Single-cell mitochondrial variant enrichment resolved clonal tracking and spatial architecture in human embryonic hematopoiesis

Yan Xue[1,2,*], Yiming Chao[1,3,*], Xinyi Lin[1], Yuanhua Huang[1,3,4], Joshua WK Ho[1,2,#], Ryohichi Sugimura [1,3,#]

1 School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, HKSAR

2 Laboratory of Data Discovery for Health Limited ($D^2$4H), Hong Kong Science Park, HKSAR

3 Centre for Translational Stem Cell Biology, HKSAR

4 Department of Statistics and Actuarial Science, Faculty of Science, The University of Hong Kong, HKSAR

* These authors contributed equally.

For correspondence: rios@hku.hk, jwkho@hku.hk

Abstract

The ability to perform lineage tracing at the single-cell level is critical to reconstructing dynamic transitions during cell differentiation. However, prospective tracing approaches inevitably encounter outstanding challenges including barcoding precision, barcode diversity, and detection efficiency, which can skew inferred lineage relationships. Human pluripotent stem cells (hPSC) even face risks of DNA-damage-induced toxicity-related cell death. We explored the use of naturally occurring somatic mutations in mitochondrial transcripts detected in single-cell RNA-seq as genetic lineage barcodes in HPSC. In this study, we used an enrichment of scRNA-seq mitochondrial reads and a robust computational method to identify clonally relevant mitochondrial variants as endogenous genetic barcodes for clonal tracking of early embryonic hematopoiesis from hPSC. We modeled the development of embryonic tissues from hPSCs and delineated the ontogeny of hematopoietic cells by mitochondrial variant lineage tracing. We identified multiple waves of erythropoiesis in time-series scRNA-seq data. We further applied mitochondrial variant lineage tracing to spatial transcriptomics and identified the clonal development of both erythro-myeloid progenitors and their niche cells. Our study highlights the power of mitochondrial variants as an endogenous marker for clonal tracking in modeling human development from stem cells.

Introduction

Recapitulation of lineage relationships during embryonic development provides key insights into cell fate. Combining molecular barcoding with single-cell transcriptomics such as CRISPR-based lineage tracer enables continuous and high-resolution lineage tracing while allowing analysis of cell states (C. S. Baron & van Oudenaarden, 2019; Bowling et al., 2020; Kebschull & Zador, 2018; Kester & van Oudenaarden, 2018; Li et al., 2023; Quinn et al., 2021; Tian et al., 2021; VanHorn & Morris, 2021, 2021; Wagner & Klein, 2020; Weinreb et al., 2020; Woodworth et al., 2017; Yang et al., 2022). The application of prospective lineage tracing involves various biological processes including early mammalian embryogenesis, hematopoiesis, neural development, and cancer progression (Bowling et al., 2020; Chan et al., 2019; Kalhor et al., 2018; Salvador-Martínez et al., 2019; He et al., 2022; Li et al., 2023). However, barcoding precision, barcode diversity, and detection efficiency could affect the resolution and accuracy of lineage reconstruction. In addition, genetic modification in multiple loci is inherently necessary in this tracing system, which is technically challenging and could intervene differentiation capacity of stem cells. Continuous endonuclease activity of Cas9 has been shown to cause cell death in human pluripotent stem cells (Ihry et al., 2018).

Naturally occurring somatic mutations can be used for mapping the lineage without genetic manipulation of organisms. Variations in the nuclear genome are a significant factor in shaping the dynamics of heteroplasmy and variation in mitochondrial copy number across different populations of humans (Gupta et al., 2023). Mitochondrial variants stand out with their high mutation rate, copy number, and heteroplasmy and low sequencing cost among various types of endogenous markers. This allows simultaneously interrogating clonality and gene expression or chromatin state profiles in single-cell level (Beneyto-Calabuig et al., 2023; Lareau et al., 2021; Ludwig et al., 2019; Penter et al., 2021; Walker et al., 2022; Wang et al., 2022; Xu et al., 2019). The development of MAESTER (mitochondrial alteration enrichment from single-cell transcriptomes to establish relatedness) expands the application of mtDNA as a natural genetic cell barcode for clonal tracking to high throughput single-cell RNA sequencing (Miller et al., 2022). We developed MQuad (mixture modelling of mitochondrial mutations) with a binomial mixture model to determine mitochondrial heteroplasmy and dissected clonal substructure using single cell mitochondrial variants (Kwok et al., 2022). Spatial transcriptomics delineated topological heterogeneity of transcriptomics and has successfully identified clonality across multiple organs (Deng et al., 2022; Erickson et al., 2022; Ji et al., 2020; Nagasawa et al., 2021; Zhao et al., 2022). Spatial transcriptomics with multiplexed ion beam imaging was used to identify clones and understand the cellular composition and architecture of cutaneous squamous cell carcinoma (cSCC), in which four tumor subpopulations, including a tumor-specific keratinocyte population unique to cancer that localized to a fibrovascular niche were identified (Ji et al., 2020). In invasive ductal carcinoma, spatial transcriptomics and targeted deep sequencing analysis revealed GATA3 mutations as potential genomic high-risk factors, with evidence suggesting the presence of GATA3-mutated clones from primary lesions in patients who experienced invasive ductal carcinoma relapse (Nagasawa et al., 2021). Combining in situ CUT&Tag chemistry and microfluidic deterministic barcoding, spatial-CUT&Tag enabled to identify specific clusters enriched with oligodendrocyte lineage cells in mouse embryos and brain tissue, offering new opportunities for studying epigenetic regulation in normal physiology and pathogenesis (Deng et al., 2022).

Erickson et al employed a systematic approach to investigate spatial genome integrity in situ and identify unique clonal patterns within tumors and nearby benign tissues through genome-wide copy number variation (Erickson et al., 2022). With slide-DNA-seq, Zhao et al captured spatially resolved DNA sequences from intact tissue sections, allowing for the identification of distinct tumour clones and their copy number alterations, and they further discovered clone-specific genetic aberrations and their associations with the local tumour microenvironment through integration with spatial transcriptomics, providing insights into the contribution of cell-intrinsic and cell-extrinsic factors to cellular phenotypes (Zhao et al., 2022). Spatial transcriptomics with lineage recorders which combines reporter barcodes with inducible CRISPR–Cas9 scarring deciphered the clonality and lineage dynamics during cerebral organoid development by linking molecular state, lineage and topological information (He et al., 2022).

Hematopoiesis is essential to life. Generation of red blood cells convey oxygen throughout tissues. Generation of myeloid and lymphoid cells combat pathogens and transformed cells in the body. Hematopoiesis begins from the yolk sac of early embryos. The yolk sac produces erythroblasts to sustain survival of embryos by supplying oxygen (Orkin & Zon, 2008). Not only direct differentiation of erythroblasts from mesodermal progenitors, but also a specific cell called erythro-myeloid progenitor (EMP) generate erythroblasts, macrophages and megakaryocytes in the yolk sac (Gomez Perdiguero et al., 2015; McGrath et al., 2015). A thorough analysis of human yolk sac identified accelerated production of macrophages independent of monocytes. PSC differentiation model recapitulated both monocyte-independent macrophages and monocyte-dependent macrophages (Goh et al., 2023). It is of paramount importance to understand if PSC differentiation model can recapitulate embryonic erythropoiesis directly from mesodermal progenitors and fetal erythropoiesis from hematopoietic progenitors (Hansen et al., 2019). Globin gene expression distinguishes embryonic vs fetal erythropoiesis, former express HBZ, HBE, and the latter express HBG (Grosso et al., 2012; Palis, 2014). The combination of globin gene expression and molecular barcodes will delineate clonal origin of each erythropoiesis.

The lineage mapping of hematopoiesis started from the robust reconstruction power of hematopoietic stem cells. Transplantation of retrovirally labelled HSCs determined lineage reconstruction capacity from a single cell (Lemischka et al., 1986), as well as defined lineage-bias and its associated epigenetic signatures in hematopoiesis at extreme stress condition, namely transplantation hematopoiesis (Challen et al., 2010; Lu et al., 2011; Martin et al., 2023). The lineage tracing with HSC-specific reporter animal models advanced the understanding of HSC commitment and locomotion in native hematopoiesis, however its application is limited to mouse, not human (Acar et al., 2015; Chen et al., 2016; Christodoulou et al., 2020; Gazit et al., 2014). The application of molecular barcoding revolutionized the understanding of HSC lineage in native hematopoiesis (Bowling et al., 2020; Pei et al., 2017; Sun et al., 2014). Recent study revealed the heterologous mesodermal populations and hematopoietic progenitors in embryos contribute to life-long hematopoiesis which was previously attributed to HSC exclusively (Patel et al., 2022; Yokomizo et al., 2022). Molecular barcoding will deepen the resolution of embryonic hematopoiesis.

The pluripotent stem cells opened the venue to model human embryonic development. Reconstruction of early embryo-like structures from stem cells recapitulate the complexity of

developing tissues and organs (Chao et al., 2023; Hislop et al., 2023; Kagawa et al., 2022; Oldak et al., 2023; Pedroza et al., 2023; Weatherbee et al., 2023). Here, we combined mitochondrial variant enrichment with robust computational calling approaches to rebuilt lineage history during embryonic hematopoiesis from human stem cells. We delineated mitochondria variants with combination of MAESTER and MQuad pipeline. We determined clonal information of embryonic hematopoiesis from human pluripotent stem cells. We identified embryonic and fetal erythropoiesis derived from different clones. Spatial transcriptomics identified co-development of both hematopoietic and their niche cells from the same clones in the yolk sac-like tissues. These results demonstrate that mitochondrial variants resolve clonal architecture of human embryonic hematopoiesis.

Results

Dichotomous erythropoiesis in human embryonic organoids

We re-analyzed erythropoiesis of human embryonic organoids at day8 (D8, extraembryonic specification and three-germ layer stage), day15 (D15, hematopoiesis emerging stage) and day18 (D18, predominantly hematopoiesis stage) (Chao et al., 2023) (Fig. S1a). The surface markers dissects erythropoiesis process (Caulier & Sankaran, 2022). CD34 expression marks the early stage of hematopoiesis while ITGA4 an TFRC expression mark proerythroblast (proE) (Fig. 1a). Combining expression of SLC4A1 and GYPA, and loss of TIGA4 can track erythroid maturation, marking late-stage orthochromatic erythroblasts (orthoE) (Fig. 1a). We used CD34, ITGA4, TFRC, GYPA, and SLC4A1 to define erythropoiesis in human embryonic organoids (HEMOs) (Fig. 1b). The erythroid proportions were significantly different, with D8 containing more late-stage orthoE (Fig. 1c,d). The pseudotime analysis located D8 at later temporal order (Fig. 1e). We applied globin switch as authentic classification of erythropoiesis (Orkin, 1990). Embryonic erythropoiesis is marked by HBZ with accelerated erythroid generation. This is replaced with fetal erythropoiesis marked by HBG driven by progenitors erythro-myeloid progenitors (EMP) and hematopoietic progenitor cells (HPC) (M. H. Baron et al., 2012; King et al., 2021). D8 expressed embryonic globin HBZ however, D15 and D18 enriched fetal globin HBG1 and HBG2 (Fig. 1f). The developmental stage-specific modulator BCL11A regulates globin switching during ontogeny and gradually increases in fetal erythropoiesis, promoting differentiation and maturation of fetal erythroid (Masuda et al., 2016; Sankaran et al., 2008; Xu et al., 2013). We detected the high expression of BCL11A at D15 and D18, but no expression was detected at D8 (Fig. 1f). These indicate the dichotomous waves of erythropoiesis of embryonic to fetal. To comprehensively understand the heterogeneity of erythropoiesis within and between stages, we performed single-cell lineage tracing build on mitochondria variants as endogenous markers. We generated Chromium single-cell transcriptomics from D8, D15 and D18, and Visium Spatial transcriptomics from D15. We enriched mitochondrial transcriptome and amplified with MAESTER. We traced lineage with Chromium-MAESTER and Visium-MAESTER libraries (Fig. 1g). We benchmarked the quality of our libraries compared with the original developer (Fig. S1b,c).

We first assured the coverage of the Chromium- and Visium-MAESTER data sets. MAESTER amplified mitochondria variant reads per cell 137 - 223 from 20 - 28 in starting library (Fig. S1d,e) We next identified informative variants to guide clonal reconstruction. MQuad identified 44 (0.27%), 41 (0.25%), and 41 (0.25%) informative mitochondrial SNVs

in the D8, D15 and D18 samples respectively (Fig. 1h). We found 9 informative mitochondrial SNVs were shared among all stages. We identified 25, 20 and 18 stage specific SNVs as well (Fig. 1h). We excluded sequencing errors by ensuring alternative allele contained at least 266 reads (Table S1). To comprehensively capture the cellular correlation of each stage, we used the total mtDNA SNVs for clone assignment. We defined 3, 5 and 4 clones by unique combinations of barcodes and reconstructed lineage at D8, D15, and D18 respectively (Fig. 1i and Fig. S1f). At D8, 31%, 40% and 29% cells were assigned into Clone A0, Clone A1 and Clone A2 respectively. At D15, 4%, 47%,1%, 44% and 4% cells were assigned into Clone B0 to Clone B4 respectively. At D18, 17%, 41%, 35% and 8% cells were assigned into each clone.

The time-series clonality reconstruction captured human erythropoiesis

We further investigated the clonality built on single cell mtDNA SNVs at individual stage. At D8, alluvial plot demonstrated that all three clones contain various cell types, including fibroblasts (Fibro), pluripotent stem cells with ectoderm specification (PSC-Ect), erythroid cells (Ery), Ectoderm-like cells (Ect-like), EMP, pluripotent stem cells-like (PSC-like), endothelial cells (EC), mesoderm-like cells (Mes-like), cardiac mesoderm-like cells (CM-like), trophoblast-like cells (TB-like), monocytes (Mono), megakaryocytes (Mk), myeloid progenitor 1 (MP-1), yolk sac endoderm (YSE), and neural ectoderm-like cells (NE-like) (Fig. 2a-c, and Fig. S2a). Fibro, PSE-Ect and Ery are the predominant cell types across all three clones at D8 (Fig. 2a, and FigS2a). This coincided with our previous single cell transcriptomic landscape (Chao et al., 2023). MP, Mono and Mk more inflowed to Clone A0. High proportion of Ect-like, Fibro, EMP and PSC-Ect lineages inflowed to Clone A1. More CM-like, NE-like and TB-like lineages inflowed to Clone A2. EC, Ery, Mes-like and PSC-like lineage distributed evenly into all three clones at D8 (Fig. 2b). Clone A0 and Clone A1 displayed a high degree of similarity, while Clone A2 was more different from Clone A0 and A1 (Fig. 2c, Fig. S2b). All three clones enriched middle to late stage earlyE and orthoE with nearly absent proE (Fig. 2d-f). Clone A2 showed the highest proportion of late stage orthoE with least contribution of EMP.The other clones A0-2 were partially contributed by EMP This suggests the embryonic erythropoiesis accelerating erythroid production at D8 predominantly with clone A2 (Fig. 2g).

We next investigated expression of factors that indicate the stage of erythropoiesis. Krüppel-like factor (KLF1) as a master transcriptional regulator that commits EMPs to the erythroid lineage (Frontelo et al., 2007; Hodge et al., 2006; Isern et al., 2008; Siatecka & Bieker, 2011; Tallack et al., 2012). We detected high expression of KLF1 in Ery of all three clones at D8 stage (Fig. S2c), suggesting EMP transition into Ery (Fig. 2g). This indicates that all three clones contributed to fetal erythropoiesis driven by EMP. Solute carrier family 4 member 1 (SLC4A1), the surface marker defines orthoE (P. Huang et al., 2020; Lee et al., 2023), was significantly increased in Ery Clone A2 compared with Clone A0 and Clone A1 Fig. (Fig. S2c),. This again support our observation that orthoE dominated Clone A2. The hematopoietic transcriptional factors GATA transcription factor 1 and 2 (GATA1 and GATA2), RUNX1 and MYB suggest once again that EMP commitment to the erythroid lineage in all three clones (Bresnick et al., 2012; Fujiwara et al., 2004; Tsai et al., 1994; Walmsley et al., 1994). Altogether, D8 stage enriched embryonic erythropoiesis and some cells were flowing to fetal erythropoiesis (Fig. 2g).

Ery, EMP and Mk became the dominant cell types at D15 (Fig. 3a, and Fig. S3a). Clone B1 and B3 occupied nearly entire population. YSE cells belonged exclusively to Clone B1. Most Ery were assigned into Clone B3. PSC-Ect and Mes-like cells were predominantly assigned into Clone B4 (Fig. 3a-c). Overall, Clones B1 and B3 shared similar cell compositions

and represented most cell types at D15 (Fig. 3a-c). Visualizing the clonality in mtDNA SNV space, Clone B0 and Clone B4 were two distinguished population and were separated well from rest of the clones (Fig. S3b). Erythropoiesis showed a striking difference from D8. All five clones labelled early-stage proE, which was absent at D8. The late-stage orthoE became minor population, unlike its dominant presence at D8 (Fig. 3d,e). These data suggest a different mode of erythropoiesis between D8 and D15. Embryonic erythropoiesis accelerated the generation of erythroblasts. Fetal erythropoiesis prolonged the erythropoiesis as if they had to take a long track from immature hematopoietic progenitors (Palis, 2014; Weintraub et al., 1971). This is consistent with the presence of EMP and HPC at D15 (Fig. 3f). The observation in the Clone B4 further reinforces the possibility of the different mode of erythropoiesis between D8 and D15 time points (Fig. 2g vs. Fig. 3g). The Clone B4 labels a minor population with distinct cellular composition (Fig. S3b) with more late-stage orthoE alike seen at D8 (Fig. 2e vs. Fig. 3e). Taken together, the Clone B4 mostly contained reminiscent clones of the accelerated embryonic erythropoiesis seen at the earlier stage D8, but majority of erythroblasts at D15 came from immature hematopoietic progenitors EMP and HPC suggesting fetal erythropoiesis (Fig. 3g).

