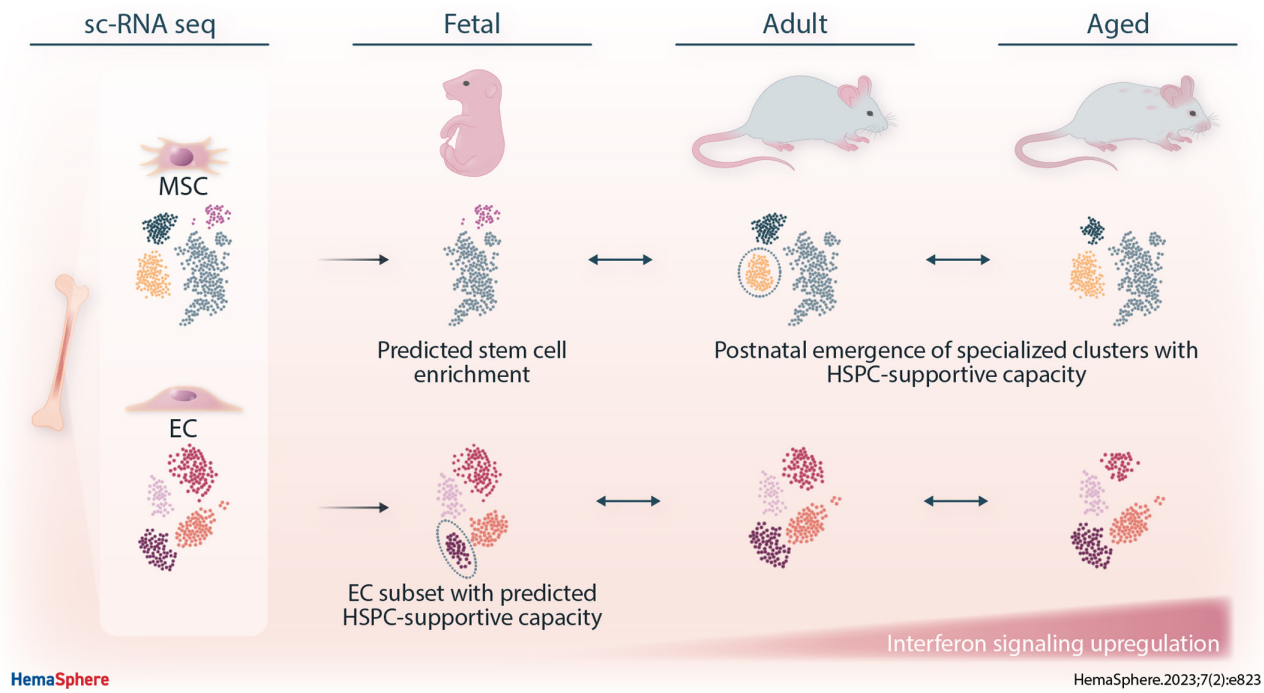
HemaSphere

EHA EUROPEAN HEMATOLOGY ASSOCIATION

Article
Open Access

Resource: A Cellular Developmental Taxonomy of the Bone Marrow Mesenchymal Stem Cell Population in Mice

Paola Pisterzi[1], Lanpeng Chen[1], Claire van Dijk[1], Michiel J. W. Wevers[1], Eric J. M. Bindels[1], Marc H. G. P. Raaijmakers[1]



HemaSphere

EHA EUROPEAN HEMATOLOGY ASSOCIATION

Article
Open Access

Resource: A Cellular Developmental Taxonomy of the Bone Marrow Mesenchymal Stem Cell Population in Mice

Paola Pisterzi[1], Lanpeng Chen[1], Claire van Dijk[1], Michiel J. W. Wevers[1], Eric J. M. Bindels[1], Marc H. G. P. Raaijmakers[1]

Correspondence: Marc H. G. P. Raaijmakers (m.h.g.raaijmakers@erasmusmc.nl).

ABSTRACT

Mesenchymal stem cells (MSCs) play pivotal roles in tissue (re)generation. In the murine bone marrow, they are thought to reside within the Sca-1$^+$ CD51$^+$ bone marrow stromal cell population. Here, using scRNAseq, we aimed to delineate the cellularheterogeneity of this MSC-enriched population throughout development. At the fetal stage, the MSC population is relatively homogeneous with subsets predicted to contain stem/progenitor cells, based on transcriptional modeling and marker expression. These subsets decline in relative size throughout life, with postnatal emergence of specialized clusters, including hematopoietic stem/progenitor cell (HSPC) niches. In fetal development, these stromal HSPC niches are lacking, but subsets of endothelial cells express HSPC factors, suggesting that they may provide initial niches for emerging hematopoiesis. This cellular taxonomy of the MSC population upon development is anticipated to provide a resource aiding the prospective identification of cellular subsets and molecular mechanisms driving bone marrow (re)generation.

INTRODUCTION

Bone marrow stromal cells and endothelial cells (ECs) play pivotal roles in bone marrow development, homeostasis, and regeneration. In the adult bone marrow, they establish favorable niches for hematopoietic stem and progenitor cells (HSPCs), ensuring the lifelong maintenance of hematopoiesis. Studies in murine bone marrow have shown that both populations coordinate osteogenesis, angiogenesis, and hematopoiesis during development and regeneration. At E17, the bone marrow cavity is being vascularized, osteoblast formation and calcification occurs, and hematopoiesis shifts from the fetal liver to the bone marrow, with the first HSPC activity being detected at E16.5.[1,2]

An important contribution to the formation of bone and bone marrow is thought to be made by a subset of stromal cells endowed with stem cell characteristics, including multilineage differentiation ability. These mesenchymal stem cells (MSCs) have the ability to self-renew and differentiate into the osteogenic, chondrogenic, and adipogenic lineage. To allow MSC isolation from the murine bone marrow, multiple combinations of phenotypical markers have been proposed, in particular PDGFRα, Sca-1, and CD51.[3,4]

PDGFRα$^+$ Sca-1$^+$ (PαS) cells are an adult stromal population, adjacent to arteries, enriched for MSC activity. Indeed, PαS cells are able to generate osteoblasts, reticular cells, and adipocytes upon transplantation *in vivo*.[3] Similarly, the majority of Nestin$^+$ cells,[5] express PDGFRα and CD51, which define a population capable of self-renewing and forming a hematopoietic niche in ectopic grafts, upon transplantation *in vivo*.[4] *In vitro* CFU-F activity and multilineage differentiation have further been described for a subset of Sca-1$^+$ CD51$^+$ stromal cells.[6]

Another proposed marker for *bona fide* MSCs in the bone marrow is the leptin receptor (LepR). LepR$^+$ cells are highly enriched for CFU-F and uniformly express PDGFRα and CD51. They arise postnatally, are the main source of new osteoblasts and adipocytes in adult murine bone marrow, and can form bony ossicles that support hematopoiesis *in vivo*.[7]

Collectively, the data indicate that *bona fide* MSCs are present within the bone marrow and reside within a subset of stromal cells expressing PDGFRα, Sca-1, and CD51. Little, however, is known about the cellular heterogeneity of this population in the mammalian bone marrow during development and aging.

Here, we used single-cell RNA sequencing (scRNAseq) to generate a cellular taxonomy of the mammalian MSC and EC populations. The data are anticipated to instruct the identification of putative *bona fide* MSCs and the factors and gene regulatory networks driving bone marrow (re)generation.

MATERIALS AND METHODS

Animal models

C57BL/6J were purchased from Charles River and housed in the Erasmus Laboratory Animal Science Center (EDC). Mice were kept in groups of maximum 4, in pathogen-free conditions, with a standard light/dark cycle, and provided with food and water ad libitum. Animal experiments were approved by the

[1]Department of Hematology, Erasmus MC Cancer Institute, Rotterdam, Netherlands

Supplemental digital content is available for this article.

