages, and their gene expression profile suggested that they were closer to long-term hematopoietic stem cells. More importantly, it was found that cellular clonal memory is associated with genome-wide DNA methylation. After many generations of cell division and migration, HSCs that are more closely related with more similar DNA methylation profile [82]. These advances in CRISPR technology paved the way for even more precise lineage tracing models in hematopoiesis.

Natural barcodes: A breakthrough in human lineage tracing

Although the diversity of DNA barcodes has continued to increase, the technology cannot be applied in vivo, which has limited research on human progenitor cells, as natural barcodes, offer a potential solution to overcome this limitation. A study published in Nature utilizes somatic mutations to construct a phylogenetic tree of fetal development. This is achieved through whole-genome sequencing of hematopoietic stem and progenitor cells from human embryos, combined with single-cell-derived cell colonies and deep targeted sequencing of tissues of known origin. This approach genealogically traces the development of the human hematopoietic system and differentiates the various stages of embryonic development [45]. The authors analyzed the constructed phylogenetic tree for asymmetry and found, consistent with previous studies, that the asymmetries contributed by the two daughter cells of the most recent common ancestor differed considerably. And the authors matched the time points of developmental acquisition of mutations with the time points of developmental events and found that the event of molecular divergence between the trophectoderm and the inner cell mass occurs at the 4–16 cell stage in the 8-week fetus, similar to the mouse developmental process. And at least 20 blood lineages were already present at the time of epiblast differentiation, a finding consistent with a number of in vitro observations. Compared with the endoderm and mesoderm, the ectoderm has only a relatively small contribution from the ancestor of the originating hematopoietic system. In contrast, the primitive blood captured in the circulatory system of the 8-week fetus was from the hypoderm, and its somatic mutations correlated more highly with those in the extraembryonic and mesodermal also tissues. Overall, this study provides a possible workflow as well as a conceptual framework for future understanding of human developmental processes through the use of naturally occurring somatic mutations as barcodes.

A recent article published in Nature proposes a new ReDeeM technique using mitochondrial DNA mutations as natural barcode that also provides clues to hematopoietic development. The ReDeeM system is based on deep sequencing of naturally occurring mitochondrial DNA mutations at the single-cell level, with simultaneous readout of transcriptional status and chromatin accessibility. Redeem utilizes a single-molecule consensus correction technique and is achieved by a modification of the 10X Genomics platform implementation, which significantly improves the sensitivity and accuracy of mutation detection. Redeem has been proven to increase the number of

mutations detected by more than 10-fold over previous methods and provide unprecedented opportunities for lineage tracing in human progenitor cells. The authors used the system to delve into the human hematopoietic process, analyzing data from bone marrow mononuclear cells (BMMCs) and hematopoietic stem cells (HSCs) from two young donors. The authors constructed phylogenetic trees through deep mtDNA mutation analysis, revealing clonal structure, functional heterogeneity, and age-related changes in hematopoietic stem cells. The study found functional differences in HSC clones that stabilized over several months, including differences in overall output and preferences for different mature cell types. With aging, the diversity of HSC clones significantly decreases, resulting in oligoclonal structures and the appearance of multiple distinctly expanding clones. This study provides the first atlas of the human hematopoietic system presented at single-cell resolution and lays the groundwork for more studies of clonal dynamics in human health and disease state [83].