Consistently, the Clones B1 and B4 Ery express the orthoE marker SLC4A1 (Fig. S3c). These align with the relatively higher presence of orthoE in Clones B1 and B4 (Fig. 3d,f). High level of erythro-megakaryocyte factor GATA1 (Cantor, 2005) indicates the active generation of Ery and MK populations at this stage (Fig. S3c). EMP express both RUNX1 and MYB reinforcing that the population is still maintained while actively generating mature Ery and MK (Fig. S3c). Taken together, all five clones underwent fetal erythropoiesis from EMP and one Clone B4 enriched the reminiscent embryonic erythropoiesis (Fig. 3g).

The dominant cell types in four clones became Ery, EMP and Mk at D18 (Fig. 4a, and Fig. S4a). Clones C1 and C2 occupied 75% of population. Clone C1 generated Ery, PSC-Ect EMP and potential niche YSE. Clone C2 generated HPC, MP-1, MP-2, and potential niche Fibro and YSE. These suggest a clonal development of both hematopoietic and niche cells (Fig. 4a-c). Clone C3 contained a minority of population and distinct from other clones (Fig. S4b). Late-stage orthoE became less that suggests reminiscent clones of embryonic erythropoiesis seen at D15 were washed out (Fig. 4d,e). Clones C1 and C2 were clonally distinct (Fig. 4c). This coincided with that Clone C1 generated EMP and early-stage proE while Clone C2 generated less proE with slight presence of orthoE (Fig. 4d,e). Expression of SLC4A1 in C2 Ery indicates the presence of orthoE (Fig. S4c). The gradual increase of HPC through D15-18 correlated with the increase of MK (Fig. 4f,g). This is consistent with Mk-generation capacity directly attributed to hematopoietic stem cells (Rodriguez-Fraticelli et al., 2020; Sanjuan-Pla et al., 2013; Yamamoto et al., 2013). Indeed, the Clone C2 enriched both HPC and MK, but C1 did not do so (Fig. 4f). Taken together, two major clones predominate hematopoiesis at D18. One clone labelled by C1 commit to erythropoiesis driven by EMP. The other clone C2 generate various types of erythroids and myeloids driven by HPC (Fig. 4h). Both C1 and C2 labelled clones may produce their own niche cells too, which we addressed by spatial-MAESTER in the following section (Fig. 5). These data demonstrate that combining scRNA-seq with mitochondrial SNV as internal barcodes decipher lineage maps with higher resolution than conventional single-cell transcriptomics.

Spatial architecture of hematopoietic and niche cell development

We applied spatial transcriptomics combined with mitochondrial variants in order to dissect clonal architecture of hematopoietic and niche cells. We enriched mitochondrial

transcripts from 10x Visium full length cDNA and dissect its clonal structure with informative mtDNA SNVs identified by MQuad. We assigned one embryonic organoid 3 clones based on the unique combination of 10 informative mtDNA SNVs (Fig. 5a and Fig. S5a). We located the clonal spatial distribution into H&E images and annotated their population. Three clones were mixed and distributed in a cross-like pattern in the organoid, with each clone presented both on the edge and in the centre of the organoid and intermingled with each other (Fig. 5b). We spatially annotated YSE, EC, Fibro, EMP, HPC and Mk by generative model SpatialScope (Wan et al., 2023) (Fig. 5c and Fig. S5b). This suggests modelling fetal erythropoiesis in the yolk sac. The CellPhoneDB spotted the interaction of EMP and its potential niche cell type EC, as well as HPC and its potential niche cell type Fibro (Fig. 5d). We dissected EMP- EC and HPC-Fibro interactions by CellPhoneDB (Efremova et al., 2020). We identified a high interactive score of NOTCH ligands and receptors between HPC and Fibro (Fig. 5e), as well as EMP and EC (Fig. 5f). Intrigued by the significance of NOTCH signaling in developmental hematopoiesis (Ditadi et al., 2015; Hadland et al., 2015; Koch et al., 2013), we scored both activation and inhibition of NOTCH signaling defined by combination of ligand-receptors. EMP expressed DLK1 and EC expressed NOTCH1 in the same area (Fig. 5g,h). Meanwhile, HPC expressed DLK1 and Fibro expressed NOTCH1 in the same area. DLK1 controls hematopoietic progenitor pool by suppressing their proliferation and differentiation (Li et al., 2005; Mirshekar-Syahkal et al., 2013; Qian et al., 2016). We achieved a powerful application of mitochondrial variants to increase a resolution of lineage mapping. We identified clonal architecture of early human hematopoiesis, from multiple waves of erythropoiesis to clonal development of hematopoietic and niche cells in the yolk sac. Notably, heterologous mitochondrial variants preceded in human PSCs serve a good platform to model organogenesis. In this study, we refined the early erythropoiesis that comprised of multiple waves. We further identified the potential role of NOTCH signaling in co-development of hematopoietic cells and their niche in the human embryonic organoid models.

Discussion:

Our study has demonstrated the potential of using mitochondrial variants as a natural genetic barcode for clonal tracking in stem cell-derived organoids. By combining mitochondrial variant enrichment with robust computational mitochondrial variant calling approaches, we were able to reconstruct the lineage history during HEMO development and observe the fate map underlying embryogenesis. Moreover, we have innovatively expanded the application of mitochondrial variant enrichment technology into spatial transcriptomics to explore the interplay of molecular and spatial continuums of HEMO clones. This gave us deep insight into spatial growth patterns, the intrinsic and extrinsic features of HEMO clones, as well as histogenomic relationships in HEMO development. Our use of mitochondrial variants for single cell lineage tracing has allowed us to characterize the attributes connecting large-scale developmental phenomena to the molecular regulation of cell fate decision-making, highlighting the potential of lineage information in facilitating the identification of key developmental events.

Sequencing coverage is critical for accurate variant calling and annotation in next-generation sequencing data (Koboldt, 2020; Song et al., 2016; Zverinova & Guryev, 2022). Considering insufficient coverage can lead to variant missing, we firstly investigated the coverage of mitochondrial genome in our Chromium- and Visium- MAESTER data sets. Consistently, we found that coverage per cell of mitochondrial genome was dramatically increased in all four MAESTER datasets compared with single cell ones, with mean of total

coverage of 20 to 28 for RNA-seq and 137 to 223 for MAESTER per cell (Fig. S1d,e), enabling reliable mtDNA variant calling and downstream clonality analysis.

Being able to map out clonal growth patterns and identify clone-specific genetic traits will play a crucial role in shedding light on the various development processes that drive haematopoiesis in the embryo. The emergence of spatially resolved genomic information has revolutionized our understanding of cancer progression and clonal evolution (Lomakin et al., 2022; Yang et al., 2022; Zhao et al., 2022). In particular, copy number variation (CNV) inference from single-cell RNA sequencing (scRNA-seq) has been widely employed to characterize malignant cells from tumors, and spatial transcriptomic analysis has adapted the CNV inference algorithm to provide robust spatial clonal evolution analysis (Erickson et al., 2022; Ji et al., 2020). In context of development, by characterizing the spatial distribution of cells, we can gain insights into clonal relationships and cell fate decisions (Chao et al., 2023). This approach can also be used to investigate the role of spatial cues in directing differentiation and tissue patterning. In our study, with integration of mitochondrial enrichment and spatial transcriptomics approaches, we found the spatial architecture of different clones at hematopoietic emerging stage and discovered the EMP-Endo niche for self-organization in embryo development through notch signalling pathways. Tissue organization controls development process. The ability to accurately locate and molecularly characterize different HEMO clones is a crucial addition to the spatial-omics toolkit. It provides a strong developmental framework that is essential for interpreting the biological significance of various spatial characteristics of hematopoiesis, which are often highly dynamic.

Furthermore, our study has highlighted the potential of stem cell-derived organoids for modeling human embryonic development and diseases. Combining lineage tracing and spatial transcriptomics approaches further advance our understanding of the complex cellular dynamics and interactions that occur during organoid development and function. Our findings on PD-L1 and macrophage development from HEMO suggest the potential of stem cell-derived organoids in advancing our understanding of disease and developing new treatments (Cao et al., 2022). It is well-established that hematopoiesis in the human embryo occurs in three waves, with the first two waves generating transient populations of blood cells and the third wave giving rise to the definitive hematopoietic stem cells that produce all blood cells throughout life (M. H. Baron et al., 2012; Dzierzak & Philipsen, 2013; Dzierzak & Speck, 2008; Lux et al., 2008). With the capacity to give rise to both embryonic and extraembryonic tissues, PSCs have the potential to revolutionize the field of regenerative medicine and provide new insights into the mechanisms underlying human development and disease.

However, our study could be further improved by incorporating additional genetic markers for lineage tracing during HEMO development. Although mitochondrial variants have proven to be a useful genetic barcode for clonal tracking, there may be limitations to their use due to the uneven distribution of mitochondria during cell division. Therefore, incorporating additional genetic markers such as somatic mutations in nuclear DNA or transposon insertions could provide more robust and accurate information for clonal tracking. Overall, our study has demonstrated the potential of using mitochondria variants as endogenous maker for clonal tracking in stem cell-derived organoids system and has provided valuable insights into the mechanisms underlying human development and disease. Future research could build on these findings to further explore the potential of stem cell-derived organoids for modeling human biology and disease.

Methods

HEMO differentiation

HEMO differentiation was done based on the previous established protocol (Chao et al., 2023). In brief, we formed EB from human PSCs and differentiated to mesoderm and hematoendothelial lineage by morphogens and cytokines.

hPSCs line. Human M1-hPSCs were gifted by Pentao Liu and maintained in the media described before (Gao et al., 2019).

hPSCs maintenance and pre-differentiation. The hPSCs were cultured and maintained in PSC medium, with medium changes occurring every other day. The composition of the medium was previously reported (Gao et al., 2019). Prior to the formation of embryoid bodies (EBs), cells were pre-differentiated in KSR medium (DMEM/F12 + 10% KSR) (Thermo Scientific, catalog no. 11320033; Thermo Scientific, catalog no. 10828028) for a period of three days.

EB formation in hanging drop. To begin the process, the KSR medium was removed and the cells were washed with PBS (Thermo Scientific, catalog no. 10010023). The cells were then digested with 500 mL of 0.05% Trypsin (Thermo Scientific, catalog no. 25300054) at 37°C for three minutes, but no longer than seven minutes. The Trypsin was removed and 2 mL of KSR medium was added to harvest the cells. Careful handling of the cells was necessary, and harsh pipetting was to be avoided. The cells were then centrifuged at 300g for three minutes, and the supernatant was carefully removed. To resuspend the cells, 1 mL of KSR medium and 1 μL of Y27632 (10 ug/ml) (Tocris, catalog no. 1254/10) were added. For each 25 μL hanging drop on the cap of a 10 cm petri dish, 4,000 cells were added. Thirty to forty drops were made for each cap, and the dish was filled with PBS to keep it moist. The cap was then gently and slightly inverted to cover the dish, and the dishes were labeled and kept at 37°C for three days.

EB collection and differentiation. After the formation of embryoid bodies (EBs), all EBs were collected and washed with PBS. The EBs were then centrifuged at 100g for one minute, and the supernatant was carefully removed. 1 mL of medium A (STEMdiff Hematopoietic Kit, catalog no. 05310) was added, and the EBs were transferred to a non-adherent 24-well plate. The date was labeled as Day 0, and the EBs were cultured with 1 mL of STEMdiff medium A for three days. On Day 2, the medium was half-changed, and the EBs were subsequently cultured with 1 mL of STEMdiff medium B for the following days. This process involved the maintenance of hPSCs, pre-differentiation, EB formation, collection, and differentiation.

10x Genomics Chromium scRNA-seq

To prepare single-cell suspensions, organoids were harvested at D8, D15, D18, D25 and D32 since EB formation in hanging drop. After washing with PBS, organoids were mechanical chopped with scissors for 20-30 times, followed by digestion with 500 μL Accumax at 37 °C for 10-15 minutes. Digestion was terminated by 500 μL PBS containing 2% FBS. Cells were filtered with 40 μm cell filter and centrifuged at 500 × g for 5 min. The collected cells were resuspended in FACS sorting buffer (1 × PBS with 2% FBS) and their concentration was adjusted to around 3k cells per μL by counting with a hemocytometer.

Subsequently, cells were stained with DAPI (BD Biosciences, catalog no. 564907) in 1:100 for 5 min at 4 °C and were sorted by BD Influx flow cytometry in CPOS at HKUMed. Cells were gated to exclude dead cells and doublets and collected in a chilled PBS with 0.04% BSA for scRNA-seq library construction. Cell concentration was adjusted to around 500-1000 cells per μL.

10x Genomics 3′ Single Cell Gene Expression v3 reagents were used for scRNA-seq library preparation. Single cell encapsulation and reverse transcription were performed at the Centre for PanorOmic Sciences, the University of Hong Kong. Briefly, 8-16k live single cells of size 30 μm or smaller and of good viability were encapsulated in gel bead-in emulsions. Captured mRNAs were transcribed into cDNAs and amplified to generate whole-transcriptome amplifications (WTAs). Libraries were constructed by fragmentation, adapter ligation and a sample index PCR. Libraries were sequenced using Illumina Novaseq 6000 for Pair-End 151bp sequencing.

10x Genomics Visium library construction and sequencing

The Visium Spatial Tissue Optimization Slide & Reagent kit (10x Genomics, catalog no. PN-1000193) was used to optimize permeabilization conditions for the tissue sections. The Visium Spatial Gene Expression Slide & Reagent kit (10X Genomics, catalog no. PN-1000184) was used to generate spatially barcoded cDNA from every 10-μm sections of D15 HEMO samples. Library construction kit (10 x Genomics, catalog no. PN-1000190) was used for library construction. Frozen HEMO sample preparation, RNA integrity detection, tissue optimization, staining and imaging, cDNA synthesis and second strand synthesis, cDNA amplification and library construction were carried out according to Chao et al. (Chao et al., 2023). Libraries were sequenced on Illumina sequencing instruments using TruSeq Read 1 and TruSeq Read 2 primers.

Mitochondrial alteration enrichment for Chromium-MAESTER and Visium-MAESTER

MAESTER was performed according to Miller et al. (Miller et al., 2022). Briefly, WTAs yield in high-throughput 10x Genomics scRNA-seq 3’ and Visium platforms served as initial templates for the enrichment of 15 mitochondrial transcripts by two rounds of PCR. Sequences of the amplified mitochondrial transcripts were obtained by standard next-generation sequencing.

In PCR1, twelve primer mixes were created using the designed primers tiled across the entire mitochondrial transcriptome (Table 1). A barcoded sample indexing i5 primer was included in each mix in which 20 ng WTAs were included. PCR was performed using the following conditions: denaturation at 95 ℃ for 3 minutes, followed by 6 cycles of 98 ℃ for 20 seconds, 65 ℃ for 15 seconds, and 72 ℃ for 3 minutes, ending with a final extension at 72 ℃ for 5 minutes. Following amplification, the PCR products were pooled in a certain ratio and purified with 1x AMPure XP beads (Beckman Coulter A63881).

The purified products were used as the template for PCR2 in which the Illumina adapters (P5, P7), dual index barcodes to identify the sample (i5, i7), and sequencing primer binding sites to the fragments were added. The programme of PCR2 was an initial denaturation at 95 ℃ for 3 minutes, then 6 cycles of 98 ℃ for 20 seconds, 60 ℃ for 30 seconds, and 72 ℃ for 3 minutes, and then a final extension at 72 ℃ for 5 minutes. After PCR2, the DNA was purified with 0.8x AMPure XP beads. The library concentration of three Chromium- and Visium-

MAESTER samples were 50, 59, 71 and 21 ng/mL respectively. Resulting MAESTER and Visium-MAESTER libraries were sequenced on the Illumina NovaSeq SP PE150 kit with run cycle 28, 8, 8, 256 for Read 1, i7, i5 and Read 2 respectively. The sequences of Nove seq primers are listed in Table 2.

Data analysis and statistics

scRNA-seq data processing and cell filtering. The sequencing data obtained from 10X Genomics was processed using CellRanger software (version 6.0.0) with the GRCh38 human reference genome and default settings for both read alignment and initial cell calling. The resulting cell by gene count matrix was loaded into the Seurat package (version 4.1.0) in R (Hao et al., 2021). Cells with fewer than 1,000 detected genes and with more than 15% mitochondrial gene expression were excluded. Additionally, genes that were expressed in fewer than 3 cells were removed. To detect doublets in an individual sample, the DoubletFinder package (version 2.0.3) was utilized with fine-tuned parameters (McGinnis et al., 2019).

Data integration, clustering, and annotation of scRNA-seq data. Following quality control and removal of doublets, the scRNA-seq datasets were integrated using fastMNN (version 1.6.3) (Haghverdi et al., 2018). High variable genes (HVGs) were identified and used for principal component analysis (PCA) to reduce dimensionality and perform clustering, all through the Seurat pipeline with default settings. Similarly, markers for each cluster were identified using the FindAllMarkers function and utilized for cluster annotation. The complete marker gene list can be found in the supplementary files.