Copyright © 2023 the Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the European Hematology Association. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

HemaSphere (2023) 7:2(e823).
http://dx.doi.org/10.1097/HS9.0000000000000823.
Received: July 8, 2022 / Accepted: November 29, 2022

Animal Welfare/Ethics Committee of the EDC in accordance with Dutch legislation.

BM isolation and FACS

Fetal bone marrow endothelial and stromal progenitor cells were isolated as described before.[8] Adult (15 weeks, n = 5) and aged (15–18 months, n = 11) mice were sacrificed by cervical dislocation and long bones were collected. Consistent with the workup of fetal bones, bones were crushed and treated with collagenase.[9] Cell suspensions were depleted using a cocktail of biotinylated hematopoietic and lineage markers, including CD45.2 (1:50), B220 (1:100), CD11b (1:100), TER119 (1:100), CD8a (1:100), CD3e (1:100), CD4 (1:100), GR1 (1:100) (Table 1). Following staining, samples were incubated with anti-biotin beads and placed on a magnetic rack to allow separation. The depleted fraction was incubated again with the biotinylated antibodies mentioned earlier, followed by Streptavidin (1:100). Samples were also stained with TER119 (1:25), CD31 (1:100), CD51 (1:50), CD144 (1:50), SCA-1 (1:100), and PDGFRa (1:100). Dead cells were discriminated by 7AAD (1:100) staining. ECs (7AAD$^-$ lin$^-$ CD45$^-$ Ter119$^-$ CD31$^+$) and MSCs (7AAD$^-$ lin$^-$ CD45$^-$ Ter119$^-$ CD31$^-$ Sca-1$^+$ CD51$^+$) were sorted in DMEM with 10% FCS. To generate the adult dataset, 13,000 ECs and 8,500 MSCs were pooled and further processed for sequencing. For the aged dataset, 26,000 ECs and 6,500 MSCs were sorted and pooled with 10,000 HSCs (live, lin$^-$, cKit$^+$, Sca-1$^+$, CD48$^-$, CD150$^+$).

10× sequencing library construction

Library construction was performed as previously described.[8] Samples were sequenced on the Illumina NovaSeq 6000 platform, paired-end mode. A sequencing depth of 47,823, 62,781 and 6507 reads/cells was obtained for the E17,[8] adult and aged sample, respectively. Computational alignment was performed using CellRanger (v 4.0.0, 10× Genomics).

Seurat clustering, integration, and subsetting

Seurat (v 4.0.5) was used for the analysis and integration of the scRNAseq datasets as described before.[10] In brief, for each dataset, cells expressing less than 200 genes, more than 5% of mitochondrial genes, or with a high ratio of counts over features (>5 in adult, >4 in aged) were filtered. Cells expressing the pan-hematopoietic markers *Spn* (Cd43) and *Ptprc* (CD45) were also removed.

To generate a comprehensive dataset including different developmental stages, the adult and aged datasets were integrated with the previously established E17 dataset.[8] Integration was performed using the functions FindIntegrationAnchors and IntegrateData with default parameters. Following data normalization and principle component analysis (PCA), unsupervised clustering was performed and cells were visualized using uniform manifold approximation and projection (UMAP). Cell identity was determined using canonical stromal or endothelial markers. MSC clusters (expressing *Col1a1* and *Prrx1*) or EC clusters (expressing *Cdh5* and *Pecam1*) were extracted using the subset function. For MSCs, PCA and unsupervised clustering (dims = 20, resolution = 0.3) were repeated and the resulting subpopulations were visualized on the UMAP. Similarly, this process was applied to generate new EC clusters (dims = 20, resolution = 0.2).

Table 1

Antibodies and Reagents

Reagent	Supplier	Cat No.
CD45.2	Biolegend	109804
B220	BD Biosciences	553086
CD11b	BD Biosciences	553309
TER119	BD Biosciences	553672
CD8a	BD Biosciences	553029
CD3e	BD Biosciences	553060
CD4	BD Biosciences	553728
GR1	BD Biosciences	553125
Antibiotin beads	Miltenyi Biotec	130-090-485
Streptavidin	Invitrogen	S11223
TER119	Biolegend	116237
CD31	BD Biosciences	565509
CD51	Biolegend	104105
CD144	Biolegend	138015
SCA-1	Biolegend	108120
PDGFRa	eBioscience	17-1401-81
7AAD	Beckman Coulter	A07704

Trajectory analyses

Lineage reconstruction of the single-cell data was determined using the StemID2 functionality from the RaceID3 package (v 0.2.3)[11,12] and the general data-structure from the Seurat package (v 4.0.5), all written in R (v 4.1.0). The trajectories were calculated using the k-nearest neighbors method in iterations with 15 neighbors, to determine the significant links between clusters. Visualization of these links were projected onto the UMAP as generated using Seurat. All other parameters for calculating the connections between clusters were kept at their default values.

Originally, StemID worked only on SingleCellExperiment (SCE) objects from the SCE package (v 1.16.0) as input for the trajectory analysis. As such, minor alterations were made to the StemID software, to accommodate the use of single-cell data in Seurat-like structures, as Seurat was used for all other data-analyses. These alterations entailed taking the gene expression and clustering information from the Seurat object and adding it to an empty SCE object.

Additionally, the visualization functions *plotgraph* and *compscore* based on the cluster-names on 1-indexing, whereas Seurat uses 0-indexing. This was also altered, to minimize differences between the visualization of the trajectory analyses and the visualization done with Seurat.

Cell trajectory analysis was performed using Monocle3 R package (https://github.com/cole-trapnell-lab/monocle3) with a resolution of 0.02. Cluster MSC-0 was set as starting point for pseudotime analysis. The trajectory was projected on the PHATE dimensionality reduction.[13]

Cell cycle score prediction

Cell cycle scores in each dataset were calculated using the CellCycleScoring function available in Seurat. Human reference genes for G_2M and S phase were obtained,[14] and gene symbols were converted from human to mouse to calculate the predicted cell cycle score in each dataset.

DE genes and GSEA analysis

DE genes for MSC and EC clusters were calculated on integrated datasets using the Seurat function FindAllMarkers, which uses the Wilcoxon Rank Sum test by default. For this analysis, only positive markers were selected (only.pos=TRUE) and genes expressed in at least 25% of cells in a cluster were included (min.pct=0.25). The resulting heat maps were plotted showing the top 5 upregulated genes ordered based on avg_log2FC.

In the integrated MSC dataset, clusters 0 and 1 were compared with all others using the function FindMarkers. Both positive and negative markers were included and min.pct was set to 0.25. From this analysis, a ranked gene list was generated as described previously[8] and preranked GSEA was performed

using the GSEA software available from the Broad Institute (v 4.1.0).

To compare DE genes in E17 and adult stage, clusters 0 and 1 of the MSC dataset were isolated using the subset function. For this comparison, the developmental stages were set as new identities and the Seurat function FindMarkers was used as described above. The result of the analysis was visualized as a volcano plot obtained with ggplot2.

Using the SplitObject function on the integrated EC dataset, E17 was isolated. The function FindMarkers was then used to compute marker genes specifically expressed in cluster 1 compared with others, in the E17 dataset.

RESULTS

Single-cell RNA sequencing of bone marrow stromal stem/progenitor and endothelial cells upon development and aging

To study the murine bone marrow niche composition in different stages of mammalian development, we performed scRNAseq on FACS purified MSCs (Ter119$^-$ CD45$^-$ CD31$^-$ CD51$^+$ Sca-1$^+$) and on the EC population (Ter119$^-$ CD45$^-$ CD31$^+$). These highly purified cell populations were isolated from pooled collagenased long bone fractions of fetal (E17, n = 9),[8] adult (15-week-old, n = 5), and aged (15- to 18-month-old, n = 11) mice, generating high quality data of 15,063 MSCs and ECs (Figure 1A, B).

Unsupervised clustering of integrated MSCs and ECs from all developmental stages, performed using Seurat, confirmed the endothelial and stromal identity of the 2 sorted populations. MSCs were defined by the expression of canonical stromal markers such as *Prrx1* and *Col1a1*, while ECs were characterized by the expression of the endothelial markers *Cdh5* (encoding VE-cadherin/CD144) and *Pecam1* (CD31) (Figure 1C).