As mentioned previously, low mutation rates in somatic cells lead to inefficiencies and make it difficult to accurately discriminate lineages. One promising approach to address this issue is lineage tracing through epigenetic mutations. A recent study published in Nature Methods introduced a significant advancement in this area with the development of MethylTree, a novel computational tool designed for single-cell lineage tracing based on DNA methylation epimutations [68]. Single-cell DNA methylation sequencing typically covers only about 5% of the genome, and its limited coverage can affect the accuracy of the genealogical information it provides. MethylTree addresses this challenge by directly calculating the Pearson correlation of overlapping CpG sites between two cellular "fragments" to construct a similarity matrix. This approach effectively resolves the sparsity issue inherent in DNA methylation data. Furthermore, MethylTree employs iterative algorithms to minimize noise between cells and filtering out DNA methylation variations that are specific to different cell types. The authors validated MethylTree's accuracy, demonstrating near 100% precision through the use of public databases and multiple experimental systems, including mouse and human hematopoietic models and early embryonic development. Additionally, MethylTree was successfully applied to investigate the model of the first cell fate decision in the human embryo, revealing this decision occurs early at the four-cell stage. The tool's capacity to count clones in vivo was demonstrated in an animal model, where approximately 250 hematopoietic stem cell clones were identified in the blood of wild-type mice. In summary, the development of MethylTree represents a breakthrough in overcoming the limitations imposed by mutations in nuclear and mitochondrial genomes. This algorithm opens new avenues for high-resolution, non-invasive, multi-omics genealogical tracing, with broad applications in human studies and beyond.

In conclusion, although different labeling methods have been chosen for different studies, SCLT technology allows us to comprehensively follow the developmental processes of multicellular systems in various states and to analyze the transcriptional regulation behind cell fate selection. SCLT provides more information and greater possibilities for hematopoietic developmental studies than traditional genealogical tracer methods and common single-cell sequencing methods.

Single-Cell lineage tracing in hematopoietic disorders: Understanding disease and stem cell dynamics

SCLT technology, in addition to being used to explore developmental processes, can likewise be used to explore disease evolution and the mechanisms that lead to treatment resistance in a wide range of cancers. mtDNA mutations themselves serve as natural markers of clonal profiling and are capable of tracking multispectral populations of cancerous and non-cancerous fractions of tissues from which they originate. Livius et al. used mtDNA as a natural barcode on patients with chronic lymphocytic leukemia for single-cell lineage tracing, revealing clonal and functional heterogeneity in human malignancies [84]. Livius et al. applied the mtscATAC-seq technique to nine patients with chronic lymphocytic leukemia in the setting of different therapeutic modalities and found that mtDNA showed long-term stability in the absence of treatment [70]. However, dramatic changes in mtDNA mutations were observed in case of major disease progression or relapse after strong therapeutic pressure. In addition, they linked chronic lymphocytic leukemia subclones to different chromatin states, further confirming the relevance of epigenomic evolution to acquired drug resistance. Another study was conducted to investigate the genealogical history of chronic lymphocytic leukemia (CLL) cells using DNA methylation epimutation data, employing the multiplexed single-cell RRBS technique (MscRRBS) [67]. The results revealed that, compared to normal B cells, cancer cells from CLL patients exhibited significantly higher levels of DNA methylation mutations. However, the differences observed among the cancer cells themselves were relatively small. A comparison of the Phylogenetic trees of both normal and cancer cells showed that cancer cells displayed earlier bifurcations and longer branches, consistent with the proliferative characteristics of tumor cells. In conclusion, these findings underscore the feasibility of utilizing natural barcodes to track the subclonal structure of disease populations. Furthermore, the continued advancements in SCLT technology hold great

potential for enabling patient-specific studies of cancer subclonal dynamics soon.

The fate of hematopoietic stem cells is largely regulated by a variety of cells of non-hematopoietic origin within the hematopoietic microenvironment, and a large number of recent studies have shown that the fate of HSCs is also regulated by "progenitor" cells of hematopoietic origin such as macrophages, and T cells [85–87]. Within the hematopoietic hierarchy, Flk2- MPPs (multipotent progenitors) are the first downstream progenitors generated by the differentiation of HSCs and also reside in the bone marrow microenvironment, and it remains unknown whether they regulate the behavior and fate of HSCs. A study utilized SCLT technology to label and dynamically track each HSC after co-transplantation of FLK2- MPPs in mice. The results revealed that co-transplantation of Flk2- MPPs significantly increased the ability of HSCs to produce T cells.Flk2- MPPs co-transplantation promoted the maintenance of multi-lineage HSCs in the early stage (2.5 months before transplantation), and lymphoid-biased/specialized HSCs formation in the late stage (5.5-0.6.5 months post-transplantation). HSCs formation. In conclusion, Flk2- MPPs influence the fate of HSCs in terms of stemness maintenance, clonal amplification, and lineage differentiation, which is informative for the reconstruction of hematopoietic and immune systems after transplantation for diseases such as aged leukemia [87].