Diffusion pseudotime analysis of scRNA-seq data. The connectivity of individual single-cell clusters was initially measured and visualized using a partition-based graph abstraction graph (PAGA graph) in Scanpy (version 1.9.1) (Gao et al., 2019; Wolf et al., 2018). To estimate pseudotime, an extended version of diffusion pseudotime (DPT) was employed (Haghverdi et al., 2016), and cells were embedded in force-directed graphs.

10x Visium spatial transcriptomics processing. The sequencing data and imaging tiff files of individual 10x Visium slides were processed using SpaceRanger software (version 1.3.1) with the GRCh38 human reference genome. Count matrices were then loaded into the Scanpy package (version 4.1.0) in Python. Spots were selected based on RNA counts and mitochondrial gene percentage. After spot selection, high variable genes (HVGs) were identified and used for principal component analysis (PCA) to reduce dimensionality and perform clustering. The Scanpy adata objects were utilized for downstream analysis individually.

Spatial transcriptomics spot deconvolution. Two packages were utilized for spatial transcriptomics spot deconvolution. The first method applied was SpatialScope, a newly developed statistical method that integrates scRNA-seq and spatial transcriptomics data to obtain the spatial distribution of the whole transcriptome at the single-cell resolution. SpatialScope initially performs nuclei segmentation to count and locate nuclei within each spatial spot. It then assigns a cell type to each located nucleus. Finally, by leveraging the paired scRNA-seq reference data and a learned deep generative model, it decomposes gene expression at each spot into gene expression of the individual cells located within the spot. The RCTD package (version 2.0.0) was also utilized to validate the cell proportion (Cable et al., 2022). It

first takes the single-cell reference and predicts the proportion of each cell type in an individual 10x Visium spot.

Cell interaction within spatial transcriptomics data. Two different packages were utilized for spatial cell-cell interaction analysis. After obtaining single-cell resolution spatial data, CellPhoneDB within the Squidpy toolbox was utilized to predict cellular ligand-receptor interactions (Efremova et al., 2020; Palla et al., 2022). All predicted pairs were pre-selected based on a p-value of 0.001. The enriched score and metadata of individual pairs were then examined, and the final results were presented in a bar plot.

Mitochondrial variant calling

Data pre-processing was performed before mitochondrial variant calling. After getting the fastq files of MAESTER and Visium-MAESTER data, umitools were first used to generate a whitelist of valid barcodes. The number of cells was estimated by the data. Based on the whitlist of valida barcodes, umitools extracted reads coming from real cells and moved UMIs and barcodes from sequence lines to the head lines. In this step, background reads and barcodes with low number of reads were filtered. Reads were aligned to the reference genome by STAR using hg38 as the reference genome. UMIs and barcodes were added as tags in the bam file by using pysam. Mitochondrial SNVs were called by cellSNV-lite mode 2a. We used minimum minor allele frequency 0.1 and minimum aggragated count as filtering criterion. MQuad(Kwok et al., 2022) was applied to call informative mitochondrial SNVs for the following analysis.

Clonality reconstruction

VireoSNP (Y. Huang et al., 2019) was used for clone assignment. We first used elbow plot to define several candidate optimal clone number based on the 'Mito Clones' tutorial. Following the same tutorial, we then plotted the assignment probability plots and mean allele ratio plots starting from the lowest number of candidate clones. We chose clone numbers that gave us the highest resolution of clonality with high confidence. We kept as many clones as possible while made sure each clone had its unique patterns of allele frequency regarding to the allele ratios. We also made sure that each clone has high assignment probability of each cell. We then used the 'anno_heat' function from VireoSNP to plot the heat map of each cell's allele frequency.

Code availability

The codes for analysis is available at https://github.com/CHAOYiming/HEMO_MAESTER.

Acknowledgement

We appreciate Dr Peter van Galen at Harvard Medical School for generous sharing of bioanalyzer positive library controls, critical discussion on this project, and comments to the manuscript. We thank CPOS at HKUMed for technical assistance in 10x Chromium, 10x Visium, and Novaseq. This study is supported in part by ECS 27109921 (R.S.) and GRF 17109022 (R.S.) by the Hong Kong Research Grant Council, and the Shenzhen-Hong Kong-Macau Technology Research Programme (Type C; SGDX2021082310356025 to J.W.K.H.). It is also supported in part by AIR@InnoHK (Laboratory of Data Discovery for Health for J.W.K.H) and Health@InnoHK (Centre for Translational Stem Cell Biology for R.S.) administered by Innovation and Technology Commission of Hong Kong.

Author contributions

Y.X, Y.C., J.W.K.H and R.S. conceived the study. Y.X. conducted experiments and generated the data. Y.C. conducted most informatics analysis with the guidance of Y.H. and J.W.K.H. Y.X and X.L. participated informatics analysis. Y.X., Y.C, J.W.K.H. and R.S. wrote the manuscript. All authors edited the manuscript.

References:

Acar, M., Kocherlakota, K. S., Murphy, M. M., Peyer, J. G., Oguro, H., Inra, C. N., Jaiyeola, C., Zhao, Z., Luby-Phelps, K., & Morrison, S. J. (2015). Deep imaging of bone marrow shows non-dividing stem cells are mainly perisinusoidal. *Nature*, *526*(7571), 126–130. https://doi.org/10.1038/nature15250

Baron, C. S., & van Oudenaarden, A. (2019). Unravelling cellular relationships during development and regeneration using genetic lineage tracing. *Nature Reviews Molecular Cell Biology*, *20*(12), 753–765. https://doi.org/10.1038/s41580-019-0186-3

Baron, M. H., Isern, J., & Fraser, S. T. (2012). The embryonic origins of erythropoiesis in mammals. *Blood*, *119*(21), 4828–4837. https://doi.org/10.1182/blood-2012-01-153486

Beneyto-Calabuig, S., Merbach, A. K., Kniffka, J.-A., Antes, M., Szu-Tu, C., Rohde, C., Waclawiczek, A., Stelmach, P., Gräßle, S., Pervan, P., Janssen, M., Landry, J. J. M., Benes, V., Jauch, A., Brough, M., Bauer, M., Besenbeck, B., Felden, J., Bäumer, S., … Velten, L. (2023). Clonally resolved single-cell multi-omics identifies routes of cellular differentiation in acute myeloid leukemia. *Cell Stem Cell*, *30*(5), 706-721.e8. https://doi.org/10.1016/j.stem.2023.04.001

Bowling, S., Sritharan, D., Osorio, F. G., Nguyen, M., Cheung, P., Rodriguez-Fraticelli, A., Patel, S., Yuan, W.-C., Fujiwara, Y., Li, B. E., Orkin, S. H., Hormoz, S., & Camargo, F. D. (2020). An Engineered CRISPR-Cas9 Mouse Line for Simultaneous Readout of Lineage Histories and Gene Expression Profiles in Single Cells. *Cell*, *181*(6), 1410-1422.e27. https://doi.org/10.1016/j.cell.2020.04.048

Bresnick, E. H., Katsumura, K. R., Lee, H.-Y., Johnson, K. D., & Perkins, A. S. (2012). Master regulatory GATA transcription factors: Mechanistic principles and emerging links to hematologic malignancies. *Nucleic Acids Research*, *40*(13), 5819–5831. https://doi.org/10.1093/nar/gks281

Cable, D. M., Murray, E., Zou, L. S., Goeva, A., Macosko, E. Z., Chen, F., & Irizarry, R. A. (2022). Robust decomposition of cell type mixtures in spatial transcriptomics. *Nature Biotechnology*, *40*(4), 517–526. https://doi.org/10.1038/s41587-021-00830-w

Cantor, A. B. (2005). GATA transcription factors in hematologic disease. *International Journal of Hematology*, *81*(5), 378–384. https://doi.org/10.1532/ijh97.04180

Cao, H., Xiang, Y., Zhang, S., Chao, Y., Guo, J., Ho, J. W. K., Huang, Y., Liu, P., & Sugimura, R. (2022). *PD-L1 regulates inflammatory macrophage development from human pluripotent stem cells by maintaining interferon-gamma signal* [Preprint]. Immunology. https://doi.org/10.1101/2022.12.14.520176

Caulier, A. L., & Sankaran, V. G. (2022). Molecular and cellular mechanisms that regulate human erythropoiesis. *Blood*, *139*(16), 2450–2459. https://doi.org/10.1182/blood.2021011044

Challen, G. A., Boles, N. C., Chambers, S. M., & Goodell, M. A. (2010). Distinct hematopoietic stem cell subtypes are differentially regulated by TGF-beta1. *Cell Stem Cell*, *6*(3), 265–278. https://doi.org/10.1016/j.stem.2010.02.002

Chan, M. M., Smith, Z. D., Grosswendt, S., Kretzmer, H., Norman, T. M., Adamson, B., Jost, M., Quinn, J. J., Yang, D., Jones, M. G., Khodaverdian, A., Yosef, N., Meissner, A., & Weissman, J. S.

(2019). Molecular recording of mammalian embryogenesis. *Nature*, *570*(7759), 77–82. https://doi.org/10.1038/s41586-019-1184-5

Chao, Y., Xiang, Y., Xiao, J., Zheng, W., Ebrahimkhani, M. R., Yang, C., Wu, A. R., Liu, P., Huang, Y., & Sugimura, R. (2023). Organoid-based single-cell spatiotemporal gene expression landscape of human embryonic development and hematopoiesis. *Signal Transduction and Targeted Therapy*, *8*(1), 230. https://doi.org/10.1038/s41392-023-01455-y

Chen, J. Y., Miyanishi, M., Wang, S. K., Yamazaki, S., Sinha, R., Kao, K. S., Seita, J., Sahoo, D., Nakauchi, H., & Weissman, I. L. (2016). Hoxb5 marks long-term haematopoietic stem cells and reveals a homogenous perivascular niche. *Nature*, *530*(7589), 223–227. https://doi.org/10.1038/nature16943

Christodoulou, C., Spencer, J. A., Yeh, S.-C. A., Turcotte, R., Kokkaliaris, K. D., Panero, R., Ramos, A., Guo, G., Seyedhassantehrani, N., Esipova, T. V., Vinogradov, S. A., Rudzinskas, S., Zhang, Y., Perkins, A. S., Orkin, S. H., Calogero, R. A., Schroeder, T., Lin, C. P., & Camargo, F. D. (2020). Live-animal imaging of native haematopoietic stem and progenitor cells. *Nature*, *578*(7794), 278–283. https://doi.org/10.1038/s41586-020-1971-z

Deng, Y., Bartosovic, M., Kukanja, P., Zhang, D., Liu, Y., Su, G., Enninful, A., Bai, Z., Castelo-Branco, G., & Fan, R. (2022). Spatial-CUT&Tag: Spatially resolved chromatin modification profiling at the cellular level. *Science (New York, N.Y.)*, *375*(6581), 681–686. https://doi.org/10.1126/science.abg7216

Ditadi, A., Sturgeon, C. M., Tober, J., Awong, G., Kennedy, M., Yzaguirre, A. D., Azzola, L., Ng, E. S., Stanley, E. G., French, D. L., Cheng, X., Gadue, P., Speck, N. A., Elefanty, A. G., & Keller, G. (2015). Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. *Nature Cell Biology*, *17*(5), 580–591. https://doi.org/10.1038/ncb3161

Dzierzak, E., & Philipsen, S. (2013). Erythropoiesis: Development and differentiation. *Cold Spring Harbor Perspectives in Medicine*, *3*(4), a011601. https://doi.org/10.1101/cshperspect.a011601

Dzierzak, E., & Speck, N. A. (2008). Of lineage and legacy: The development of mammalian hematopoietic stem cells. *Nature Immunology*, *9*(2), 129–136. https://doi.org/10.1038/ni1560

Efremova, M., Vento-Tormo, M., Teichmann, S. A., & Vento-Tormo, R. (2020). CellPhoneDB: Inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nature Protocols*, *15*(4), 1484–1506. https://doi.org/10.1038/s41596-020-0292-x

Erickson, A., He, M., Berglund, E., Marklund, M., Mirzazadeh, R., Schultz, N., Kvastad, L., Andersson, A., Bergenstråhle, L., Bergenstråhle, J., Larsson, L., Alonso Galicia, L., Shamikh, A., Basmaci, E., Díaz De Ståhl, T., Rajakumar, T., Doultsinos, D., Thrane, K., Ji, A. L., … Lundeberg, J. (2022). Spatially resolved clonal copy number alterations in benign and malignant tissue. *Nature*, *608*(7922), 360–367. https://doi.org/10.1038/s41586-022-05023-2

Frontelo, P., Manwani, D., Galdass, M., Karsunky, H., Lohmann, F., Gallagher, P. G., & Bieker, J. J. (2007). Novel role for EKLF in megakaryocyte lineage commitment. *Blood*, *110*(12), 3871–3880. https://doi.org/10.1182/blood-2007-03-082065

Fujiwara, Y., Chang, A. N., Williams, A. M., & Orkin, S. H. (2004). Functional overlap of GATA-1 and GATA-2 in primitive hematopoietic development. *Blood*, *103*(2), 583–585. https://doi.org/10.1182/blood-2003-08-2870

Gao, X., Nowak-Imialek, M., Chen, X., Chen, D., Herrmann, D., Ruan, D., Chen, A. C. H., Eckersley-Maslin, M. A., Ahmad, S., Lee, Y. L., Kobayashi, T., Ryan, D., Zhong, J., Zhu, J., Wu, J., Lan, G., Petkov, S., Yang, J., Antunes, L., … Liu, P. (2019). Establishment of porcine and human expanded potential stem cells. *Nature Cell Biology*, *21*(6), 687–699. https://doi.org/10.1038/s41556-019-0333-2

Gazit, R., Mandal, P. K., Ebina, W., Ben-Zvi, A., Nombela-Arrieta, C., Silberstein, L. E., & Rossi, D. J. (2014). Fgd5 identifies hematopoietic stem cells in the murine bone marrow. *The Journal of Experimental Medicine*, *211*(7), 1315–1331. https://doi.org/10.1084/jem.20130428

Goh, I., Botting, R. A., Rose, A., Webb, S., Engelbert, J., Gitton, Y., Stephenson, E., Quiroga Londoño, M., Mather, M., Mende, N., Imaz-Rosshandler, I., Yang, L., Horsfall, D., Basurto-Lozada, D., Chipampe, N.-J., Rook, V., Lee, J. T. H., Ton, M.-L., Keitley, D., … Haniffa, M. (2023). Yolk sac cell atlas reveals multiorgan functions during human early development. *Science*, *381*(6659), eadd7564. https://doi.org/10.1126/science.add7564

Gomez Perdiguero, E., Klapproth, K., Schulz, C., Busch, K., Azzoni, E., Crozet, L., Garner, H., Trouillet, C., de Bruijn, M. F., Geissmann, F., & Rodewald, H.-R. (2015). Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. *Nature*, *518*(7540), 547–551. https://doi.org/10.1038/nature13989

Grosso, M., Sessa, R., Puzone, S., Rosaria, M., & Izzo, P. (2012). Molecular Basis of Thalassemia. In D. Silverberg (Ed.), *Anemia*. InTech. https://doi.org/10.5772/31362

Gupta, R., Kanai, M., Durham, T. J., Tsuo, K., McCoy, J. G., Kotrys, A. V., Zhou, W., Chinnery, P. F., Karczewski, K. J., Calvo, S. E., Neale, B. M., & Mootha, V. K. (2023). Nuclear genetic control of mtDNA copy number and heteroplasmy in humans. *Nature*, *620*(7975), 839–848. https://doi.org/10.1038/s41586-023-06426-5

Hadland, B. K., Varnum-Finney, B., Poulos, M. G., Moon, R. T., Butler, J. M., Rafii, S., & Bernstein, I. D. (2015). Endothelium and NOTCH specify and amplify aorta-gonad-mesonephros-derived hematopoietic stem cells. *The Journal of Clinical Investigation*, *125*(5), 2032–2045. https://doi.org/10.1172/JCI80137

Haghverdi, L., Büttner, M., Wolf, F. A., Buettner, F., & Theis, F. J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. *Nature Methods*, *13*(10), 845–848. https://doi.org/10.1038/nmeth.3971

Hansen, M., Von Lindern, M., Van Den Akker, E., & Varga, E. (2019). Human-induced pluripotent stem cell-derived blood products: State of the art and future directions. *FEBS Letters*, *593*(23), 3288–3303. https://doi.org/10.1002/1873-3468.13599

He, Z., Maynard, A., Jain, A., Gerber, T., Petri, R., Lin, H.-C., Santel, M., Ly, K., Dupré, J.-S., Sidow, L., Sanchis Calleja, F., Jansen, S. M. J., Riesenberg, S., Camp, J. G., & Treutlein, B. (2022). Lineage recording in human cerebral organoids. *Nature Methods*, *19*(1), 90–99. https://doi.org/10.1038/s41592-021-01344-8

Hislop, J., Alavi, A., Song, Q., Schoenberger, R., Keshavarz F., K., LeGraw, R., Velazquez, J., Mokhtari, T., Taheri, M. N., Rytel, M., De Sousa Lopes, S. M. C., Watkins, S., Stolz, D., Kiani, S., Sozen, B., Bar-Joseph, Z., & Ebrahimkhani, M. R. (2023). *Modelling Human Post-Implantation Development via Extra-Embryonic Niche Engineering* [Preprint]. Developmental Biology. https://doi.org/10.1101/2023.06.15.545118