Cellular heterogeneity and postnatal specification of the bone marrow MSC population in mice

To interrogate the cellular heterogeneity of the Sca-1$^+$ CD51$^+$ MSC population and better define putative stem cell subsets within the population, integrated MSC data from the different developmental stages were subsetted and reclustered, generating a dataset of 5123 cells (1633 E17, 1474 adult, 2016 aged). This analysis revealed the presence of 11 cell types/clusters within the MSC population (Figure 2A), and differential expression (DE) analysis was performed on the integrated dataset to define marker genes for each cluster (Figure 2B).

The composition of the MSC population appeared relatively homogeneous at the fetal stage, where 3 clusters (cluster 0, 1, and 4) constituted the vast majority (85.9%) of cells within the population (Figure 2C, D). Of these clusters, cluster 4 was uniquely present in the fetal stage and was defined by the expression of proliferative markers such as *Birc5*, *Mki67*,

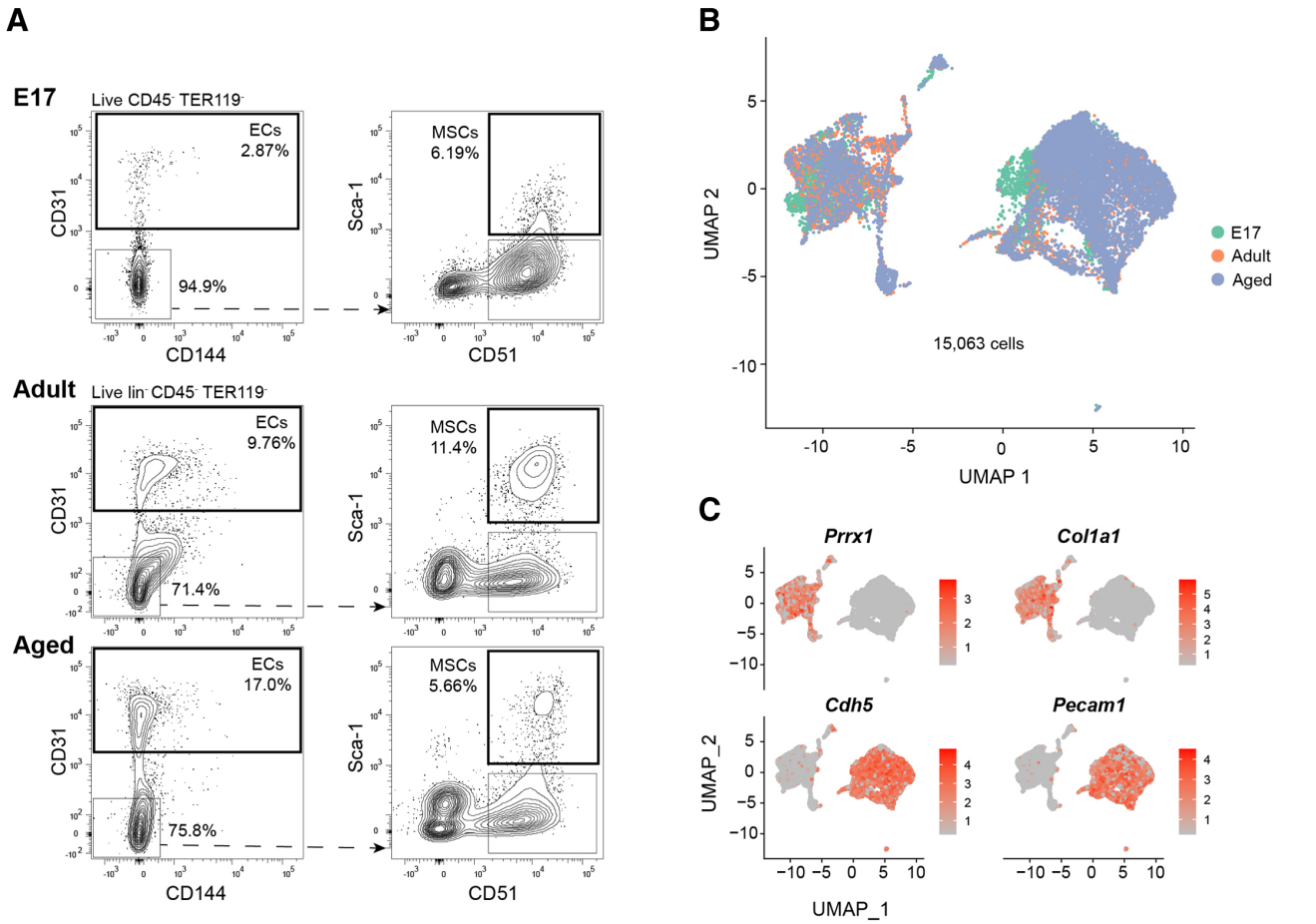


Figure 1. ScRNAseq of the mesenchymal stem cell and endothelial populations in the murine bone marrow upon development. A, Representative flow cytometry plots depicting the sorting strategy used to purify ECs and MSCs from pooled collagenased long bones. Sorted populations are highlighted in bolded boxes. E17 populations (n = 9) had been previously isolated[8] and were reanalyzed to ensure consistent gating with populations in adult (15-week-old, n = 5) and aged (15- to 18-month-old, n = 11) mice. B, UMAP of the integrated datasets shows the distribution of the sequenced populations throughout life stages. C, Feature plots showing the expression pattern of canonical stromal markers (*Prrx1* and *Col1a1*) and endothelial markers (*Cdh5* and *Pecam1*) separating MSCs and ECs. ECs = endothelial cells; MSC = mesenchymal stem cell; UMAP = uniform manifold approximation and projection.