Conclusions

Lineage tracing is an effective technique for tracking cellular progeny and providing important insights into cell behavior and lineage dynamics. This review describes lineage tracing, highlighting the disadvantages of traditional methods and the advantages of single-cell lineage tracing technologies, including integration barcoding, CRISPR barcoding, and base editors. Applications of lineage tracing in hematologic research are explored, including applications in hematopoietic development, hematologic oncology, and therapeutics. These studies have greatly advanced our understanding of hematologic development and its evolution in disease progression. With the evolution of single-cell omics techniques like single-cell RNA sequencing, the resolution and scale of lineage tracking have improved, allowing for the simultaneous detection of cell state and clonal relationships. It can be confidently predicted that these studies will gain further momentum, extending lineage-tracking applications to various diseases, not only in mice but also in humans.

Abbreviations

SCLT	single-cell lineage tracing
HSCs	Hematopoietic stem cells
InDels	insertions and deletions
mtDNA	mitochondrial DNA
scRNA-seq	single-cell RNA sequencing
scSeq	single-cell RNA sequencing
MDPs	Mo-dendritic progenitors
GMPs	granulocyte- Mo progenitors
TdT	terminal deoxynucleotidyl transferase
BMMCs	bone marrow mononuclear cells
MPPs	multipotent progenitors
CLL	chronic lymphocytic leukemia

Acknowledgements
Not applicable.

Author contributions
All authors contributed to the work equally.

Funding
National Natural Science Foundation of China (No.82070184, 82350105, 82270228).
National Science and Technology Major Project (No. 2023ZD0501200).
Science, Technology& Innovation Project of Xiongan New area (No. 2023XACX0004).
Beijing Outstanding Young Scientists Project(JWZQ20240101001).
Beijing Natural Science Foundation(Z240019).

Data availability
Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no conflicts of interest.

Received: 18 January 2025 / Accepted: 23 February 2025
Published online: 04 March 2025

References

1. Kretzschmar K, Watt FM. Lineage tracing. Cell. 2012;148(1–2):33–45.
2. McKenna A, Findlay GM, Gagnon JA, Horwitz MS, Schier AF, Shendure J. Whole-organism lineage tracing by combinatorial and cumulative genome editing. Science. 2016;353(6298):aaf7907.
3. Saelens W, Cannoodt R, Todorov H, Saeys Y. A comparison of single-cell trajectory inference methods. Nat Biotechnol. 2019;37(5):547–54.
4. Child C. The Organization and Cell-lineage of the Ascidian Egg. By, Professor EG, Conklin I. 1905. Science. 1906;23(583):340-4.
5. Kimmel CB, Warga RM, Schilling TF. Origin and organization of the zebrafish fate map. Development. 1990;108(4):581–94.
6. Kimmel CB, Law RD. Cell lineage of zebrafish blastomeres. III. Clonal analyses of the blastula and gastrula stages. Dev Biol. 1985;108(1):94–101.
7. Keller RE. Vital dye mapping of the gastrula and Neurula of Xenopus laevis. II. Prospective areas and morphogenetic movements of the deep layer. Dev Biol. 1976;51(1):118–37.
8. Weisblat DA, Sawyer RT, Stent GS. Cell lineage analysis by intracellular injection of a tracer enzyme. Science. 1978;202(4374):1295–8.
9. Le Douarin NM, Teillet MA. Experimental analysis of the migration and differentiation of neuroblasts of the autonomic nervous system and of neurectodermal mesenchymal derivatives, using a biological cell marking technique. Dev Biol. 1974;41(1):162–84.
10. Dymecki SM, Tomasiewicz H. Using Flp-recombinase to characterize expansion of Wnt1-expressing neural progenitors in the mouse. Dev Biol. 1998;201(1):57–65.