Hodge, D., Coghill, E., Keys, J., Maguire, T., Hartmann, B., McDowall, A., Weiss, M., Grimmond, S., & Perkins, A. (2006). A global role for EKLF in definitive and primitive erythropoiesis. *Blood*, *107*(8), 3359–3370. https://doi.org/10.1182/blood-2005-07-2888

Huang, P., Zhao, Y., Zhong, J., Zhang, X., Liu, Q., Qiu, X., Chen, S., Yan, H., Hillyer, C., Mohandas, N., Pan, X., & Xu, X. (2020). Putative regulators for the continuum of erythroid differentiation revealed by single-cell transcriptome of human BM and UCB cells. *Proceedings of the National Academy of Sciences*, *117*(23), 12868–12876. https://doi.org/10.1073/pnas.1915085117

Huang, Y., McCarthy, D. J., & Stegle, O. (2019). Vireo: Bayesian demultiplexing of pooled single-cell RNA-seq data without genotype reference. *Genome Biology*, *20*(1), 273. https://doi.org/10.1186/s13059-019-1865-2

Ihry, R. J., Worringer, K. A., Salick, M. R., Frias, E., Ho, D., Theriault, K., Kommineni, S., Chen, J., Sondey, M., Ye, C., Randhawa, R., Kulkarni, T., Yang, Z., McAllister, G., Russ, C., Reece-Hoyes, J., Forrester, W., Hoffman, G. R., Dolmetsch, R., & Kaykas, A. (2018). P53 inhibits CRISPR–

Cas9 engineering in human pluripotent stem cells. *Nature Medicine*, *24*(7), 939–946. https://doi.org/10.1038/s41591-018-0050-6

Isern, J., Fraser, S. T., He, Z., & Baron, M. H. (2008). The fetal liver is a niche for maturation of primitive erythroid cells. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(18), 6662–6667. https://doi.org/10.1073/pnas.0802032105

Ji, A. L., Rubin, A. J., Thrane, K., Jiang, S., Reynolds, D. L., Meyers, R. M., Guo, M. G., George, B. M., Mollbrink, A., Bergenstråhle, J., Larsson, L., Bai, Y., Zhu, B., Bhaduri, A., Meyers, J. M., Rovira-Clavé, X., Hollmig, S. T., Aasi, S. Z., Nolan, G. P., … Khavari, P. A. (2020). Multimodal Analysis of Composition and Spatial Architecture in Human Squamous Cell Carcinoma. *Cell*, *182*(2), 497-514.e22. https://doi.org/10.1016/j.cell.2020.05.039

Kagawa, H., Javali, A., Khoei, H. H., Sommer, T. M., Sestini, G., Novatchkova, M., Scholte Op Reimer, Y., Castel, G., Bruneau, A., Maenhoudt, N., Lammers, J., Loubersac, S., Freour, T., Vankelecom, H., David, L., & Rivron, N. (2022). Human blastoids model blastocyst development and implantation. *Nature*, *601*(7894), 600–605. https://doi.org/10.1038/s41586-021-04267-8

Kalhor, R., Kalhor, K., Mejia, L., Leeper, K., Graveline, A., Mali, P., & Church, G. M. (2018). Developmental barcoding of whole mouse via homing CRISPR. *Science*, *361*(6405), eaat9804. https://doi.org/10.1126/science.aat9804

Kebschull, J. M., & Zador, A. M. (2018). Cellular barcoding: Lineage tracing, screening and beyond. *Nature Methods*, *15*(11), 871–879. https://doi.org/10.1038/s41592-018-0185-x

Kester, L., & van Oudenaarden, A. (2018). Single-Cell Transcriptomics Meets Lineage Tracing. *Cell Stem Cell*, *23*(2), 166–179. https://doi.org/10.1016/j.stem.2018.04.014

King, A. J., Songdej, D., Downes, D. J., Beagrie, R. A., Liu, S., Buckley, M., Hua, P., Suciu, M. C., Marieke Oudelaar, A., Hanssen, L. L. P., Jeziorska, D., Roberts, N., Carpenter, S. J., Francis, H., Telenius, J., Olijnik, A.-A., Sharpe, J. A., Sloane-Stanley, J., Eglinton, J., … Babbs, C. (2021). Reactivation of a developmentally silenced embryonic globin gene. *Nature Communications*, *12*(1), 4439. https://doi.org/10.1038/s41467-021-24402-3

Koboldt, D. C. (2020). Best practices for variant calling in clinical sequencing. *Genome Medicine*, *12*(1), 91. https://doi.org/10.1186/s13073-020-00791-w

Koch, U., Lehal, R., & Radtke, F. (2013). Stem cells living with a Notch. *Development (Cambridge, England)*, *140*(4), 689–704. https://doi.org/10.1242/dev.080614

Kwok, A. W. C., Qiao, C., Huang, R., Sham, M.-H., Ho, J. W. K., & Huang, Y. (2022). MQuad enables clonal substructure discovery using single cell mitochondrial variants. *Nature Communications*, *13*(1), 1205. https://doi.org/10.1038/s41467-022-28845-0

Lareau, C. A., Ludwig, L. S., Muus, C., Gohil, S. H., Zhao, T., Chiang, Z., Pelka, K., Verboon, J. M., Luo, W., Christian, E., Rosebrock, D., Getz, G., Boland, G. M., Chen, F., Buenrostro, J. D., Hacohen, N., Wu, C. J., Aryee, M. J., Regev, A., & Sankaran, V. G. (2021). Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling. *Nature Biotechnology*, *39*(4), 451–461. https://doi.org/10.1038/s41587-020-0645-6

Lee, S.-J., Jung, C., Oh, J. E., Kim, S., Lee, S., Lee, J. Y., & Yoon, Y. (2023). Generation of Red Blood Cells from Human Pluripotent Stem Cells—An Update. *Cells*, *12*(11), 1554. https://doi.org/10.3390/cells12111554

Lemischka, I. R., Raulet, D. H., & Mulligan, R. C. (1986). Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell*, *45*(6), 917–927. https://doi.org/10.1016/0092-8674(86)90566-0

Li, L., Bowling, S., Yu, Q., McGeary, S. E., Alcedo, K., Lemke, B., Ferreira, M., Klein, A. M., Wang, S.-W., & Camargo, F. D. (2023). *A mouse model with high clonal barcode diversity for joint lineage, transcriptomic, and epigenomic profiling in single cells* [Preprint]. Developmental Biology. https://doi.org/10.1101/2023.01.29.526062

Li, L., Forman, S. J., & Bhatia, R. (2005). Expression of DLK1 in hematopoietic cells results in inhibition of differentiation and proliferation. *Oncogene*, *24*(27), 4472–4476. https://doi.org/10.1038/sj.onc.1208637

Lomakin, A., Svedlund, J., Strell, C., Gataric, M., Shmatko, A., Rukhovich, G., Park, J. S., Ju, Y. S., Dentro, S., Kleshchevnikov, V., Vaskivskyi, V., Li, T., Bayraktar, O. A., Pinder, S., Richardson, A. L., Santagata, S., Campbell, P. J., Russnes, H., Gerstung, M., … Yates, L. R. (2022). Spatial genomics maps the structure, nature and evolution of cancer clones. *Nature*, *611*(7936), 594–602. https://doi.org/10.1038/s41586-022-05425-2

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. https://doi.org/10.1038/nbt.1977

Ludwig, L. S., Lareau, C. A., Ulirsch, J. C., Christian, E., Muus, C., Li, L. H., Pelka, K., Ge, W., Oren, Y., Brack, A., Law, T., Rodman, C., Chen, J. H., Boland, G. M., Hacohen, N., Rozenblatt-Rosen, O., Aryee, M. J., Buenrostro, J. D., Regev, A., & Sankaran, V. G. (2019). Lineage Tracing in Humans Enabled by Mitochondrial Mutations and Single-Cell Genomics. *Cell*, *176*(6), 1325-1339.e22. https://doi.org/10.1016/j.cell.2019.01.022

Lux, C. T., Yoshimoto, M., McGrath, K., Conway, S. J., Palis, J., & Yoder, M. C. (2008). All primitive and definitive hematopoietic progenitor cells emerging before E10 in the mouse embryo are products of the yolk sac. *Blood*, *111*(7), 3435–3438. https://doi.org/10.1182/blood-2007-08-107086

Martin, E. W., Rodriguez Y Baena, A., Reggiardo, R. E., Worthington, A. K., Mattingly, C. S., Poscablo, D. M., Krietsch, J., McManus, M. T., Carpenter, S., Kim, D. H., & Forsberg, E. C. (2023). Dynamics of Chromatin Accessibility During Hematopoietic Stem Cell Differentiation Into Progressively Lineage-Committed Progeny. *Stem Cells (Dayton, Ohio)*, *41*(5), 520–539. https://doi.org/10.1093/stmcls/sxad022

Masuda, T., Wang, X., Maeda, M., Canver, M. C., Sher, F., Funnell, A. P. W., Fisher, C., Suciu, M., Martyn, G. E., Norton, L. J., Zhu, C., Kurita, R., Nakamura, Y., Xu, J., Higgs, D. R., Crossley, M., Bauer, D. E., Orkin, S. H., Kharchenko, P. V., & Maeda, T. (2016). Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. *Science (New York, N.Y.)*, *351*(6270), 285–289. https://doi.org/10.1126/science.aad3312

McGrath, K. E., Frame, J. M., Fegan, K. H., Bowen, J. R., Conway, S. J., Catherman, S. C., Kingsley, P. D., Koniski, A. D., & Palis, J. (2015). Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. *Cell Reports*, *11*(12), 1892–1904. https://doi.org/10.1016/j.celrep.2015.05.036

Miller, T. E., Lareau, C. A., Verga, J. A., DePasquale, E. A. K., Liu, V., Ssozi, D., Sandor, K., Yin, Y., Ludwig, L. S., El Farran, C. A., Morgan, D. M., Satpathy, A. T., Griffin, G. K., Lane, A. A., Love, J. C., Bernstein, B. E., Sankaran, V. G., & van Galen, P. (2022). Mitochondrial variant enrichment from high-throughput single-cell RNA sequencing resolves clonal populations. *Nature Biotechnology*. https://doi.org/10.1038/s41587-022-01210-8

Mirshekar-Syahkal, B., Haak, E., Kimber, G. M., van Leusden, K., Harvey, K., O'Rourke, J., Laborda, J., Bauer, S. R., de Bruijn, M. F. T. R., Ferguson-Smith, A. C., Dzierzak, E., & Ottersbach, K. (2013). Dlk1 is a negative regulator of emerging hematopoietic stem and progenitor cells. *Haematologica*, *98*(2), 163–171. https://doi.org/10.3324/haematol.2012.070789

Nagasawa, S., Kashima, Y., Suzuki, A., & Suzuki, Y. (2021). Single-cell and spatial analyses of cancer cells: Toward elucidating the molecular mechanisms of clonal evolution and drug resistance acquisition. *Inflammation and Regeneration*, *41*(1), 22. https://doi.org/10.1186/s41232-021-00170-x

Oldak, B., Wildschutz, E., Bondarenko, V., Aguilera-Castrejon, A., Zhao, C., Tarazi, S., Comar, M.-Y., Ashouokhi, S., Lokshtanov, D., Roncato, F., Viukov, S., Ariel, E., Rose, M., Livnat, N., Shani, T., Joubran, C., Cohen, R., Addadi, Y., Kedmi, M., … Hanna, J. H. (2023). *Transgene-Free Ex Utero*

Derivation of A Human Post-Implantation Embryo Model Solely from Genetically Unmodified Naïve PSCs [Preprint]. Developmental Biology. https://doi.org/10.1101/2023.06.14.544922

Orkin, S. H. (1990). Globin gene regulation and switching: Circa 1990. *Cell*, *63*(4), 665–672. https://doi.org/10.1016/0092-8674(90)90133-y

Orkin, S. H., & Zon, L. I. (2008). Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. *Cell*, *132*(4), 631–644. https://doi.org/10.1016/j.cell.2008.01.025

Palis, J. (2014). Primitive and definitive erythropoiesis in mammals. *Frontiers in Physiology*, *5*. https://doi.org/10.3389/fphys.2014.00003

Palla, G., Spitzer, H., Klein, M., Fischer, D., Schaar, A. C., Kuemmerle, L. B., Rybakov, S., Ibarra, I. L., Holmberg, O., Virshup, I., Lotfollahi, M., Richter, S., & Theis, F. J. (2022). Squidpy: A scalable framework for spatial omics analysis. *Nature Methods*, *19*(2), 171–178. https://doi.org/10.1038/s41592-021-01358-2

Patel, S. H., Christodoulou, C., Weinreb, C., Yu, Q., da Rocha, E. L., Pepe-Mooney, B. J., Bowling, S., Li, L., Osorio, F. G., Daley, G. Q., & Camargo, F. D. (2022). Lifelong multilineage contribution by embryonic-born blood progenitors. *Nature*, *606*(7915), 747–753. https://doi.org/10.1038/s41586-022-04804-z

Pedroza, M., Gassaloglu, S. I., Dias, N., Zhong, L., Hou, T.-C. J., Kretzmer, H., Smith, Z. D., & Sozen, B. (2023). Self-patterning of human stem cells into post-implantation lineages. *Nature*. https://doi.org/10.1038/s41586-023-06354-4

Pei, W., Feyerabend, T. B., Rössler, J., Wang, X., Postrach, D., Busch, K., Rode, I., Klapproth, K., Dietlein, N., Quedenau, C., Chen, W., Sauer, S., Wolf, S., Höfer, T., & Rodewald, H.-R. (2017). Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature*, *548*(7668), 456–460. https://doi.org/10.1038/nature23653

Penter, L., Gohil, S. H., Lareau, C., Ludwig, L. S., Parry, E. M., Huang, T., Li, S., Zhang, W., Livitz, D., Leshchiner, I., Parida, L., Getz, G., Rassenti, L. Z., Kipps, T. J., Brown, J. R., Davids, M. S., Neuberg, D. S., Livak, K. J., Sankaran, V. G., & Wu, C. J. (2021). Longitudinal Single-Cell Dynamics of Chromatin Accessibility and Mitochondrial Mutations in Chronic Lymphocytic Leukemia Mirror Disease History. *Cancer Discovery*, *11*(12), 3048–3063. https://doi.org/10.1158/2159-8290.CD-21-0276

Qian, P., He, X. C., Paulson, A., Li, Z., Tao, F., Perry, J. M., Guo, F., Zhao, M., Zhi, L., Venkatraman, A., Haug, J. S., Parmely, T., Li, H., Dobrowsky, R. T., Ding, W.-X., Kono, T., Ferguson-Smith, A. C., & Li, L. (2016). The Dlk1-Gtl2 Locus Preserves LT-HSC Function by Inhibiting the PI3K-mTOR Pathway to Restrict Mitochondrial Metabolism. *Cell Stem Cell*, *18*(2), 214–228. https://doi.org/10.1016/j.stem.2015.11.001

Quinn, J. J., Jones, M. G., Okimoto, R. A., Nanjo, S., Chan, M. M., Yosef, N., Bivona, T. G., & Weissman, J. S. (2021). Single-cell lineages reveal the rates, routes, and drivers of metastasis in cancer xenografts. *Science*, *371*(6532), eabc1944. https://doi.org/10.1126/science.abc1944

Rodriguez-Fraticelli, A. E., Weinreb, C., Wang, S.-W., Migueles, R. P., Jankovic, M., Usart, M., Klein, A. M., Lowell, S., & Camargo, F. D. (2020). Single-cell lineage tracing unveils a role for TCF15 in haematopoiesis. *Nature*, *583*(7817), 585–589. https://doi.org/10.1038/s41586-020-2503-6

Salvador-Martínez, I., Grillo, M., Averof, M., & Telford, M. J. (2019). Is it possible to reconstruct an accurate cell lineage using CRISPR recorders? *eLife*, *8*, e40292. https://doi.org/10.7554/eLife.40292

Sanjuan-Pla, A., Macaulay, I. C., Jensen, C. T., Woll, P. S., Luis, T. C., Mead, A., Moore, S., Carella, C., Matsuoka, S., Bouriez Jones, T., Chowdhury, O., Stenson, L., Lutteropp, M., Green, J. C. A., Facchini, R., Boukarabila, H., Grover, A., Gambardella, A., Thongjuea, S., … Jacobsen, S. E. W. (2013). Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature*, *502*(7470), 232–236. https://doi.org/10.1038/nature12495

Sankaran, V. G., Menne, T. F., Xu, J., Akie, T. E., Lettre, G., Van Handel, B., Mikkola, H. K. A., Hirschhorn, J. N., Cantor, A. B., & Orkin, S. H. (2008). Human Fetal Hemoglobin Expression Is

Regulated by the Developmental Stage-Specific Repressor *BCL11A*. *Science*, *322*(5909), 1839–1842. https://doi.org/10.1126/science.1165409

Siatecka, M., & Bieker, J. J. (2011). The multifunctional role of EKLF/KLF1 during erythropoiesis. *Blood*, *118*(8), 2044–2054. https://doi.org/10.1182/blood-2011-03-331371

Song, K., Li, L., & Zhang, G. (2016). Coverage recommendation for genotyping analysis of highly heterologous species using next-generation sequencing technology. *Scientific Reports*, *6*(1), 35736. https://doi.org/10.1038/srep35736