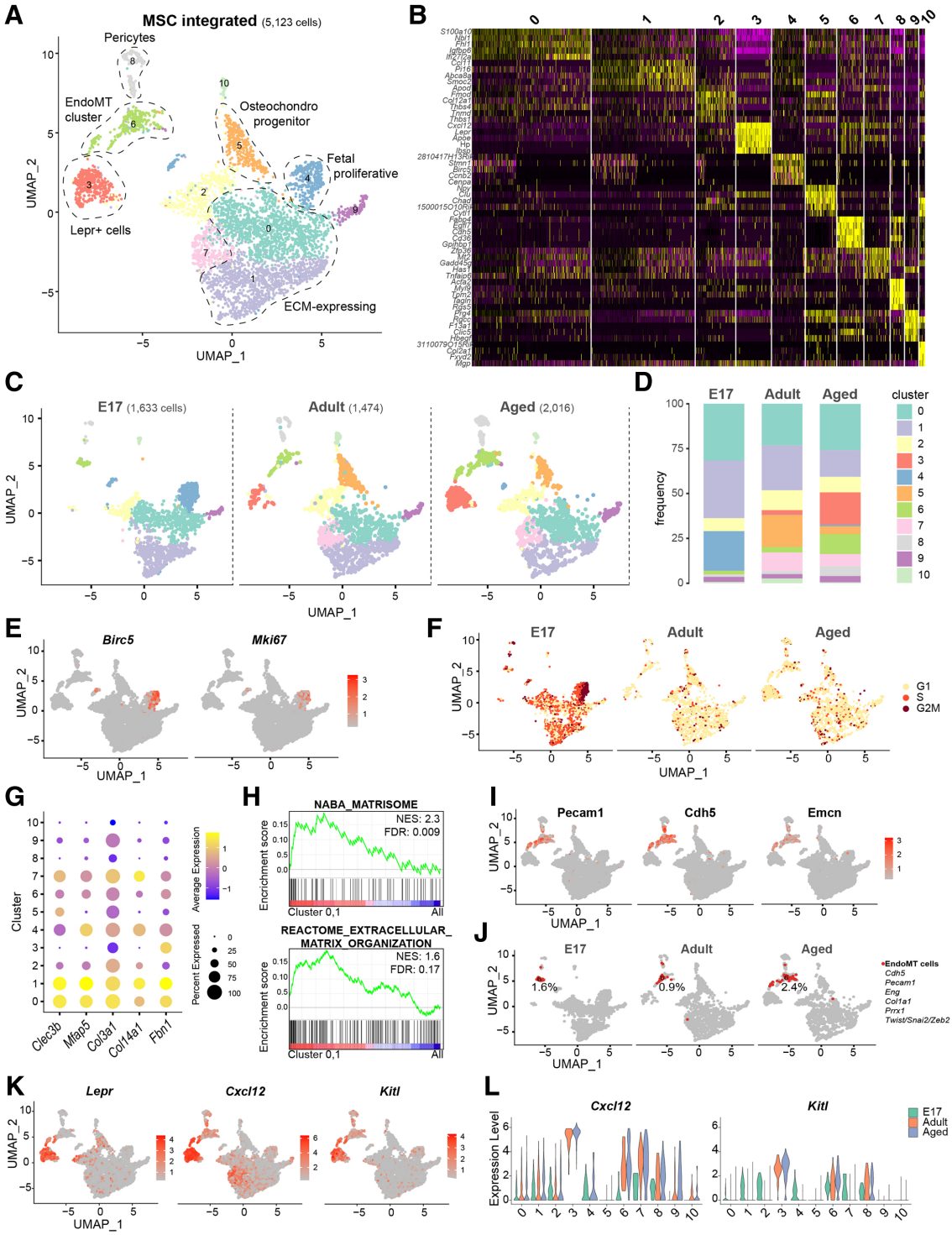


Figure 2. Cellular heterogeneity and specification of the MSC population upon development. A, UMAP representation of the integrated dataset with MSC cluster annotation. B, Heat map of the integrated dataset showing the top 5 marker genes for each MSC cluster. C, D, Cellular heterogeneity of the MSC population in the fetal and postnatal stages. C, UMAP showing MSC cluster composition in fetal, adult and aged stage. D, Bar plot depicting the frequency of each cluster for every stage. Note the postnatal relative reduction of cluster 0–1 and the disappearance of cluster 4. E, F, Identification of a proliferative MSC cluster restricted to fetal development. E, Feature plot showing the expression of proliferation-associated genes *Birc5* and *Mki67* in cluster 4, (F) UMAP representation of cell cycle status shows that in E17, cluster 4 is enriched for genes of S and G_2M phase, while clustering of cells in the adult and aged datasets is not affected by cell cycle state. G, H, Enrichment of genes and transcriptional programs associated with ECM generation in clusters 0–1. G, Dot plot showing the expression of extracellular matrix associated genes. Dot color represents the level of expression, dot size depicts the percentage of cells expressing the marker. H, GSEA showing that clusters 0 and 1 are enriched in extracellular matrix transcriptional signatures. I, J, Identification of EndoMT cells, characterized by coexpression both *bona fide* stromal and endothelium-defining genes, and their relative decline in frequency over time. I, Feature plots showing the expression of the endothelial markers *Pecam1, Cdh5,* and *Emcn* in cluster 6. J, UMAP highlighting the presence of cells coexpressing transcripts consistent with EndoMT. K, L, Identification of a *Lepr*-expressing cluster (cluster 3) marked by genes encoding HSPC regulatory factors. K, Feature plot showing the expression of *Lepr, Cxcl12,* and *Kitl*, underlying cluster 3. L, Violin plots showing the expression level of HSPC regulatory genes throughout development. BM = bone marrow; ECM = extracellular matrix; EndoMT = endothelial to mesenchymal transition; GSEA = Gene Set Enrichment Analysis; MSC = mesenchymal stem cell; UMAP = uniform manifold approximation and projection.

Ccnb2 (Cyclin B2), and *Cenpa* (Figure 2B, E). Cell cycle analysis confirmed that this population is predicted to be highly cycling, as shown by marked expression of genes related to S and G_2M phases (Figure 2F), indicative of expansion of stromal cell subsets, uniquely restricted to fetal development.

Clusters 0 and 1 were conserved throughout life, although their relative frequency within the MSC population diminished postnatally (Figure 2C, D). These clusters were enriched in genes and transcriptional programs related to the deposition of extracellular matrix (ECM) proteins (Figure 2G, H), suggesting that they may play a role in establishment of the marrow matrix and calcification occurring in this phase of development.[15]

The fetal MSC population further consisted of several small subsets, including cluster 8, that we annotated as pericytes based on the expression of *Acta2*, *Tagln*, and *Rgs5* (Suppl. Figure S1A) and a subset (cluster 6) that harbored cells that coexpressed both stromal (*Prrx1*, *Col1a1*) and endothelial (*Pecam1*, *Cdh5*, *Emcn*) markers (Figure 2I, J). We have previously reported that these cells represent *bona fide* stem/progenitor cells undergoing endothelial-mesenchymal transition (EndoMT), capable of generating (*Lepr*-expressing) stromal cells, and reconstituting the entire bone marrow niche.[8] The frequency of EndoMT cells, coexpressing endothelial (*Pecam1*, *Cdh5*), stromal (*Col1a1*, *Prrx1*), and genes encoding EndoMT transcription factors (*Twist*/*Snai2*/*Zeb2*) was relatively stable within the MSC population throughout development (E17: 1.6%, adult: 0.9%, aged: 2.4%).

While the fetal MSC population appeared relatively homogeneous, postnatal stages (adult and aged) were characterized by an increased heterogeneity of the population and the emergence of distinct clusters (Figure 2C, D), in particular clusters 3 and 5.

Cluster 5 expressed both osteogenic (*Spp1*) (Suppl. Figure S1B) and cartilage (*Chad*, *Comp*, *Slurp1*, *Angptl7*) markers[16] (Suppl. Figure S1C), suggesting that it may represent a committed osteochondroprogenitor population, perhaps consistent with the notion that bone formation occurs abundantly in the perinatal and early adult phase, with a relative decrease of the population size upon aging (Figure 2C, D). The expression of the proteoglycan encoding gene *Fmod* and *Nt5e* (CD73) (Suppl. Figure S1D), previously demonstrated to mark *bona fide* osteochondroprogenitor cells, capable of bone regeneration by differentiating into chondrocytes and osteocytes[17–19] further supports this view.

Cells in cluster 3 highly expressed *Lepr*, *Kitl* and *Cxcl12*, in line with the notion that they represent a subset of *Lepr*$^+$ stromal cells that constitutes supportive niches for HSPCs, as previously established (Figure 2K, L).[20,21] These niches have also been referred to as Cxcl12-abundant reticular (CAR) cells,[7,22] marked by the expression of adipogenic genes like *Apoe*, *Adipoq*, and *Lpl* (Figure 2B and Suppl. Figure S1E), but also genes reflecting their predicted osteogenic potential like *Spp1* and *Runx2* (Suppl. Figure S1B, F). Interestingly, this specialized subset of stromal cells with high expression of key HSPC regulatory factors emerged only postnatally, expanded upon aging, and was not present during fetal development, in line with previous reports (Figure 2C, D).[7,23,24]

Taken together, the data suggest that a relatively homogeneous primordial fetal MSC population may specify after E17, with differentiation into distinct committed/specialized subsets, including proposed (*Lepr*$^+$ and *Nt5e*$^+$) bone/cartilage/adipocyte progenitor populations[7,18] and HSPC supportive (*Lepr*$^+$) niches, likely consistent with the notion that bone formation and the establishment of hematopoietic stem cells in the bone marrow occurs in this phase of development.