11. Zinyk DL, Mercer EH, Harris E, Anderson DJ, Joyner AL. Fate mapping of the mouse midbrain-hindbrain constriction using a site-specific recombination system. Curr Biol. 1998;8(11):665–8.
12. Walsh C, Cepko CL. Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. Science. 1992;255(5043):434–40.
13. Porter SN, Baker LC, Mittelman D, Porteus MH. Lentiviral and targeted cellular barcoding reveals ongoing clonal dynamics of cell lines in vitro and in vivo. Genome Biol. 2014;15(5):R75.
14. Holland EC, Varmus HE. Basic fibroblast growth factor induces cell migration and proliferation after glia-specific gene transfer in mice. Proc Natl Acad Sci U S A. 1998;95(3):1218–23.
15. Livet J, Weissman TA, Kang H, Draft RW, Lu J, Bennis RA, et al. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature. 2007;450(7166):56–62.
16. Egawa R, Hososhima S, Hou X, Katow H, Ishizuka T, Nakamura H, et al. Optogenetic probing and manipulation of the calyx-type presynaptic terminal in the embryonic chick ciliary ganglion. PLoS ONE. 2013;8(3):e59179.
17. Heap LA, Goh CC, Kassahn KS, Scott EK. Cerebellar output in zebrafish: an analysis of Spatial patterns and topography in eurydendroid cell projections. Front Neural Circuits. 2013;7:53.
18. Kang H, Lichtman JW. Motor axon regeneration and muscle reinnervation in young adult and aged animals. J Neurosci. 2013;33(50):19480–91.
19. Robles E, Filosa A, Baier H. Precise lamination of retinal axons generates multiple parallel input pathways in the tectum. J Neurosci. 2013;33(11):5027–39.
20. Förster D, Luschnig S. Src42A-dependent polarized cell shape changes mediate epithelial tube elongation in Drosophila. Nat Cell Biol. 2012;14(5):526–34.
21. Wachsman G, Heidstra R, Scheres B. Distinct cell-autonomous functions of RETINOBLASTOMA-RELATED in Arabidopsis stem cells revealed by the brother of brainbow clonal analysis system. Plant Cell. 2011;23(7):2581–91.
22. Loulier K, Barry R, Mahou P, Le Franc Y, Supatto W, Matho KS, et al. Multiplex cell and lineage tracking with combinatorial labels. Neuron. 2014;81(3):505–20.
23. Keller G, Paige C, Gilboa E, Wagner EF. Expression of a foreign gene in myeloid and lymphoid cells derived from multipotent Haematopoietic precursors. Nature. 1985;318(6042):149–54.
24. Gerrits A, Dykstra B, Kalmykowa OJ, Klauke K, Verovskaya E, Broekhuis MJ, et al. Cellular barcoding tool for clonal analysis in the hematopoietic system. Blood. 2010;115(13):2610–8.
25. Sun J, Ramos A, Chapman B, Johnnidis JB, Le L, Ho YJ, et al. Clonal dynamics of native haematopoiesis. Nature. 2014;514(7522):322–7.
26. Williams DA, Lemischka IR, Nathan DG, Mulligan RC. Introduction of new genetic material into pluripotent Haematopoietic stem cells of the mouse. Nature. 1984;310(5977):476–80.
27. Joyner A, Keller G, Phillips RA, Bernstein A. Retrovirus transfer of a bacterial gene into mouse Haematopoietic progenitor cells. Nature. 1983;305(5934):556–8.