Sun, J., Ramos, A., Chapman, B., Johnnidis, J. B., Le, L., Ho, Y.-J., Klein, A., Hofmann, O., & Camargo, F. D. (2014). Clonal dynamics of native haematopoiesis. *Nature*, *514*(7522), 322–327. https://doi.org/10.1038/nature13824

Tallack, M. R., Magor, G. W., Dartigues, B., Sun, L., Huang, S., Fittock, J. M., Fry, S. V., Glazov, E. A., Bailey, T. L., & Perkins, A. C. (2012). Novel roles for KLF1 in erythropoiesis revealed by mRNA-seq. *Genome Research*, *22*(12), 2385–2398. https://doi.org/10.1101/gr.135707.111

Tian, L., Tomei, S., Schreuder, J., Weber, T. S., Amann-Zalcenstein, D., Lin, D. S., Tran, J., Audiger, C., Chu, M., Jarratt, A., Willson, T., Hilton, A., Pang, E. S., Patton, T., Kelly, M., Su, S., Gouil, Q., Diakumis, P., Bahlo, M., … Naik, S. H. (2021). Clonal multi-omics reveals Bcor as a negative regulator of emergency dendritic cell development. *Immunity*, *54*(6), 1338-1351.e9. https://doi.org/10.1016/j.immuni.2021.03.012

Tsai, F. Y., Keller, G., Kuo, F. C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F. W., & Orkin, S. H. (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*, *371*(6494), 221–226. https://doi.org/10.1038/371221a0

VanHorn, S., & Morris, S. A. (2021). Next-Generation Lineage Tracing and Fate Mapping to Interrogate Development. *Developmental Cell*, *56*(1), 7–21. https://doi.org/10.1016/j.devcel.2020.10.021

Wagner, D. E., & Klein, A. M. (2020). Lineage tracing meets single-cell omics: Opportunities and challenges. *Nature Reviews Genetics*, *21*(7), 410–427. https://doi.org/10.1038/s41576-020-0223-2

Walker, B. L., Cang, Z., Ren, H., Bourgain-Chang, E., & Nie, Q. (2022). Deciphering tissue structure and function using spatial transcriptomics. *Communications Biology*, *5*(1), 220. https://doi.org/10.1038/s42003-022-03175-5

Walmsley, M. E., Guille, M. J., Bertwistle, D., Smith, J. C., Pizzey, J. A., & Patient, R. K. (1994). Negative control of Xenopus GATA-2 by activin and noggin with eventual expression in precursors of the ventral blood islands. *Development (Cambridge, England)*, *120*(9), 2519–2529. https://doi.org/10.1242/dev.120.9.2519

Wan, X., Xiao, J., Tam, S. S. T., Cai, M., Sugimura, R., Wang, Y., Wan, X., Lin, Z., Wu, A. R., & Yang, C. (2023). *SpatialScope: A unified approach for integrating spatial and single-cell transcriptomics data using deep generative models* [Preprint]. Bioinformatics. https://doi.org/10.1101/2023.03.14.532529

Wang, R., Li, J., Zhou, X., Mao, Y., Wang, W., Gao, S., Wang, W., Gao, Y., Chen, K., Yu, S., Wu, X., Wen, L., Ge, H., Fu, W., & Tang, F. (2022). Single-cell genomic and transcriptomic landscapes of primary and metastatic colorectal cancer tumors. *Genome Medicine*, *14*(1), 93. https://doi.org/10.1186/s13073-022-01093-z

Weatherbee, B. A. T., Gantner, C. W., Iwamoto-Stohl, L. K., Daza, R. M., Hamazaki, N., Shendure, J., & Zernicka-Goetz, M. (2023). Pluripotent stem cell-derived model of the post-implantation human embryo. *Nature*. https://doi.org/10.1038/s41586-023-06368-y

Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F. D., & Klein, A. M. (2020). Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science (New York, N.Y.)*, *367*(6479), eaaw3381. https://doi.org/10.1126/science.aaw3381

Weintraub, H., Campbell, G. le M., & Holtzer, H. (1971). Primitive erythropoiesis in early chick embryogenesis. I. Cell cycle kinetics and the control of cell division. *The Journal of Cell Biology*, *50*(3), 652–668. https://doi.org/10.1083/jcb.50.3.652

Wolf, F. A., Angerer, P., & Theis, F. J. (2018). SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biology*, *19*(1), 15. https://doi.org/10.1186/s13059-017-1382-0

Woodworth, M. B., Girskis, K. M., & Walsh, C. A. (2017). Building a lineage from single cells: Genetic techniques for cell lineage tracking. *Nature Reviews Genetics*, *18*(4), 230–244. https://doi.org/10.1038/nrg.2016.159

Xu, J., Bauer, D. E., Kerenyi, M. A., Vo, T. D., Hou, S., Hsu, Y.-J., Yao, H., Trowbridge, J. J., Mandel, G., & Orkin, S. H. (2013). Corepressor-dependent silencing of fetal hemoglobin expression by BCL11A. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(16), 6518–6523. https://doi.org/10.1073/pnas.1303976110

Xu, J., Nuno, K., Litzenburger, U. M., Qi, Y., Corces, M. R., Majeti, R., & Chang, H. Y. (2019). Single-cell lineage tracing by endogenous mutations enriched in transposase accessible mitochondrial DNA. *eLife*, *8*, e45105. https://doi.org/10.7554/eLife.45105

Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K. L., Ema, H., & Nakauchi, H. (2013). Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell*, *154*(5), 1112–1126. https://doi.org/10.1016/j.cell.2013.08.007

Yang, D., Jones, M. G., Naranjo, S., Rideout, W. M., Min, K. H. (Joseph), Ho, R., Wu, W., Replogle, J. M., Page, J. L., Quinn, J. J., Horns, F., Qiu, X., Chen, M. Z., Freed-Pastor, W. A., McGinnis, C. S., Patterson, D. M., Gartner, Z. J., Chow, E. D., Bivona, T. G., … Weissman, J. S. (2022). Lineage tracing reveals the phylodynamics, plasticity, and paths of tumor evolution. *Cell*, *185*(11), 1905-1923.e25. https://doi.org/10.1016/j.cell.2022.04.015

Yokomizo, T., Ideue, T., Morino-Koga, S., Tham, C. Y., Sato, T., Takeda, N., Kubota, Y., Kurokawa, M., Komatsu, N., Ogawa, M., Araki, K., Osato, M., & Suda, T. (2022). Independent origins of fetal liver haematopoietic stem and progenitor cells. *Nature*, *609*(7928), 779–784. https://doi.org/10.1038/s41586-022-05203-0

Zhao, T., Chiang, Z. D., Morriss, J. W., LaFave, L. M., Murray, E. M., Del Priore, I., Meli, K., Lareau, C. A., Nadaf, N. M., Li, J., Earl, A. S., Macosko, E. Z., Jacks, T., Buenrostro, J. D., & Chen, F. (2022). Spatial genomics enables multi-modal study of clonal heterogeneity in tissues. *Nature*, *601*(7891), 85–91. https://doi.org/10.1038/s41586-021-04217-4

Zverinova, S., & Guryev, V. (2022). Variant calling: Considerations, practices, and developments. *Human Mutation*, *43*(8), 976–985. https://doi.org/10.1002/humu.24311

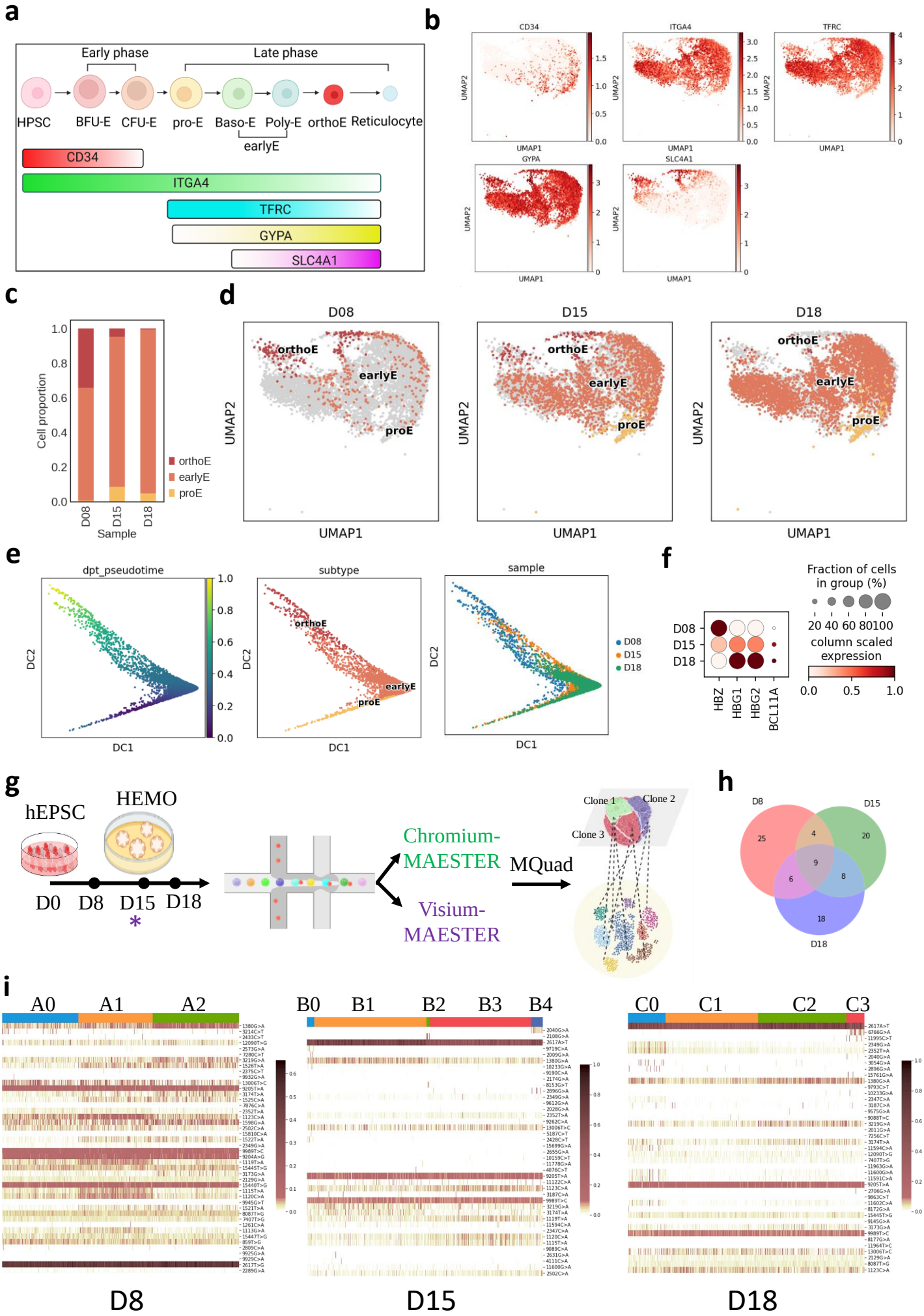
a
Early phase
Late phase
HPSC
BFU-E
CFU-E
pro-E
Baso-E
Poly-E
orthoE
Reticulocyte
earlyE
CD34
ITGA4
TFRC
GYPA
SLC4A1
b
CD34
ITGA4
TFRC
GYPA
SLC4A1
UMAP1
UMAP2
c
Cell proportion
Sample
D08
D15
D18
orthoE
earlyE
proE
d
D08
D15
D18
orthoE
earlyE
proE
UMAP1
UMAP2
e
dpt_pseudotime
subtype
sample
DC1
DC2
orthoE
earlyE
proE
D08
D15
D18
f
Fraction of cells in group (%)
20 40 60 80100
column scaled expression
HBZ
HBG1
HBG2
BCL11A
g
hEPSC
HEMO
D0
D8
D15
D18
*
Chromium-MAESTER
Visium-MAESTER
MQuad
Clone 1
Clone 2
Clone 3
h
D8
D15
D18
25
4
20
9
6
8
18
i
A0
A1
A2
B0
B1
B2
B3
B4
C0
C1
C2
C3
D8
D15
D18

Fig. 1 The dichotomous waves of erythropoiesis in human embryonic organoids.

a, Human erythropoiesis overview.

b, UMAP plots shows the expression pattern of erythropoiesis markers CD34, ITGA4, TFRC, GYPA and SLC4A1.

c, Erythroid subtypes across three time points of human embryonic organoids (HEMOs) indicate the late-stage population orthoE enrichment.

d, UMAP visualizes erythroid subtypes at D8, D15 and D18 of HEMOs. D8 enriched the late-stage population orthoE.

e, Pseudotime analysis put D8 at most “mature” stage among three time points.

f, Dotplot indicates the expression of embryonic globin HBZ, fetal globins HBG1 and HBG2, as well as fetal globin modulator BCL11A, suggesting globin switch between D8 and D15.

g, Overall study design of mitochondrial variant analysis. D8, D15, and D18 HEMOs were harvested for mitochondrial variant enrichment to generate Chromium-MAESTER libraries. D15 HEMOs were also harvested for 10X Visium spatial transcriptomics based mitochondrial variant enrichment to generate Visium-MAESTER libraries.

h, Venn diagram of mtDNA variants identified in D8, D15 and D18 HEMO samples via Chromium-MAESTER.

i, Allele frequency heatmap shows the 43, 41, and 41 informative mtDNA SNVs detected by MQuad in each clone in D8, D15 and D18 HEMOs. Each row is a variant, each column is a barcode. Heatmap color indicates the value of the allele frequency.

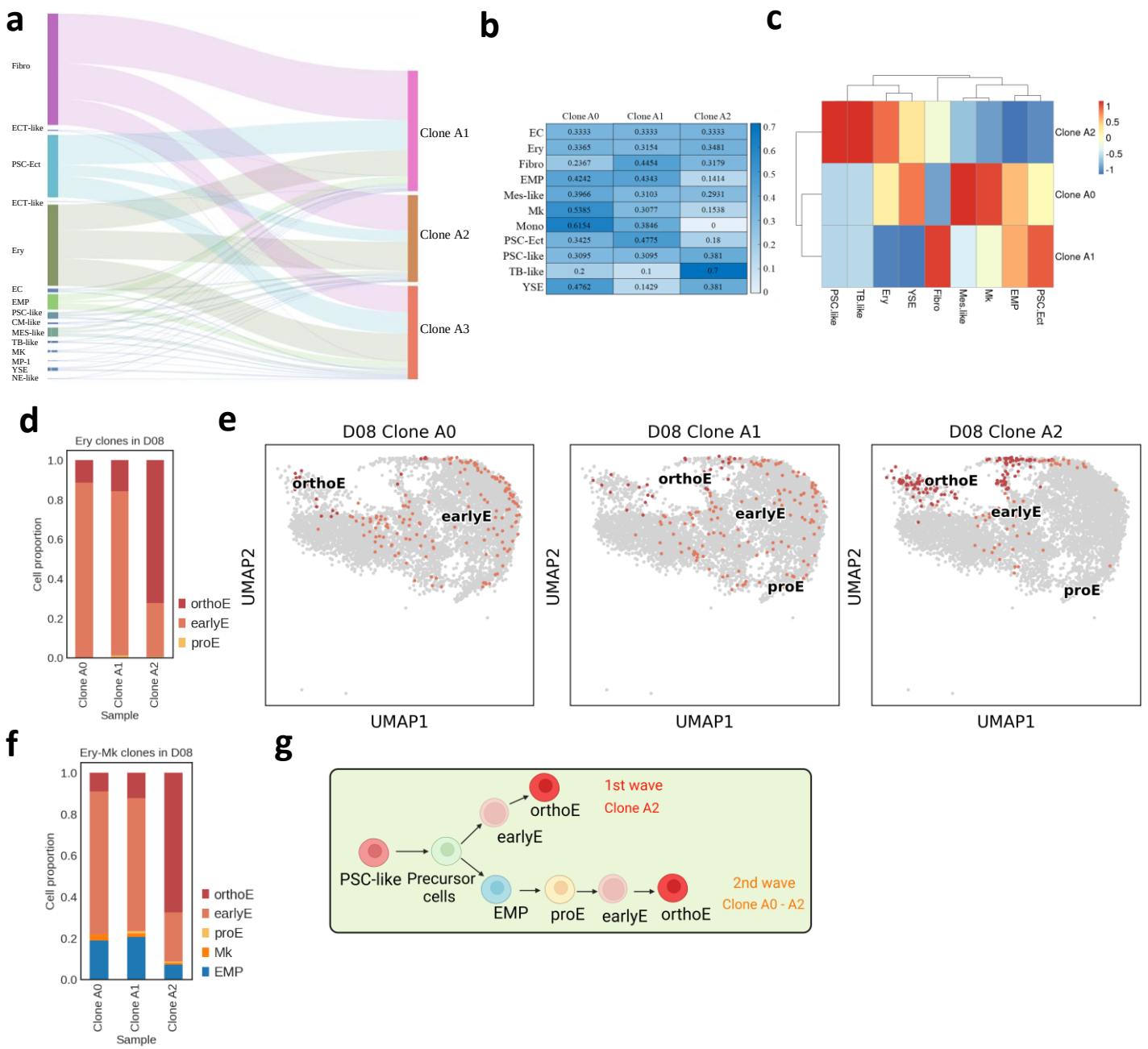

Fig. 2 Single-cell mitochondrial SNV enrichment captured two erythropoiesis waves in D8 HEMOs.

a, Alluvial plot demonstrates that Fibro, Ery and PSC-Ect are the predominant lineages in three clones traced by mtDNA variant enrichment.

b, Heatmap shows clone inflow trends of different lineages. High proportion of MP, Mono and Mk lineages tend to inflow towards Clone A0. More MP, Mono and Mk lineages tend to inflow towards Clone A1. More Ect-like and Fibro lineages tend to inflow towards Clone A2.

c, Heatmap shows hierarchy structure of three clones. Clone A0 has a closer relationship with Clone A1, while Clone A2 is much more remote.

d, Bar chart of erythroid subtypes in individual clones indicates that early-stage proE missing, whilst late-stage OrthE enriched.

e, UMAP of erythroid subtypes in individual clones enriched orthoE.

f, Bar chart shows the emerging of EMP in all three clones.

g, An overview of multiple erythropoiesis waves captured by mitochondrial SNV enrichment enabled clonal tracking at single cell resolution.