Identification of candidate stem/progenitor cell subsets within the stromal Sca-1$^+$ CD51$^+$ cell population in early development

One important challenge in the field is the prospective isolation of *bona fide* stem cells from the MSC population, that may be used for regenerative purposes. To assist in identifying candidate stem/progenitor subsets and establish a predicted lineage relationship between MSC subsets, we inferred a lineage tree using StemID2 on the integrated dataset. This algorithm can predict the presence of stem/progenitor states in scRNA-seq data, taking into account delta entropy and number of links between clusters. These 2 parameters are multiplied to calculate a StemID2 score. High scores reflect predicted stem/progenitor states, while lower scores are associated to more differentiated cells.[25]

StemID2 assigned the highest scores to cluster 0–1 (the dominant fetal clusters), 4 (the fetal proliferative cluster), and 6 (the EndoMT-enriched cluster), predicting that these populations might be enriched for progenitor/stem cells (Figure 3A). Pseudotime trajectory analysis suggested close relationships between cluster 0–1 and 4 with these populations, as well as the EndoMT population, being "hubs" of lineage differentiation toward osteochondroprogenitors, *Lepr*$^+$ cells, and pericytes (Suppl. Figure S2A, B).

We have previously shown that EndoMT cells within cluster 6 have the capacity to generate an entire bone marrow organ upon transplantation.[8] To further characterize the stromal subsets represented in cluster 0–1, their transcriptome was compared to all other clusters in the integrated dataset. This revealed expression of genes previously associated with MSC fates, including expression of *Cd34*, *Ly6a*, *and Thy1* (Figure 3B), recently reported to represent a subset of early mesenchymal progenitor cells (EMPs).[24] In particular, both clusters were defined by the expression of *Clec3b*, *Mfap5,* and *Gsn* (Figure 2G and data not shown), described by Zhong et al[24] as additional markers of EMPs. Hematopoietic contamination of these *Cd34*$^+$ subsets was excluded by lack of expression of hematopoietic markers *Gypa* (Ter119), *Ptprc* (CD45), and *Spn* (CD43) (data not shown).

Comparison of the transcriptome of the clusters 0 and 1 at E17 to their postnatal transcriptomes (week 15) (Figure 3C) revealed differential expression of genes encoding proteins implicated in osteogenesis, angiogenesis and hematopoiesis (*Dlk1*, *Igf2*, *Mdk*, and *Ptn*)[26–30] (Figure 3D), processes that are highly active and tightly coupled at E17 in development. The expression of these genes was temporally restricted to the developmental phase and not restricted to the stromal cells in clusters 0 and 1 (Figure 3D).

Collectively, the presence and relative abundance of clusters 0, 1, 4, and 6 in early development, prediction modeling of progenitor state, as well as differential expression of genes related to stemness and developmental processes, suggest that these clusters may represent populations enriched for *bona fide* MSCs. Their identification and the elucidation of their transcriptome may thus instruct the discovery of new markers for their prospective isolation and characterization.

ECs and putative HSPC supportive capacity upon development

ECs, together with MSCs, are major components of the bone marrow niche and play an important role in coordinating osteogenesis, angiogenesis, and hematopoiesis during bone and bone marrow development and regeneration.

To study the dynamic changes in the EC subpopulations upon development, ECs were subsetted from the integrated data, generating a dataset of 9940 total cells (1657 E17, 1630 adult, 6653 aged). Cells were assigned to 8 EC clusters (Figure 4A) and annotated based on differentially expressed genes and the expression of canonical markers for each population (Figure 4B).

Like in MSCs, some endothelial subsets (clusters 5 and 6) were predominantly present at the fetal stage and characterized by the expression of genes associated with cell cycling, such as *Mki67*, *Birc5*, and *Top2a* (Suppl. Figure S3A). The proliferative nature of these clusters was confirmed by cell cycle analysis (Suppl. Figure S3B).

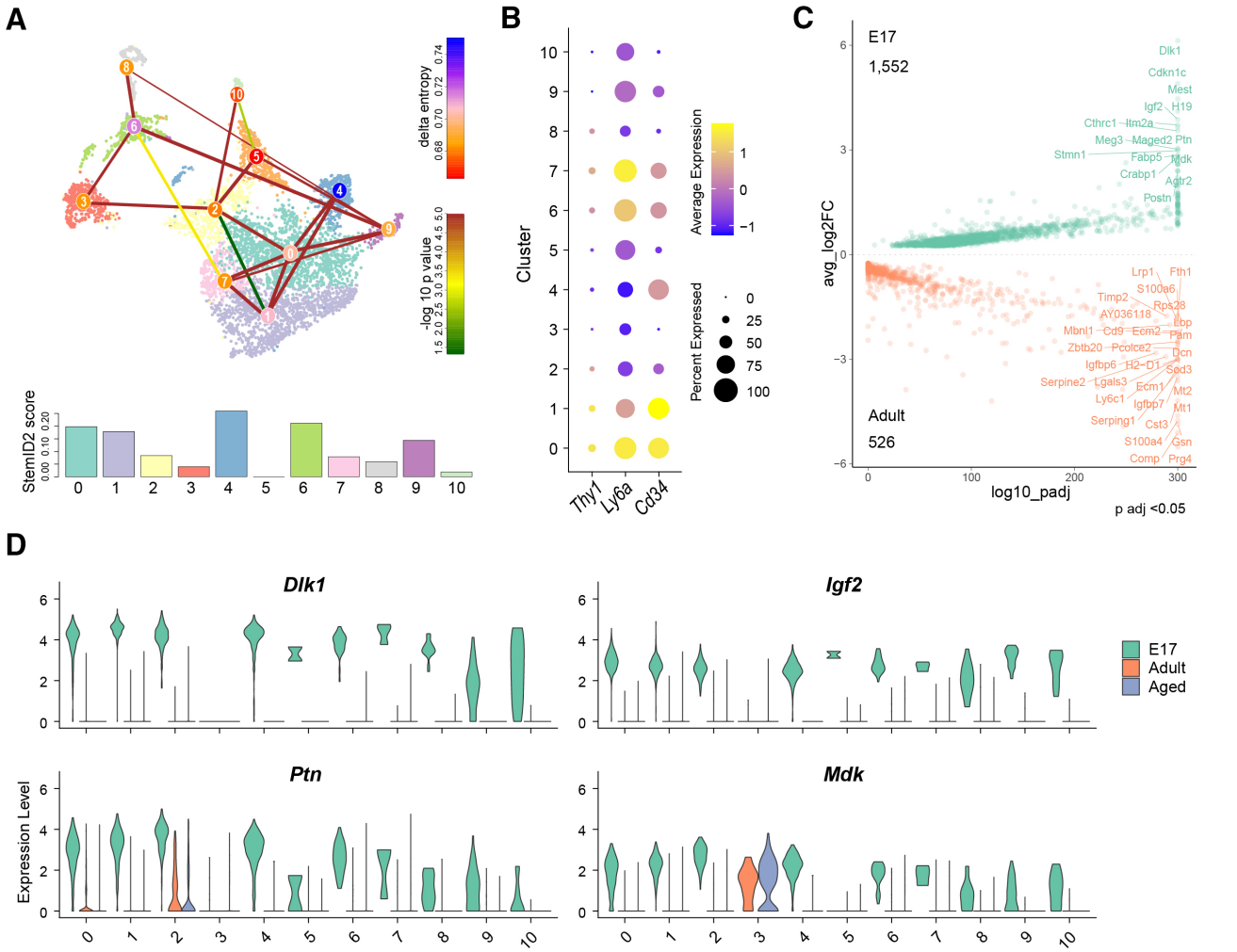


Figure 3. Predictive modeling of stem cell states and their transcriptional wiring within bone marrow MSC subsets. A, Inferred MSC lineage tree in the integrated dataset using StemID2. The color of the circle indicates the delta entropy, a parameter predictive of a stem/progenitor state. The higher the entropy, the more likely the cells in the cluster are stem cells/progenitors. The color of the link between clusters corresponds to the –log10 *P* value, and it indicates the statistical significance of the connection. The thickness of the lines reflects the density of cells covering the link. The bar plot below depicts the StemID2 score calculated multiplying delta entropy by number of links. Clusters 0, 1, 4, and 6 show the highest StemID2 score, predictive of stem/progenitor states. B, Dot plot showing the expression of genes encoding proposed stem cell markers and their distribution among subsets predicted to be enriched for *bona fide* stem/progenitor cells. C, Volcano plot depicting the differential expression of genes in putative MSC-enriched clusters 0–1 comparing fetal development to postnatal (adult) stages. D, Violin plots depicting the expression of genes previously functionally associated with bone marrow development (*Dlk1, Igf2, Mdk,* and *Ptn*) revealing expression temporally restricted to fetal development. MSC = mesenchymal stem cell.