28. Miller AD, Eckner RJ, Jolly DJ, Friedmann T, Verma IM. Expression of a retrovirus encoding human HPRT in mice. Science. 1984;225(4662):630–2.
29. Lu R, Neff NF, Quake SR, Weissman IL. Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. Nat Biotechnol. 2011;29(10):928–33.
30. Naik SH, Schumacher TN, Perié L. Cellular barcoding: a technical appraisal. Exp Hematol. 2014;42(8):598–608.
31. Weinreb C, Rodriguez-Fraticelli A, Camargo FD, Klein AM. Lineage tracing on transcriptional landscapes links state to fate during differentiation. Science. 2020;367(6479).
32. Capel B, Hawley RG, Mintz B. Long- and short-lived murine hematopoietic stem cell clones individually identified with retroviral integration markers. Blood. 1990;75(12):2267–70.
33. Jordan CT, Lemischka IR. Clonal and systemic analysis of long-term hematopoiesis in the mouse. Genes Dev. 1990;4(2):220–32.
34. Roeder I, Loeffler M. A novel dynamic model of hematopoietic stem cell organization based on the concept of within-tissue plasticity. Exp Hematol. 2002;30(8):853–61.
35. Roeder I, Horn K, Sieburg HB, Cho R, Muller-Sieburg C, Loeffler M. Characterization and quantification of clonal heterogeneity among hematopoietic stem cells: a model-based approach. Blood. 2008;112(13):4874–83.
36. Alvarez-Dolado M, Pardal R, Garcia-Verdugo JM, Fike JR, Lee HO, Pfeffer K, et al. Fusion of bone-marrow-derived cells with purkinje neurons, cardiomyocytes and hepatocytes. Nature. 2003;425(6961):968–73.
37. Wang X, Willenbring H, Akkari Y, Torimaru Y, Foster M, Al-Dhalimy M, et al. Cell fusion is the principal source of bone-marrow-derived hepatocytes. Nature. 2003;422(6934):897–901.
38. Ying QL, Nichols J, Evans EP, Smith AG. Changing potency by spontaneous fusion. Nature. 2002;416(6880):545–8.
39. Terada N, Hamazaki T, Oka M, Hoki M, Mastalerz DM, Nakano Y, et al. Bone marrow cells adopt the phenotype of other cells by spontaneous cell fusion. Nature. 2002;416(6880):542–5.
40. Weissman IL. Normal and neoplastic stem cells. Novartis Found Symp. 2005;265:35–50; discussion– 4, 92– 7.
41. Sternberg N, Hamilton D. Bacteriophage P1 site-specific recombination. I. Recombination between LoxP sites. J Mol Biol. 1981;150(4):467–86.
42. Rajewsky K, Gu H, Kühn R, Betz UA, Müller W, Roes J, et al. Conditional gene targeting. J Clin Invest. 1996;98(3):600–3.
43. Kalhor R, Kalhor K, Mejia L, Leeper K, Graveline A, Mali P et al. Developmental barcoding of whole mouse via homing CRISPR. Science. 2018;361(6405).
44. Pei W, Feyerabend TB, Rössler J, Wang X, Postrach D, Busch K, et al. Polylox barcoding reveals Haematopoietic stem cell fates realized in vivo. Nature. 2017;548(7668):456–60.
45. Spencer Chapman M, Ranzoni AM, Myers B, Williams N, Coorens THH, Mitchell E, et al. Lineage tracing of human development through somatic mutations. Nature. 2021;595(7865):85–90.
46. Espinosa-Medina I, Garcia-Marques J, Cepko C, Lee T. High-throughput dense reconstruction of cell lineages. Open Biol. 2019;9(12):190229.