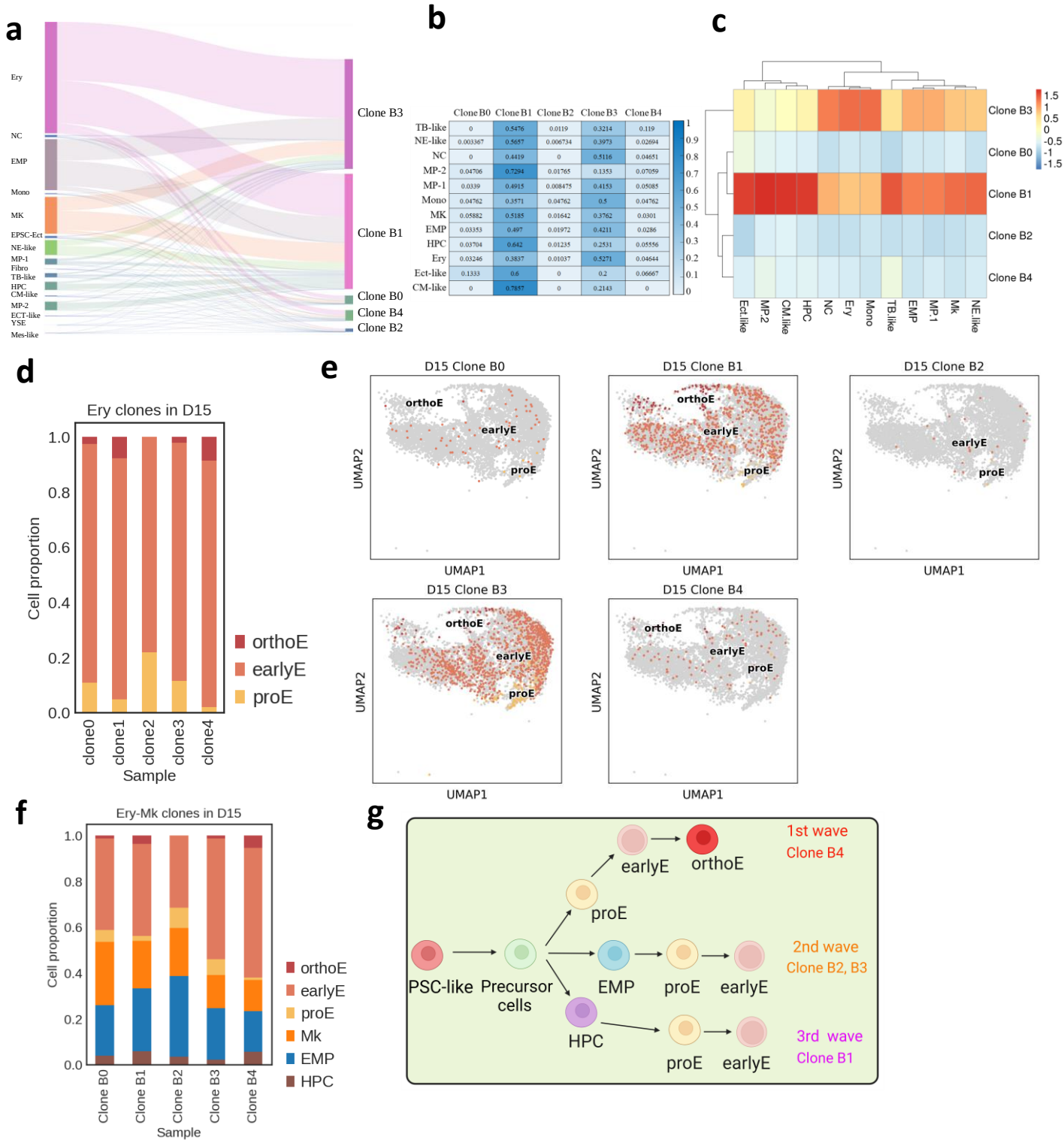

Fig. 3 Single-cell mitochondrial SNV enrichment captured multiple erythropoiesis waves in D15 HEMOs.

a, Alluvial plot demonstrates that Ery, EMP and Mk are the predominant lineages in five clones traced by mtDNA variant enrichment.

b, Heatmap shows clone inflow trends of different lineages. Most lineages prefer inflow towards Clone B1 and B3.

c, Heatmap shows the hierarchy structure of 5 clones. Clone B0 and B1 are closer in relationship, and Clone B3 and B4 are closer.

d, Bar chart of erythroid subtypes in individual clones indicates that the presence of early-stage proE.

e, UMAP visualizes erythroid subtypes in individual clone.

f, Bar chart shows the emerging of EMP and HPC in all five clones at hematopoiesis emerging stage.

g, An overview of multiple erythropoiesis waves captured by mitochondrial SNV enrichment enabled clonal tracking at single cell resolution.

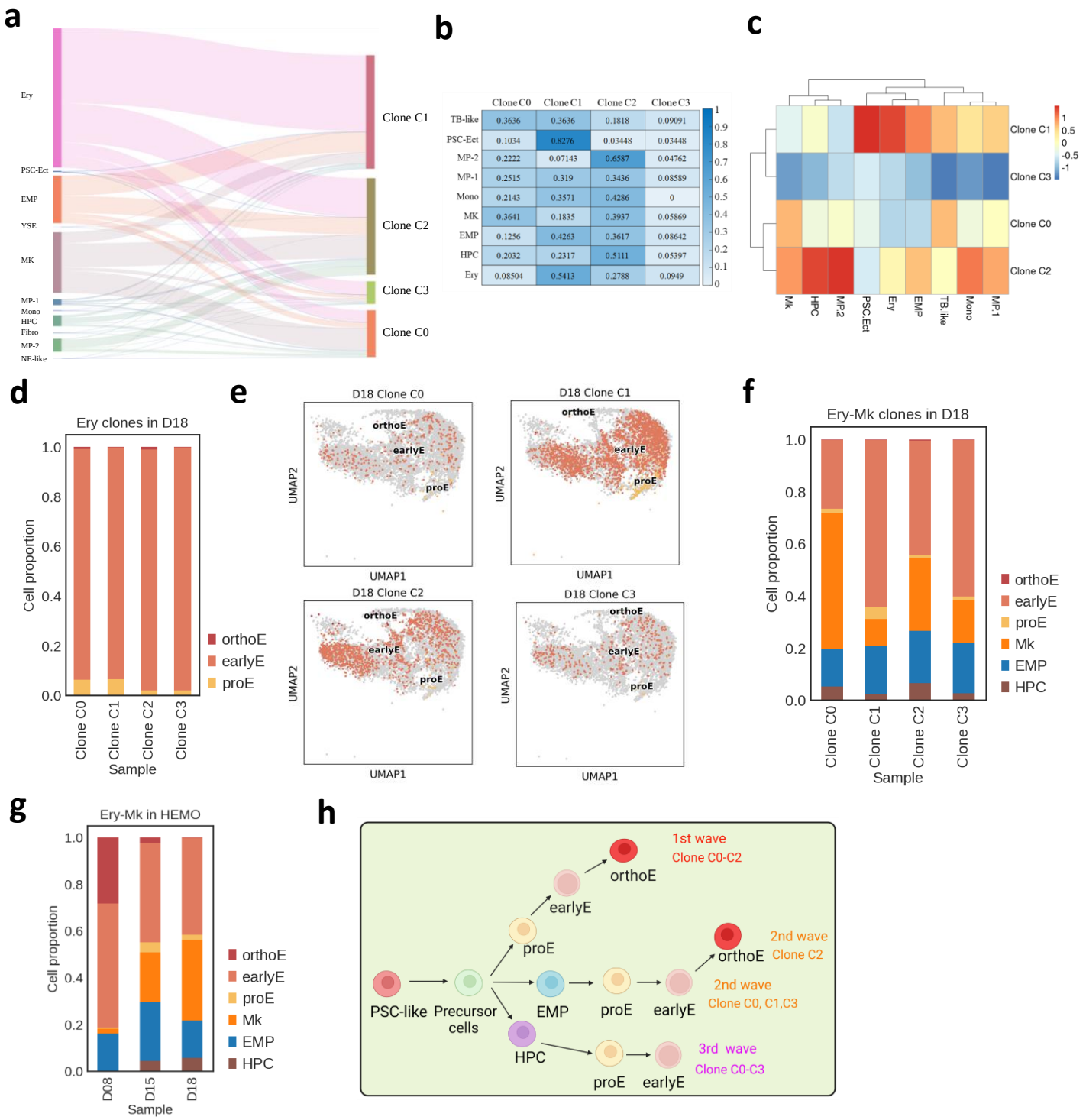

Fig. 4 Single-cell mitochondrial SNV enrichment captured multiple erythropoiesis waves in D18 HEMOs.

a, Alluvial plot demonstrates that Ery, EMP and Mk are the predominant lineages in four clones traced by mtDNA variant enrichment.

b, Heatmap shows clone inflow trends of different lineages. Most lineages prefer inflow towards Clone C1 and C2.

c, Heatmap shows the hierarchy structure of 4 clones. Clone C0 and C2 are closer in relationship, and Clone C1 and C3 are closer.

d, Bar chart of erythroid subtypes in individual clone.

e, UMAP visualization of erythroid subtypes in individual clone.

f, Bar chart shows the increase of Mk in all four clones.

g, Bar chart shows the dynamic composition changes of HPC, EMP, Ery and Mk lineages across three developmental stages.

h, An overview of multiple erythropoiesis waves captured by mitochondrial SNV enrichment enabled clonal tracking at single cell resolution.

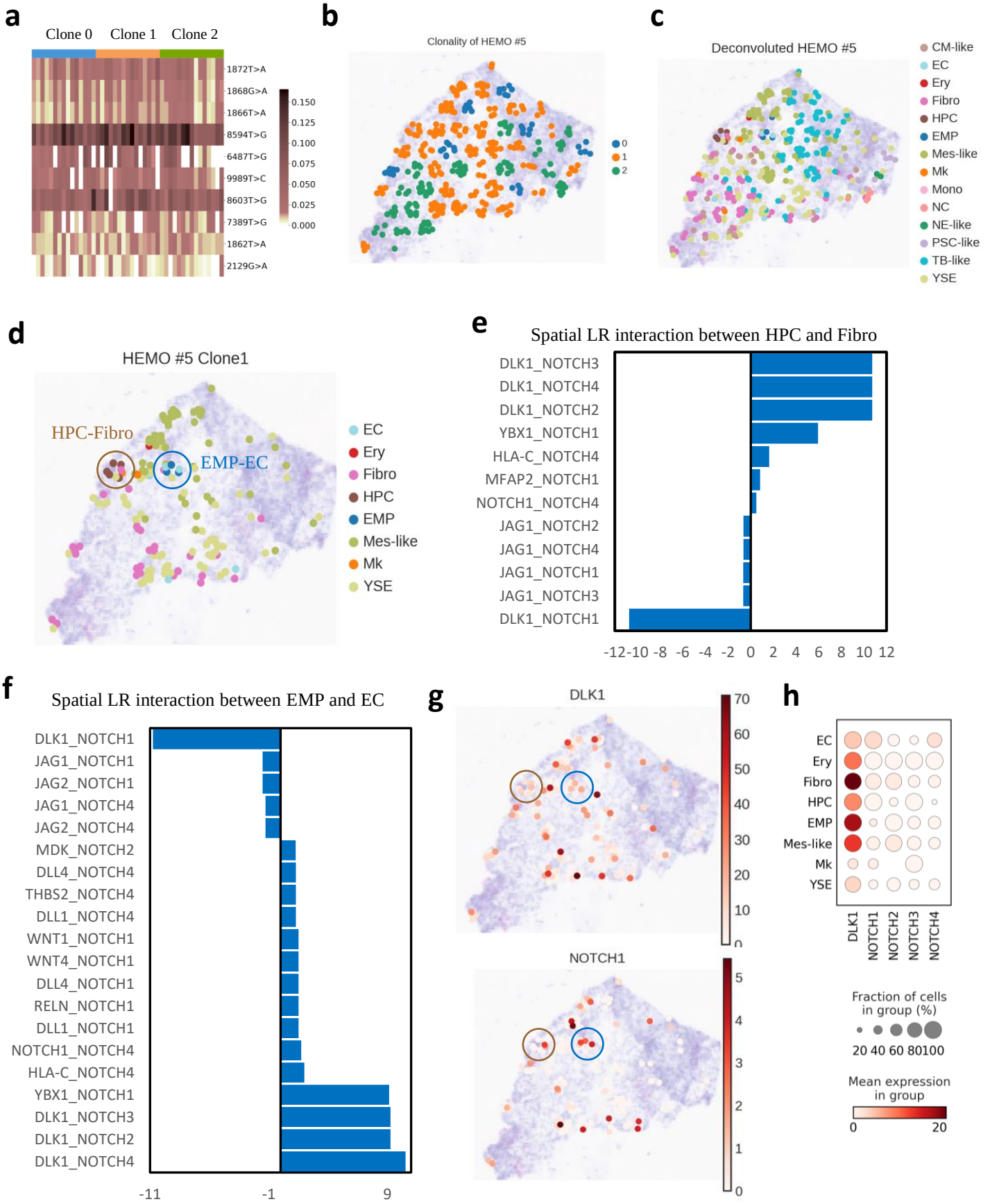
a
Clone 0
Clone 1
Clone 2
1872T>A
1868G>A
1866T>A
8594T>G
6487T>G
9989T>C
8603T>G
7389T>G
1862T>A
2129G>A
0.150
0.125
0.100
0.075
0.050
0.025
0.000
b
Clonality of HEMO #5
0
1
2
c
Deconvoluted HEMO #5
CM-like
EC
Ery
Fibro
HPC
EMP
Mes-like
Mk
Mono
NC
NE-like
PSC-like
TB-like
YSE
d
HEMO #5 Clone1
HPC-Fibro
EMP-EC
EC
Ery
Fibro
HPC
EMP
Mes-like
Mk
YSE
e
Spatial LR interaction between HPC and Fibro
DLK1_NOTCH3
DLK1_NOTCH4
DLK1_NOTCH2
YBX1_NOTCH1
HLA-C_NOTCH4
MFAP2_NOTCH1
NOTCH1_NOTCH4
JAG1_NOTCH2
JAG1_NOTCH4
JAG1_NOTCH1
JAG1_NOTCH3
DLK1_NOTCH1
-12 -10 -8 -6 -4 -2 0 2 4 6 8 10 12
f
Spatial LR interaction between EMP and EC
DLK1_NOTCH1
JAG1_NOTCH1
JAG2_NOTCH1
JAG1_NOTCH4
JAG2_NOTCH4
MDK_NOTCH2
DLL4_NOTCH4
THBS2_NOTCH4
DLL1_NOTCH4
WNT1_NOTCH1
WNT4_NOTCH1
DLL4_NOTCH1
RELN_NOTCH1
DLL1_NOTCH1
NOTCH1_NOTCH4
HLA-C_NOTCH4
YBX1_NOTCH1
DLK1_NOTCH3
DLK1_NOTCH2
DLK1_NOTCH4
-11
-1
9
g
DLK1
70
60
50
40
30
20
10
0
NOTCH1
5
4
3
2
1
0
h
EC
Ery
Fibro
HPC
EMP
Mes-like
Mk
YSE
DLK1
NOTCH1
NOTCH2
NOTCH3
NOTCH4
Fraction of cells in group (%)
20 40 60 80100
Mean expression in group
0
20

Fig 5. Mitochondrial SNVs based spatial clonal tracking in D15 HEMOs.

a, Allele frequency heatmap shows the 10 informative mtDNA SNVs detected by MQuad in each clone in HEMO #5. Each row is a variant, each column is a barcode. Heatmap color indicates the value of the allele frequency.

b, Spatial architecture of 3 clones identified by mtDNA SNVs from MQuad in Visium-MAESTER library within HEMO #5.

c, Spatial transcriptomics map of HEMO #5 section, colored by clusters.

d, Spatial plot of EC,Ery, Fibro,HPC,EMP,Mes-like,Mk and YSE population within Clone 1 of HEMO #5.

e,f, Spatial cell-cell interaction analysis conducted by CellPhoneDB within Clone 1 between HPC and Fibro (e) and EMP and EC (f) . Pairs with a mean value above 0 indicate the activation of the ligand-receptor pairs. Pairs with a mean value below 0 indicate the inhibition of the pairs. Among these, DLK1_NOTCH1 exhibit the highest enriched score between HPC-Fibro and EMP-EC.

g, Gene expression pattern of DLK1 and NOTCH1 within Clone 1 on HEMO #5 in a spatial slice. DLK1 highly expressed in EMP within Clone 1. NOTCH1 highly expressed in EC and EMP.

h, Dotplot shows the expression of DLK1 and NOTCH1.