The other clusters could be separated into predicted arterial subsets (clusters 0, 1, and 2), characterized by the expression of *Ly6a*, *Cav1*, and *Sox17* (Suppl. Figure S3C) and sinusoidal cells (cluster 4), characterized by the expression of sinusoidal markers *Stab2*, *Gm1673*, *Lrg1*, and *Tfpi* (Suppl. Figure S3D). The relative abundance of arterial populations in our dataset may be explained by the cell isolation method employed, as enzymatic digestion is known to facilitate the release of cell populations enriched at the endosteal surface, including arterial/arteriolar ECs (Figure 4C, D).[22,31]

At E17, HSPCs migrate from the fetal liver to the bone marrow, followed by their expansion and the establishment of hematopoiesis. Since specialized *Lepr*$^{+}$ stromal HSPC supportive niches develop only after E17, we asked the question whether putative HSPC supportive niches could be identified within EC subsets at the fetal stage (E17).

Interestingly, *Cxcl12*, a gene encoding a factor critically implicated in HSPC homing, was expressed predominantly and at the highest level in arterial cluster 1 in fetal development (Figure 4E), at levels exceeding the expression in stromal cells. ECs within the population also differentially expressed the Notch ligand *Jag1* (Figure 4E), which has a recognized role as a regulator of angiogenesis[32] and HSPC maintenance,[33] suggesting that a subset of ECs within cluster 1 may be implicated in HSPC homing to the bone marrow. In contrast, *Kitl*, an HSPC regulatory factor associated with HSPC proliferation and expansion, was more broadly expressed throughout EC subsets in fetal development (Figure 4E).

Within the arterial populations, cluster 2 highly expressed genes encoding members of the AP-1 transcription factor family (*Jun*, *Junb*, *Jund*, *Fos*, *Fosb*, *Atf3*, and *Atf4*), transcripts reflecting the activation status of NF-κB (*Nfkbia* and *Nfkbiz*), and genes with a role in embryonic vascular development and maintenance of endothelial integrity (*Klf2* and *Klf4*) (Suppl. Figure S3E and data not shown). These genes may be induced by shear stress, a biomechanical force generated by blood flow in the vessel, able to modulate gene expression.[34] The expansion of this arterial cluster might thus be related to the maturation of the respiratory and circulatory systems in postnatal stages[35] (Figure 4C, D).

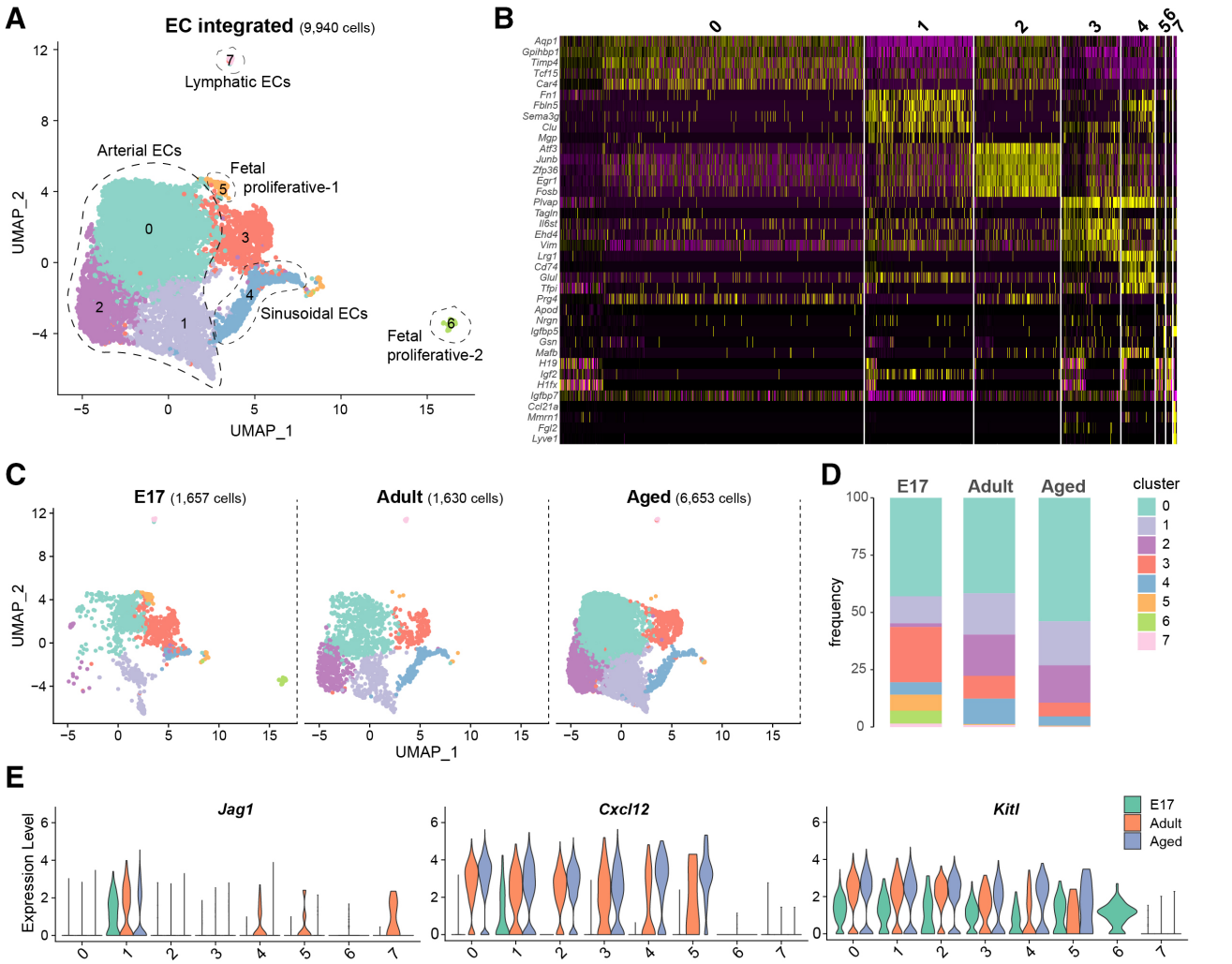


Figure 4. Cellular heterogeneity of bone marrow endothelial cells upon development. A, UMAP representation of the integrated dataset with EC cluster annotation. B, Heat map of the integrated EC dataset showing the top 5 marker genes for each cluster. C, UMAP showing EC cluster composition in fetal, adult and aged stage. D, Bar plot depicting the frequency of each cluster for every stage. E, Violin plots showing the expression level of HSPC regulatory genes in ECs. *Jag1* and *Cxcl12* are specifically expressed in fetal arterial cluster 1. *Kitl* expression is also depicted but is not restricted to a specific fetal EC subpopulation. EC = endothelial cell; UMAP = uniform manifold approximation and projection.

EC cluster 3 coexpressed *Pecam1*, *Emcn*, and *Hif1a* (Suppl. Figure S3F), and its relative abundance declined upon aging (Figure 4C, D); these features were reminiscent of type H endothelium, a specialized EC subset shown to couple angiogenesis and osteogenesis in the bone marrow.[36] A small lymphatic cluster (0.36% of cells in the integrated dataset) was also present and expressed *Thy1*, *Pdpn*, and *Lyve1* (Suppl. Figure S3G), likely of periosteal origin, as the bone and bone marrow do not contain lymphatic vessels.[37]

Taken together, the data indicate that the composition of the EC population changes upon development with the presence of distinct proliferative and candidate HSPC supportive subsets in early development.