47. Baron CS, van Oudenaarden A. Unravelling cellular relationships during development and regeneration using genetic lineage tracing. Nat Rev Mol Cell Biol. 2019;20(12):753–65.
48. Behjati S, Huch M, van Boxtel R, Karthaus W, Wedge DC, Tamuri AU, et al. Genome sequencing of normal cells reveals developmental lineages and mutational processes. Nature. 2014;513(7518):422–5.
49. Luo T, He X, Xing K. Lineage analysis by microsatellite loci deep sequencing in mice. Mol Reprod Dev. 2016;83(5):387–91.
50. Kalhor R, Mali P, Church GM. Rapidly evolving homing CRISPR barcodes. Nat Methods. 2017;14(2):195–200.
51. Frieda KL, Linton JM, Hormoz S, Choi J, Chow KK, Singer ZS, et al. Synthetic recording and in situ readout of lineage information in single cells. Nature. 2017;541(7635):107–11.
52. Alemany A, Florescu M, Baron CS, Peterson-Maduro J, van Oudenaarden A. Whole-organism clone tracing using single-cell sequencing. Nature. 2018;556(7699):108–12.
53. Raj B, Wagner DE, McKenna A, Pandey S, Klein AM, Shendure J, et al. Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. Nat Biotechnol. 2018;36(5):442–50.
54. Spanjaard B, Hu B, Mitic N, Olivares-Chauvet P, Janjuha S, Ninov N, et al. Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. Nat Biotechnol. 2018;36(5):469–73.
55. Chan MM, Smith ZD, Grosswendt S, Kretzmer H, Norman TM, Adamson B, et al. Molecular recording of mammalian embryogenesis. Nature. 2019;570(7759):77–82.
56. Hwang B, Lee W, Yum SY, Jeon Y, Cho N, Jang G, et al. Lineage tracing using a Cas9-deaminase barcoding system targeting endogenous L1 elements. Nat Commun. 2019;10(1):1234.
57. Ye C, Chen Z, Liu Z, Wang F, He X. Defining endogenous barcoding sites for CRISPR/Cas9-based cell lineage tracing in zebrafish. J Genet Genomics. 2020;47(2):85–91.
58. Chen H, Liu S, Padula S, Lesman D, Griswold K, Lin A, et al. Efficient, continuous mutagenesis in human cells using a pseudo-random DNA editor. Nat Biotechnol. 2020;38(2):165–8.
59. Wagner DE, Klein AM. Lineage tracing Meets single-cell omics: opportunities and challenges. Nat Rev Genet. 2020;21(7):410–27.
60. Bowling S, Sritharan D, Osorio FG, Nguyen M, Cheung P, Rodriguez-Fraticelli A, et al. An engineered CRISPR-Cas9 mouse line for simultaneous readout of lineage histories and gene expression profiles in single cells. Cell. 2020;181(6):1410–e2227.
61. Liu K, Deng S, Ye C, Yao Z, Wang J, Gong H, et al. Mapping single-cell-resolution cell phylogeny reveals cell population dynamics during organ development. Nat Methods. 2021;18(12):1506–14.
62. Sankaran VG, Weissman JS, Zon LI. Cellular barcoding to Decipher clonal dynamics in disease. Science. 2022;378(6616):eabm5874.
63. Yang D, Jones MG, Naranjo S, Rideout WM 3rd, Min KHJ, Ho R, et al. Lineage tracing reveals the phylodynamics, plasticity, and paths of tumor evolution. Cell. 2022;185(11):1905–e2325.