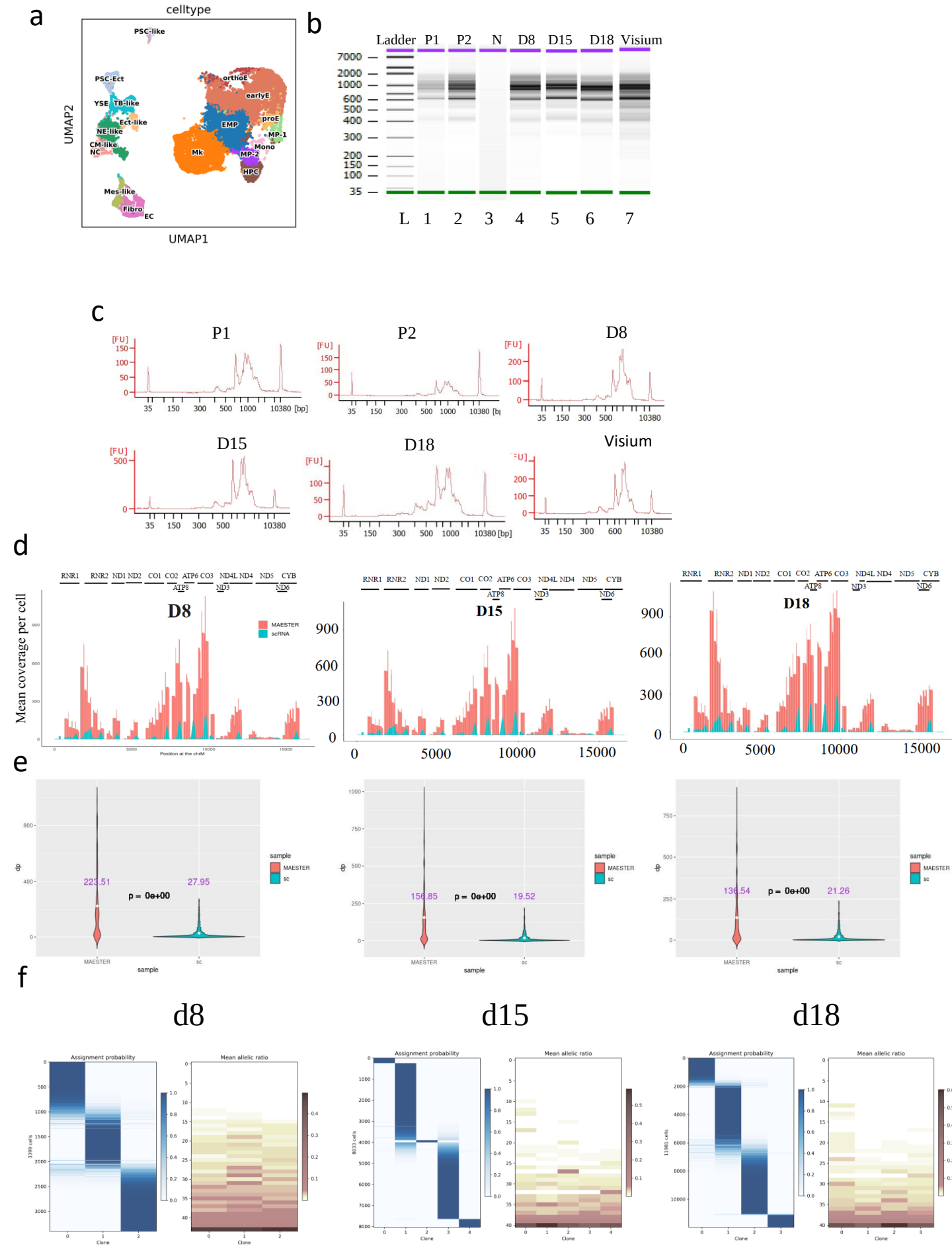
a
celltype
PSC-like
PSC-Ect
YSE
TB-like
Ect-like
NE-like
CM-like
NC
orthoE
earlyE
EMP
proE
MP-1
Mono
Mk
MP-2
HPC
Mes-like
Fibro
EC
UMAP2
UMAP1
b
Ladder P1 P2 N D8 D15 D18 Visium
7000
2000
1000
600
500
400
300
200
150
100
35
L 1 2 3 4 5 6 7
c
P1
P2
D8
D15
D18
Visium
[FU]
35 150 300 500 1000 10380 [bp]
d
RNR1 RNR2 ND1 ND2 CO1 CO2 ATP6 CO3 ND4L ND4 ND5 CYB
ATP8 ND3 ND6
D8
D15
D18
Mean coverage per cell
MAESTER
scRNA
Position at the chrM
e
223.51
27.95
p = 0e+00
156.85
19.52
136.54
21.26
sample
MAESTER
sc
f
d8
d15
d18
Assignment probability
Mean allelic ratio
Clone

Fig S1

Fig. S1 Enrichment of mitochondrial transcripts and confident mtDNA variant calling enabled clonal tracking across different stages during HEMO development.

a, UMAP of major cell type in HEMO captured by scRNA-seq. earlyE early erythroblast, proE proerythroblast, orthoE orthochromatic erythroblasts, CM-like cardiac mesoderm-like cells, Ect-like Ectoderm-like cells, EC endothelium, Fibro fibroblasts, HPC hematopoietic progenitor cells, MP-1 myeloid progenitor 1, EMP erythro-myeloid progenitors, MP-2 myeloid progenitor 2, Mes-like mesoderm-like cells, Mk megakaryocytes, Mono: monocytes, NC neural crest, NE-like neural ectoderm-like cells, PSC-Ect pluripotent stem cells with ectoderm specification, PSC-like pluripotent stem cells-like, TB-like trophoblast-like cells, YSE yolk sac endoderm.

b,c High sensitivity DNA assay (Bioanalyzer) on Chromium-MAESTER and Visium-MAESTER libraries. L, DNA ladder; P1 and P2, positive control contributed by MEASTER developer; N, negative control; D8, D15 and D18, Chromium-MAESTER of D8, D15 and D18 HEMOs; Visium, Visium-MAESTER library of D15 HEMOs.

d,e Bar plot and Violin plot show increased coverage of the mitochondrial genome with MAESTER amplification from 10x Genome libraries.

f, (Left) Heatmap shows barcode assignment with informative mtDNA variants detected with MQuad. Each row is a cell, each column is a barcode, heatmap color indicates assignment probability of cells in D8, D15 and D18 HEMOs.
(Right) Allele frequency heatmap shows the 43, 41, and 41 informative mtDNA SNVs detected by MQuad ranked from lowest to highest score of difference in Bayesian Information Criterion (ΔBIC) in each clone in D8, D15 and D18 HEMOs. Each row is a variant, each column is a barcode. Heatmap color indicates the value of the ΔBIC which is an indicator for clonal informativeness of each SNP with higher ΔBIC being more informative.

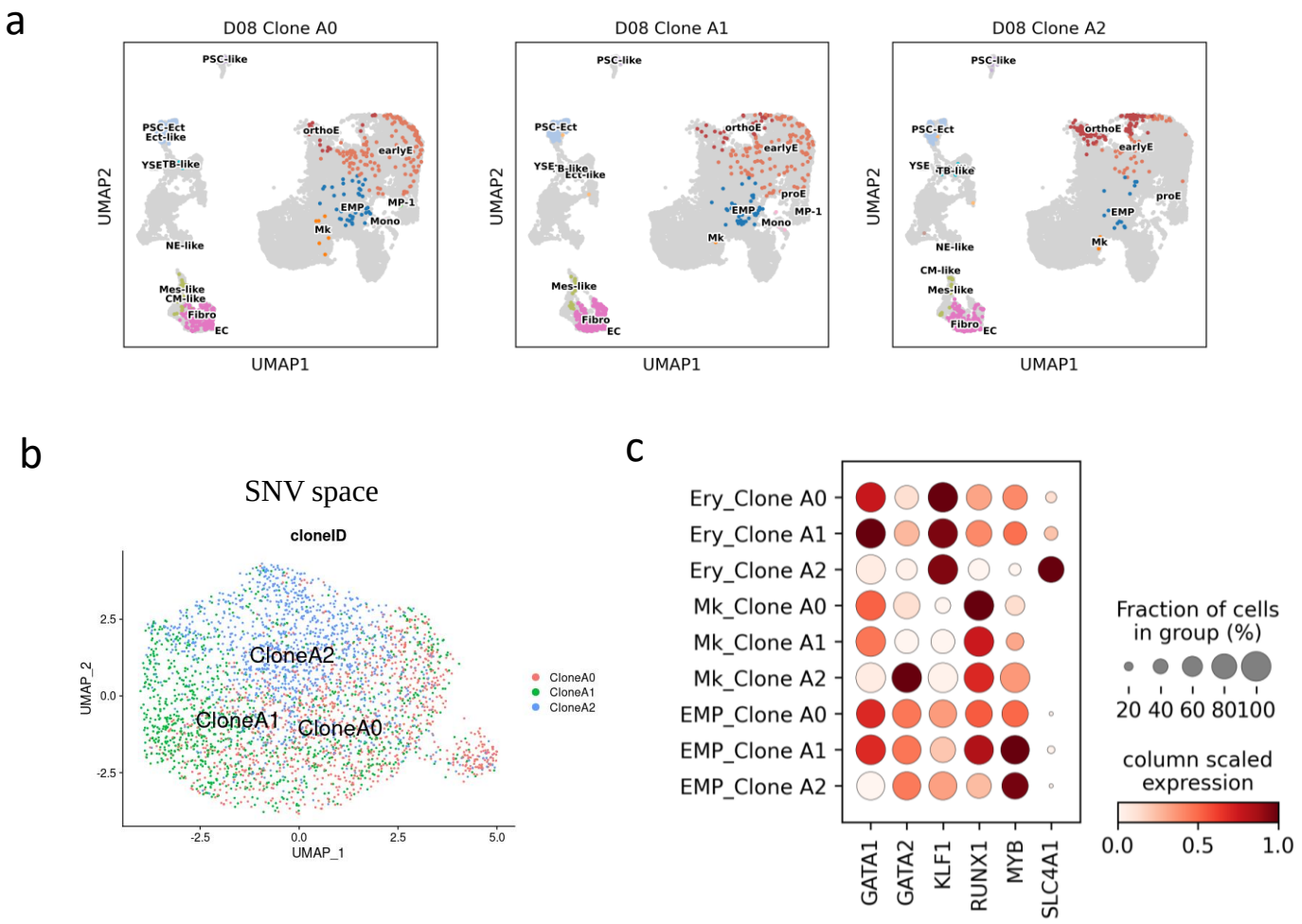

Fig. S2 Landscape of HEMO cells and mtDNA SNVs in each clone at D8 stage.

a, UMAP shows various cell types in each clone.

b, UMAP plot of D8 HEMO cells in mtDNA SNV space, colored by clone ID.

c, Dotplot of a core set of transcription factors regulating erythropoiesis.

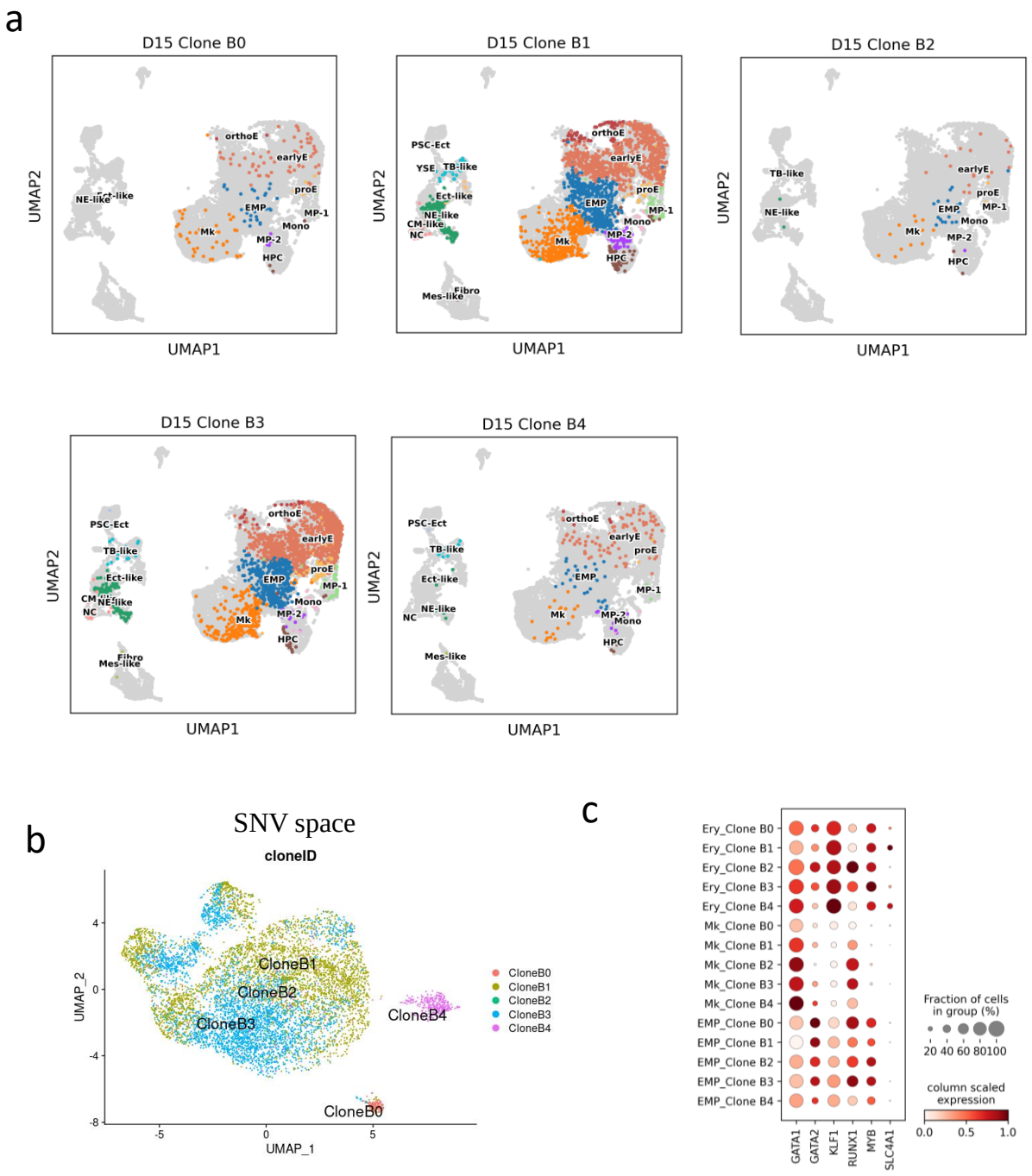

Fig. S3 Landscape of HEMO cells and mtDNA SNVs in each clone at D15 stage.

a, UMAP shows various cell types in each clone.

b, UMAP plot of D15 HEMO cells in mtDNA SNV space, colored by clone ID.

c, Dotplot of a core set of transcription factors regulating erythropoiesis.

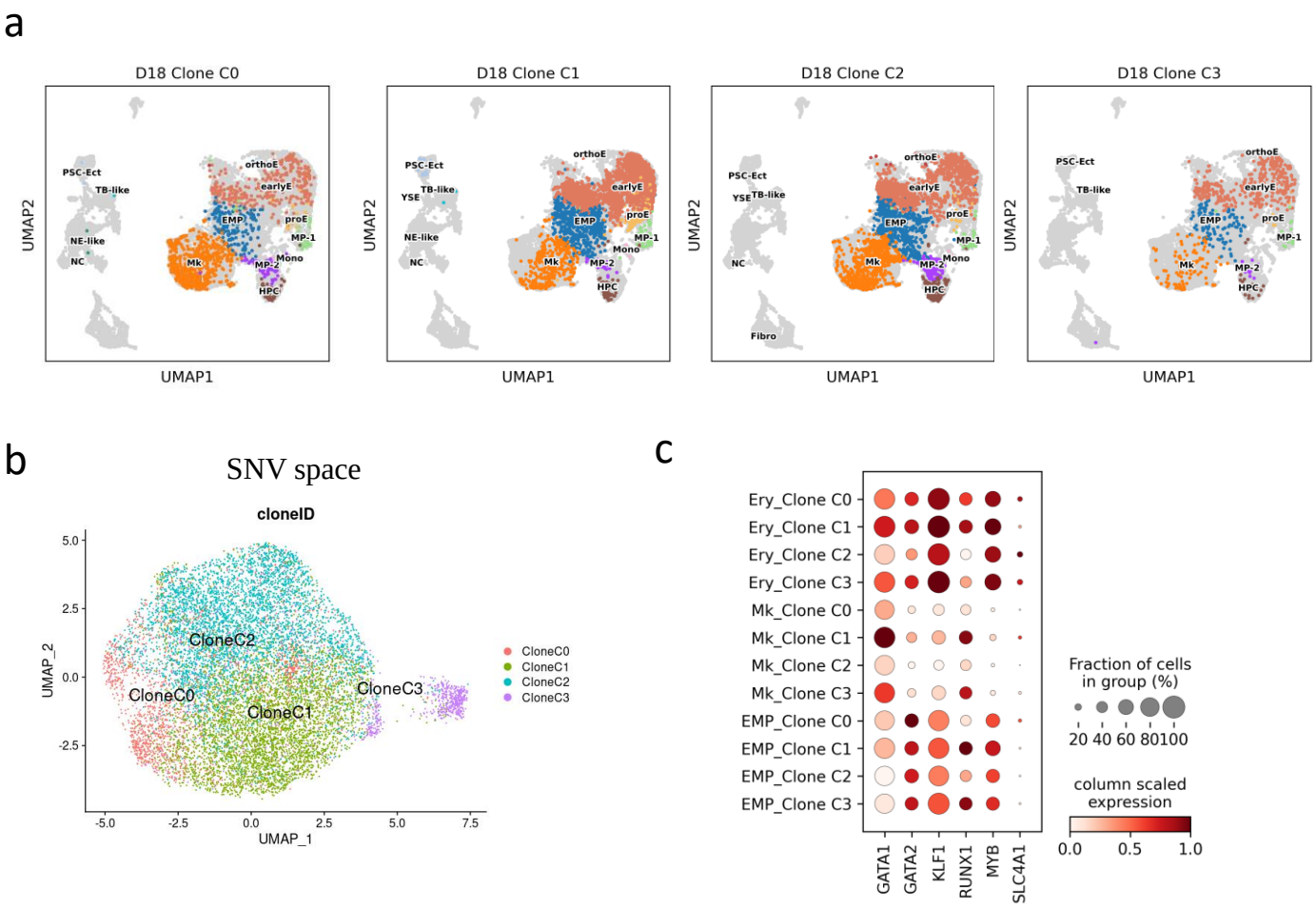

Fig. S4 Landscape of HEMO cells and mtDNA SNVs in each clone at D18 stage.

a, UMAP shows various cell types in each clone.

b, UMAP plot of D15 HEMO cells in mtDNA SNV space, colored by clone ID.

c, Dotplot of a core set of transcription factors regulating erythropoiesis.