Transcriptional alterations in MSCs and ECs upon aging

To obtain insight into the transcriptional alterations in mesenchymal cell populations upon development and aging, we analyzed the transcriptome of MSCs and ECs in the transition from E17 to adult and from adult to aged stage. Interestingly, the number of differentially expressed genes was much higher comparing E17 to adult MSC and EC populations (2230 and 1847 genes, respectively), in comparison to the number of differentially expressed genes comparing adult versus aged populations (318 and 211 for MSCs and ECs, respectively) (Figure 5A and Suppl. Table S1), indicative of a substantial remodeling of the transcriptional landscape in early development, with relatively modest transcriptional alterations upon aging, reminiscent of an earlier report comparing early postnatal development to aging-related stromal changes.[38]

In aged MSCs, 217 genes were significantly downregulated and 101 significantly upregulated (padj < 0.05) in comparison to adult MSCs (Figure 5B). Among the most significantly upregulated genes upon aging were genes associated with interferon signaling (including *Ifi27l2a*, *B2m*, *Bst2*, *Ifitm3*, *C1s1*, *Lgals3bp*, *and Ly6e*) and GSEA revealed upregulation of gene programs associated with activation of interferon signaling (Hallmark Interferon Alpha Response and Hallmark Interferon Gamma Response), as the most significantly enriched gene sets in aged MSCs (Figure 5C, Suppl. Figure S4A, Suppl. Table S1). This finding seems congruent with the aging associated upregulation of inflammatory signaling (and in particular interferon-associated signaling) in PαS cells[38] and may be relevant as IFN signaling in mesenchymal niche cells has been reported to negatively affect their maintenance and hematopoietic support capacity.[39] Upregulation of interferon-related genes did not seem to be restricted to distinct subsets but rather common to most subpopulations (Suppl. Figure S4A).

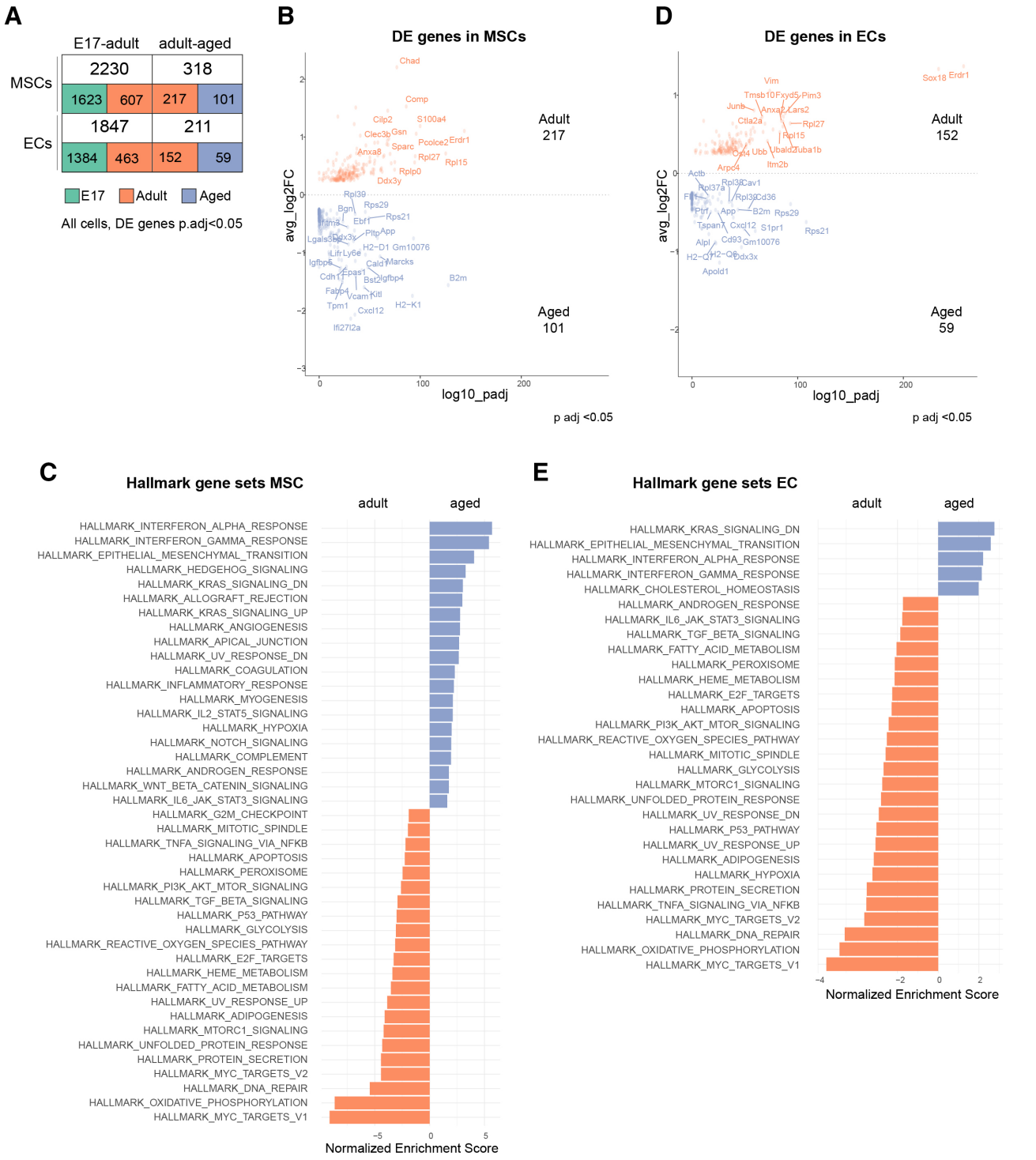


Figure 5. Transcriptional alterations in MSCs and ECs upon aging. A, B, D, A depicts the number of differentially expressed genes in the transition E17 to adult and adult to aged stage in the MSC and EC dataset. Genes were filtered based on avg_log2FC>0.25 and p_val_adj<0.05 and represented on volcano plots for MSCs and ECs, shown in B and D, respectively. C, E, GSEA showing signatures enriched in aged and adult MSCs (C) and in aged and adult ECs (E). Genes and signatures enriched in adult cells are shown in orange, while those enriched in aged cells are represented in violet. EC = endothelial cell; GSEA = gene set enrichment analysis; MSC = mesenchymal stem cell.

Aging was further linked to significant downregulation of genes associated with deposition and organization of the extracellular matrix (*Col1a1*, *Col1a2*, *Sparc*, *Dcn*, *Ecm1*, *Plod2*, *Clec3b*, *Pcolce2*) (Suppl. Figure S4B, Suppl. Table S1) and a Hallmark signature protein secretion (Figure 5C), recapitulating earlier findings in aged PαS cells[38] and suggesting impaired capacity of matrix deposition of aged MSCs.

Other differentially expressed gene sets in aged MSCs included downregulation of genes associated with osteochondroprogenitor cells (*Nt5e* and *Fmod*) (Suppl. Figure S4C) and chondrogenic lineage fates (*Chad*, *Comp*, *Clu*, *Cilp2*, *Angptl7*) (Suppl. Figure S4D), upregulation of genes associated with antigen presentation (*H2-K1*, *H2-D1*, *H2-Q4*) (Suppl. Figure S4E) and upregulation of genes implicated in the complement cascade

(*C3*, *C4b*) (Suppl. Figure S4F), in line with earlier findings in aged PαS cells.[38] Notably, expression of genes encoding complement factors, which have been implicated in the retention of HSCs in the bone marrow[40] was highest in the stromal subsets 3 and 7, which emerge after E17, and its upregulation most prominent in the *Lepr*-expressing subset 3.

Comparison of differentially expressed genes and gene programs in our dataset with those reported to be differentially expressed in *ex vivo* cultured bone marrow stromal cells upon aging (26-month-old mice)[41] revealed only limited overlap in significantly dysregulated genes (3/80 downregulated genes in cultured stromal cells were downregulated in aged MSCs in our study), although enrichment of gene signatures associated with cell cycle regulation and apoptosis in adult versus aged stromal cells were common between the studies. Limited overlap between these data sources likely reflects differences in sensitivity of expression detection between methods, inherent differences between the specific in situ MSC population and bone marrow derived stromal cells, and culture condition-induced transcriptional alterations. In addition, it needs to be emphasized that aging is a continuous process and there may be differences between 15- and 18-month-old mice (in our study) and truly geriatric mice (>24 months old), warranting future studies.