64. Lee-Six H, Øbro NF, Shepherd MS, Grossmann S, Dawson K, Belmonte M, et al. Population dynamics of normal human blood inferred from somatic mutations. Nature. 2018;561(7724):473–8.
65. Coorens THH, Moore L, Robinson PS, Sanghvi R, Christopher J, Hewinson J, et al. Extensive phylogenies of human development inferred from somatic mutations. Nature. 2021;597(7876):387–92.
66. Park S, Mali NM, Kim R, Choi JW, Lee J, Lim J, et al. Clonal dynamics in early human embryogenesis inferred from somatic mutation. Nature. 2021;597(7876):393–7.
67. Gaiti F, Chaligne R, Gu H, Brand RM, Kothen-Hill S, Schulman RC, et al. Epigenetic evolution and lineage histories of chronic lymphocytic leukaemia. Nature. 2019;569(7757):576–80.
68. Chen M, Fu R, Chen Y, Li L, Wang S-W. High-resolution, noninvasive single-cell lineage tracing in mice and humans based on DNA methylation epimutations. Nat Methods. 2025.
69. Ludwig LS, Lareau CA, Ulirsch JC, Christian E, Muus C, Li LH, et al. Lineage tracing in humans enabled by mitochondrial mutations and Single-Cell genomics. Cell. 2019;176(6):1325–e3922.
70. Lareau CA, Ludwig LS, Muus C, Gohil SH, Zhao T, Chiang Z, et al. Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling. Nat Biotechnol. 2021;39(4):451–61.
71. Mimitou EP, Lareau CA, Chen KY, Zorzetto-Fernandes AL, Hao Y, Takeshima Y, et al. Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. Nat Biotechnol. 2021;39(10):1246–58.
72. Xu J, Nuno K, Litzenburger UM, Qi Y, Corces MR, Majeti R et al. Single-cell lineage tracing by endogenous mutations enriched in transposase accessible mitochondrial DNA. Elife. 2019;8.
73. Miller TE, Lareau CA, Verga JA, DePasquale EAK, Liu V, Ssozi D, et al. Mitochondrial variant enrichment from high-throughput single-cell RNA sequencing resolves clonal populations. Nat Biotechnol. 2022;40(7):1030–4.
74. Lareau CA, Liu V, Muus C, Praktiknjo SD, Nitsch L, Kautz P, et al. Mitochondrial single-cell ATAC-seq for high-throughput multi-omic detection of mitochondrial genotypes and chromatin accessibility. Nat Protoc. 2023;18(5):1416–40.
75. Lynch M. Evolution of the mutation rate. Trends Genet. 2010;26(8):345–52.
76. Liu Y, Li XC, Rashidi Mehrabadi F, Schäffer AA, Pratt D, Crawford DR, et al. Single-cell methylation sequencing data reveal succinct metastatic migration histories and tumor progression models. Genome Res. 2023;33(7):1089–100.
77. Hwang B, Lee JH, Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp Mol Med. 2018;50(8):1–14.
78. Lafzi A, Moutinho C, Picelli S, Heyn H. Tutorial: guidelines for the experimental design of single-cell RNA sequencing studies. Nat Protoc. 2018;13(12):2742–57.
79. 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.
80. Potter SS. Single-cell RNA sequencing for the study of development, physiology and disease. Nat Rev Nephrol. 2018;14(8):479–92.
81. Pei W, Shang F, Wang X, Fanti AK, Greco A, Busch K, et al. Resolving fates and Single-Cell transcriptomes of hematopoietic stem cell clones by polyloxexpress barcoding. Cell Stem Cell. 2020;27(3):383–e958.
82. Li L, Bowling S, McGeary SE, Yu Q, Lemke B, Alcedo K, et al. A mouse model with high clonal barcode diversity for joint lineage, transcriptomic, and epigenomic profiling in single cells. Cell. 2023;186(23):5183–e9922.
83. Weng C, Yu F, Yang D, Poeschla M, Liggett LA, Jones MG, et al. Deciphering cell States and genealogies of human haematopoiesis. Nature. 2024;627(8003):389–98.
84. Penter L, Gohil SH, Lareau C, Ludwig LS, Parry EM, Huang T, et al. Longitudinal Single-Cell dynamics of chromatin accessibility and mitochondrial mutations in chronic lymphocytic leukemia mirror disease history. Cancer Discov. 2021;11(12):3048–63.
85. Fujisaki J, Wu J, Carlson AL, Silberstein L, Putheti P, Larocca R, et al. In vivo imaging of Treg cells providing immune privilege to the Haematopoietic stem-cell niche. Nature. 2011;474(7350):216–9.
86. Bruns I, Lucas D, Pinho S, Ahmed J, Lambert MP, Kunisaki Y, et al. Megakaryocytes regulate hematopoietic stem cell quiescence through CXCL4 secretion. Nat Med. 2014;20(11):1315–20.
87. Zhao M, Perry JM, Marshall H, Venkatraman A, Qian P, He XC, et al. Megakaryocytes maintain homeostatic quiescence and promote post-injury regeneration of hematopoietic stem cells. Nat Med. 2014;20(11):1321–6.

Publisher's note