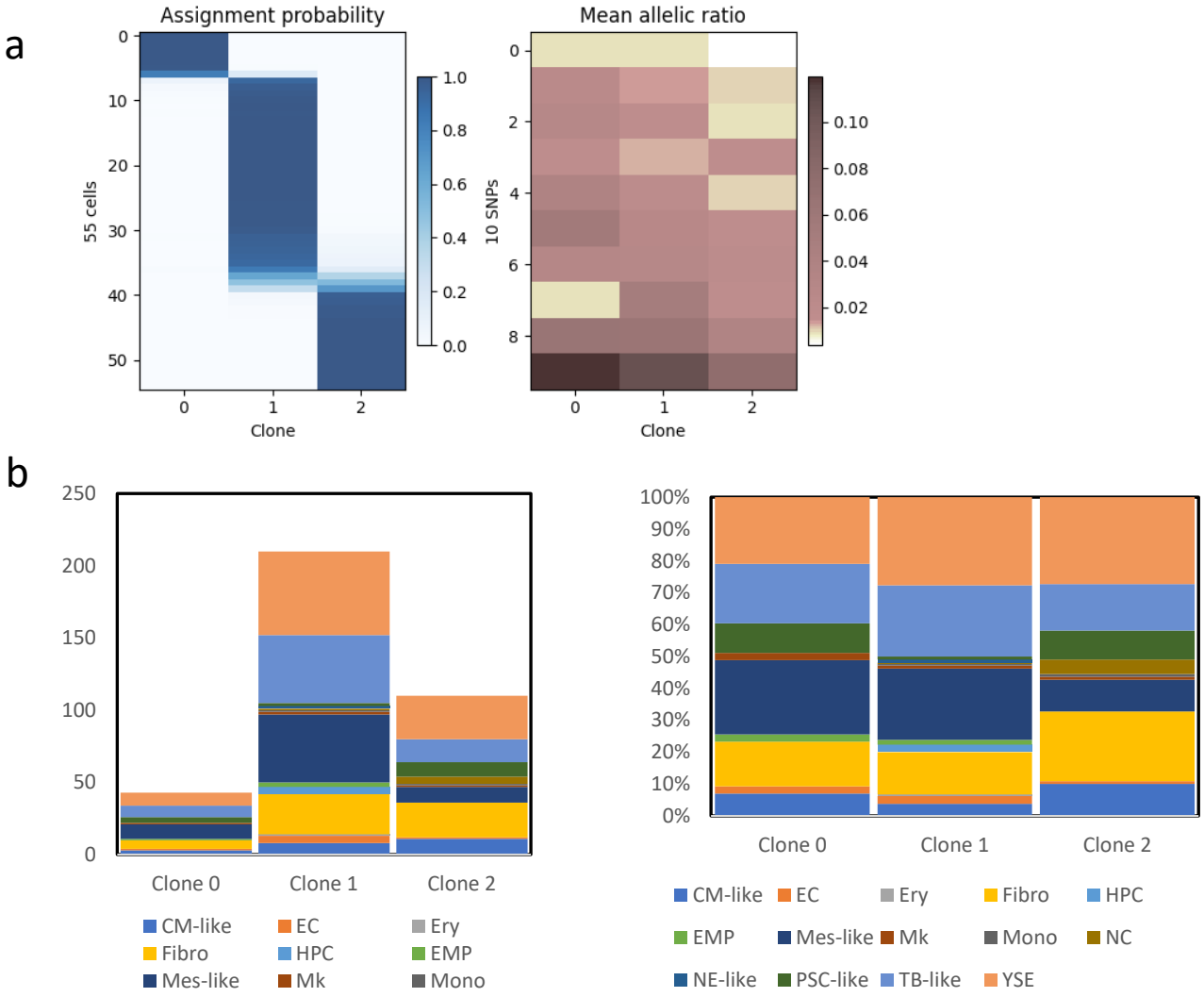

Fig. S5 Enriched mitochondrial variants spatially resolved clonal architecture in hematopoietic HEMO.

a, (Left) Heatmap shows barcode assignment with informative mtDNA variants detected with MQuad. Each row is a cell, each column is a barcode, heatmap color indicates assignment probability of cells in D8, D15 and D18 HEMOs.

(Right) Allele frequency heatmap shows the 10 informative mtDNA SNVs detected by MQuad ranked from lowest to highest score of difference in Bayesian Information Criterion (ΔBIC) in each clone in D8, D15 and D18 HEMOs. Each row is a variant, each column is a barcode. Heatmap color indicates the value of the ΔBIC which is an indicator for clonal informativeness of each SNP with higher ΔBIC being more informative.

b, Bar plots show the amount (left) and proportion (right) of different cell types in individual clone tracked by Visium-MAESTER.

Table 1 Primers used for mitochondrial variants enrichment.

PCR Mix	Primer name	Complete sequence (5'to 3')
1	PvG1218_MT-ND1_4009	CACCCGAGAATTCCAAACACCCTCACCACTACAATCT
1	PvG1223_MT-ND2_5363	CACCCGAGAATTCCACTCCACCTCAATCACACTACTCC
1	PvG1230_MT-CO1_7184	CACCCGAGAATTCCAACAACACTTTCTCGGCCTATCC
1	PvG1234_MT-ATP8_8367	CACCCGAGAATTCCATGCCCCAACTAAATACTACCG
1	PvG1241_MT-CO3_9756	CACCCGAGAATTCCATCTCCCTTCACCATTTCCGAC
1	PvG1243_MT-ND3_10264	CACCCGAGAATTCCATTGCCCTCCTTTTACCCCTAC
1	PvG1244_MT-ND4L_10496	CACCCGAGAATTCCAACTAGCATTTACCATCTCACTTCT
1	PvG1251_MT-ND4_11900	CACCCGAGAATTCCAGTGCTAGTAACCACGTTCTCCT
1	PvG1259_MT-ND5_13926	CACCCGAGAATTCCATAGCATCACACACCGCACAA
1	PvG1260_MT-ND6_14263	CACCCGAGAATTCCAGGATCCTATTGGTGCGGGG
1	PvG1268_MT-CYB_15643	CACCCGAGAATTCCACATCCTAGCAATAATCCCCATCCT
2	PvG1217_MT-ND1_3777	CACCCGAGAATTCCATGGCTCCTTTAACCTCTCCAC
2	PvG1222_MT-ND2_5145	CACCCGAGAATTCCAACGACCCTACTACTATCTCGCA
2	PvG1229_MT-CO1_6957	CACCCGAGAATTCCAGGCCTGACTGGCATTGTATT
2	PvG1232_MT-CO2_7852	CACCCGAGAATTCCAGGTCAACGATCCCTCCCTTAC
2	PvG1236_MT-ATP6_8766	CACCCGAGAATTCCACACAACTAACCTCCTCGGACT
2	PvG1240_MT-CO3_9535	CACCCGAGAATTCCACCCAATTAGGAGGGCACTGG
2	PvG1242_MT-ND3_10127	CACCCGAGAATTCCAACTACCACAACTCAACGGCTAC
2	PvG1250_MT-ND4_11684	CACCCGAGAATTCCATTCACCGGCGCAGTCATT
2	PvG1258_MT-ND5_13758	CACCCGAGAATTCCACGCATCCCCCTTCCAAACA
2	PvG1261_MT-ND6_14492	CACCCGAGAATTCCAGGGGAATGATGGTTGTCTTTGG
2	PvG1267_MT-CYB_15432	CACCCGAGAATTCCACCCTCGGCTTACTTCTCTTCC
3	PvG1216_MT-ND1_3537	CACCCGAGAATTCCAAGCTCTCACCATCGCTCTTC
3	PvG1221_MT-ND2_4923	CACCCGAGAATTCCAAGCCTTCTCCTCACTCTCTCAA
3	PvG1228_MT-CO1_6742	CACCCGAGAATTCCATTGGCTTCCTAGGGTTTATCGTG
3	PvG1231_MT-CO2_7609	CACCCGAGAATTCCATCTACAAGACGCTACTTCCCC
3	PvG1235_MT-ATP6_8541	CACCCGAGAATTCCAGTTCGCTTCATTCATTGCCCC
3	PvG1239_MT-CO3_9316	CACCCGAGAATTCCATCCACTCCATAACGCTCCTC
3	PvG1249_MT-ND4_11491	CACCCGAGAATTCCAACGCCTCACACTCATTCTCAA
3	PvG1257_MT-ND5_13515	CACCCGAGAATTCCACCACATCATCGAAACCGCAAA
3	PvG1262_MT-ND6_14664	CACCCGAGAATTCCAGCTTTGTTTCTGTTGAGTGTGG
3	PvG1266_MT-CYB_15260	CACCCGAGAATTCCAAGTCCCACCCTCACACGAT
4	PvG1204_MT-RNR1_656	CACCCGAGAATTCCATGGTCCTAGCCTTTCTATTAGCTC
4	PvG1215_MT-ND1_3398	CACCCGAGAATTCCATACAACTACGCAAAGGCCCC
4	PvG1220_MT-ND2_4711	CACCCGAGAATTCCACCGGACAATGAACCATAACCAA
4	PvG1227_MT-CO1_6547	CACCCGAGAATTCCATCAACACCACCTTCTTCGACC
4	PvG1238_MT-CO3_9210	CACCCGAGAATTCCAACCCACCAATCACATGCCTATC
4	PvG1248_MT-ND4_11410	CACCCGAGAATTCCATAAAGCCCATGTCGAAGCCC
4	PvG1256_MT-ND5_13288	CACCCGAGAATTCCAGGCATCAACCAACCACACCT
4	PvG1265_MT-CYB_15088	CACCCGAGAATTCCACATCGGCATTATCCTCCTGCT
5	PvG1219_MT-ND2_4483	CACCCGAGAATTCCACCCAACCCGTCATCTACTCTAC
5	PvG1226_MT-CO1_6324	CACCCGAGAATTCCAGCCTCCGTAGACCTAACCATC
5	PvG1247_MT-ND4_11223	CACCCGAGAATTCCATAGGCTCCCTTCCCCTACTC
5	PvG1255_MT-ND5_13069	CACCCGAGAATTCCAGCCCTACTCCACTCAAGCAC
5	PvG1264_MT-CYB_14937	CACCCGAGAATTCCACATCAATCGCCCACATCACTC
6	PvG1212_MT-RNR2_2524	CACCCGAGAATTCCAACCAGTATTAGAGGCACCGC
6	PvG1225_MT-CO1_6124	CACCCGAGAATTCCATAATCGGAGGCTTTGGCAACT
6	PvG1246_MT-ND4_10994	CACCCGAGAATTCCAGCAAGCCAACGCCACTTATC

6	PvG1254_MT-ND5_12831	CACCCGAGAATTCCACACAGCAGCCATTCAAGCAA
6	PvG1263_MT-CYB_14789	CACCCGAGAATTCCAAACCACTCATTCATCGACCTCC
7	PvG1210_MT-RNR2_2110	CACCCGAGAATTCCAACAGCTCTTTGGACACTAGGAA
7	PvG1224_MT-CO1_5910	CACCCGAGAATTCCAGCCGACCGTTGACTATTCTCT
7	PvG1245_MT-ND4_10761	CACCCGAGAATTCCATGCTAAAACTAATCGTCCCAACAA
7	PvG1253_MT-ND5_12601	CACCCGAGAATTCCATTCATCCCTGTAGCATTGTTCGT
8	PvG1208_MT-RNR2_1679	CACCCGAGAATTCCATAGCCCCAAACCCACTCCAC
8	PvG1252_MT-ND5_12360	CACCCGAGAATTCCACACCCTAACCCTGACTTCCC
9	PvG1207_MT-RNR1_1347	CACCCGAGAATTCCAGGTGGCAAGAAATGGGCTACA
9	PvG1214_MT-RNR2_2985	CACCCGAGAATTCCACCTCGATGTTGGATCAGGAC
9	PvG1237_MT-ATP6_8992	CACCCGAGAATTCCACTGGCCGTACGCCTAACC
10	PvG1206_MT-RNR1_1127	CACCCGAGAATTCCAACTGCTCGCCAGAACACTAC
10	PvG1213_MT-RNR2_2757	CACCCGAGAATTCCAAGTACCTAACAAACCCACAGGTC
11	PvG1205_MT-RNR1_899	CACCCGAGAATTCCAGCGGTCACACGATTAACCCA
11	PvG1211_MT-RNR2_2323	CACCCGAGAATTCCAATTCTCCTCCGCATAAGCCTG
12	PvG1209_MT-RNR2_1895	CACCCGAGAATTCCACTAAGACCCCCGAAACCAGA

Table 2: Primers used for Illumina NovaSeq.

Primer name	**Complete sequence (5'to 3')**
D8-i7-BC	CAAGCAGAAGACGGCATACGAGATGGTCCAGAGTGACTGGAGTTCCTTGG CACCCGAGAATTCCA
D15-i7-BC	CAAGCAGAAGACGGCATACGAGATGCACATCTGTGACTGGAGTTCCTTGG CACCCGAGAATTCCA
D18-i7-BC	CAAGCAGAAGACGGCATACGAGATTTCGCTGAGTGACTGGAGTTCCTTGG CACCCGAGAATTCCA
Visium-i7-BC	CAAGCAGAAGACGGCATACGAGATAGCAATTCGTGACTGGAGTTCCTTGG CACCCGAGAATTCCA
D8-i5-BC	AATGATACGGCGACCACCGAGATCTACACCGGTTCTTTCTTTCCCTACACG ACGCTC
D15-i5-BC	AATGATACGGCGACCACCGAGATCTACACAACCTCTTTCTTTCCCTACACG ACGCTC
D18-i5-BC	AATGATACGGCGACCACCGAGATCTACACCGCATATTTCTTTCCCTACACG ACGCTC
Visium-i5-BC	AATGATACGGCGACCACCGAGATCTACACCTGCTCCTTCTTTCCCTACACG ACGCTC

Table S1 Sequencing depth and alternative allele depth for different single nucleotide variant (SNV) among samples.

Sample	SNV	min_ad	max_ad	median_ad	min_dp	max_dp	median_dp	n
D8	G>A	585	56596	8737	302725	3177496	1210780	11
	C>A	230	22293	1311	410636	2978545	509345	9
	T>A	4443	119270	10660	408683	1881729	456710	8
	T>G	17985	373100	39945.5	772066	2799698	1111136	8
	C>T	454	2041	574.5	500064	1974934	523341.5	4
	T>C	8185	190953	99569	202044	2334536	1268290	2
	A>G	109115	109115	109115	1324645	1324645	1324645	1
	G>T	1864	1864	1864	2975153	2975153	2975153	1
D15	G>A	354	48787	1730	592201	6736624	1863460	16
	C>A	424	31326	6933.5	219243	6715789	1365450	12
	T>A	12252	216326	20818	669645	2607320	686125	5
	C>T	266	638	419	772331	1590765	1128598	4
	T>C	17656	405852	211754	472786	4779280	2626033	2
	A>T	566330	566330	566330	1150153	1150153	1150153	1
	G>T	761	761	761	3622737	3622737	3622737	1
D18	G>A	457	121132	2368	642801	7883466	2696035	19
	C>A	2011	35420	14075	785291	3444078	1864754	6
	C>T	669	3295	1449.5	2242932	7032808	5078846	4
	T>C	565	461100	9220	437242	5228996	2787804	4
	T>G	31989	164518	51326	2244801	6714234	2847517	4
	T>A	19890	278091	69482	821492	3409369	2132301	3
	A>T	596365	596365	596365	1210494	1210494	1210494	1