In aged ECs 152 genes were downregulated, while 59 were upregulated (Figure 5A, D, Suppl. Table S2). Among the most significantly downregulated genes were genes encoding proteins implicated in vessel formation and wound healing (*Erdr1*, *Sox17*, *Sox18*, *Junb, Anxa2, CD9*) (Suppl. Figure S5A), while upregulated genes included transcripts related to antigen presentation (*H2-Q7*, *H2-Q6*) (Suppl. Figure S5B) and endothelial integrity and function (*Apold1*, *Cd36*, and *Cd93*) (Suppl. Figure S5C). Altered expression of these genes upon aging was not limited to distinct subsets of ECs in the bone marrow, but rather seemed to reflect more general transcriptional remodeling of ECs upon aging (Suppl. Figure S5).

GSEA revealed signatures associated with IFN alpha and gamma response upon aging, while signatures related to DNA repair, p53 pathway, adipogenesis, and apoptosis were present in the adult stage (Figure 5E), indicating shared transcription programs upon aging in MSCs and ECs.

DISCUSSION

MSCs are thought to play pivotal roles in bone marrow (re) generation. They reside within the Sca-1$^+$ CD51$^+$ population in mice, but their cellular identity and transcriptional wiring remain largely elusive, in part due to the presumed heterogeneity of this stromal population.

Previous single-cell transcriptional analyses of the bone marrow microenvironment in mice have provided unprecedented insights into the heterogeneity of stromal populations.[22,23,42] In these datasets, the entire (endosteal) nonhematopoietic lin$^-$ Ter119$^-$ CD71$^-$[42] or lin$^-$ (CD45$^-$) fraction[22,23] was studied, containing only a minor (<10%) population of cells with high expression of *Ly6a* (Sca-1) and expression of *Itgav*(CD51)/*Pdgfra*, residing within populations relegated either "fibroblast-1 and 2,"[42] "endosteal/stromal/arteriolal fibroblasts"[22] or eBMSC-1/2.[23] Interestingly, the fibroblast-1 and fibroblast-2 clusters were suggested to be "MSC-like" based on expression of the progenitor marker *Cd34* and MSC markers (*Ly6a*, *Pdgfra*, *Thy1*, and *Cd44*),[42] but the low representation of the putative Sca-1$^+$CD51$^+$ MSC population within these datasets limits insights into its heterogeneity.

Here, employing scRNAseq, we generated a cellular taxonomy of the Sca-1$^+$ CD51$^+$ population throughout mammalian development, revealing relative homogeneity of MSCs in fetal development, with the postnatal emergence of distinct cellular subsets, possibly reflecting committed progenitor populations.

The focus on the small Sca-1$^+$ CD51$^+$ subset in our dataset, and the ensuing depth of sequencing, allowed further specification of its heterogeneity and the identification of putative stem cells subsets within the population based on bioinformatic modeling. This revealed a unique proliferative cluster and 3 other clusters predicted by bioinformatic analyses to contain putative stem cells, including a subset expressing both stromal and endothelial-associated genes that we have recently shown to have the capacity to generate an entire bone marrow organ.[8] The fetal proliferative cluster and clusters 0–1 were predicted by the StemID2 analysis to represent stem/progenitor cell populations, expressed genes encoding markers previously associated with MSCs (such as *Clec3b*, *Cd34*, and *Ly6a*) and declined in relative size in aging, a characteristic associated with other stem cell populations. Formal experimental demonstration of stem cell potential of these populations, however, awaits experimental definition, including lineage tracing studies, which may be instructed by the gene expression profiles revealed herein. In contrast with the relatively homogeneous MSC cluster composition defining fetal development, the postnatal MSC population is more heterogeneous and characterized by the emergence of clusters with transcriptional profiles associated with specialized functions.

The postnatal emergence of subsets within the MSC population includes a *Lepr*$^+$ stromal cell population, previously shown to comprise adult stromal progenitor cells with HSPC niche function,[7] in line with earlier published data.[23,43] Of note, the *Lepr*$^+$ cells that we captured within the MSC population is only reflecting a minor subset of *Lepr*$^+$ stromal cells in the murine bone marrow because most *Lepr*$^+$ cells express CD51, but only a small fraction express Sca-1. It may be hypothesized that Sca-1$^+$ LepR$^+$ cells represent primitive progenitors of the Sca-1$^-$ LepR$^+$ cells, but this requires future experimental definition.

Within the postnatal MSC population, we further identified a proposed osteochondroprogenitor population (cluster 5) expressing *Nt5e* (CD73). This marker defined a perivascular stromal population capable of *in vitro* clonal expansion and trilineage differentiation and showing regenerative potential in a murine bone injury model.[18,19] Our lineage tree reconstruction based on StemID2 suggests that this osteochondroprogenitor population might reflect more committed progenitors (with lower StemID2 scores).

In our analyses, we have focused on a bone marrow stromal cell subset that is present in the fetal stage and throughout development and thus likely to contain *bona fide* mesenchymal stem/progenitor cells that are important for bone marrow generation. High granularity analysis of this MSC subset throughout development has the potential to reveal hierarchical structures reflecting early "branching" events in primitive stromal cells, but it needs to be stressed that other bone marrow stromal subsets, including the vast majority of LepR$^+$ stromal cells (which are CD51$^-$) are not represented in this taxonomy, and that these may also contain mesenchymal progenitor cell subsets.[7]

The absence of the LepR$^+$ niche subsets within the MSC population in fetal development indicated that other niche subsets may be present at that stage to support the nascent hematopoiesis, including the homing and expansion of HSPCs. Liu et al[23] recently elegantly demonstrated that arterial endothelial cell-derived Wnt signaling (including Wnt2) control the initial expansion of HSPC in the fetal bone marrow. Unfortunately, we could not detect *Wnt2* expression in our dataset, likely due to the inherently limited depth of sequencing in scRNAseq, but we did identify an arterial cluster in our dataset preferentially expressing *Cxcl12*, warranting future investigations on a potential contribution of the EC subset to early hematopoiesis.

Taken together, our results shed further light on the cellular diversity of the bone marrow MSC population and ECs from

different developmental phases, providing a cellular taxonomy of these populations at single-cell resolution. This resource is anticipated to instruct the identification of key cellular and mechanistic players driving bone marrow generation.

ACKNOWLEDGMENTS

We thank Eline Pronk, Jacqueline Feyen, and Martijn Ernst for providing technical assistance, Bella Banjanin and Remco Hoogenboezem for support with the analysis, and the members of the department of Hematology at Erasmus MC for scientific discussions.

AUTHOR CONTRIBUTIONS

PP and MHGPR did conceptualization. PP, CvD, and EJMB did investigation. LC and CW did software. PP, LC, and CW did formal analysis. PP, LC, and CW did data curation. PP and MHGPR did writing—original draft. PP and MHGPR did writing—review and editing. PP did visualization. MHGPR did supervision.

DISCLOSURES

The authors have no conflicts of interest to disclose.

REFERENCES

1. Coskun S, Chao H, Vasavada H, et al. Development of the fetal bone marrow niche and regulation of HSC quiescence and homing ability by emerging osteolineage cells. *Cell Rep*. 2014;9:581–590.
2. Chan CK, Chen CC, Luppen CA, et al. Endochondral ossification is required for haematopoietic stem-cell niche formation. *Nature*. 2009;457:490–494.
3. Morikawa S, Mabuchi Y, Kubota Y, et al. Prospective identification, isolation, and systemic transplantation of multipotent mesenchymal stem cells in murine bone marrow. *J Exp Med*. 2009;206:2483–2496.
4. Pinho S, Lacombe J, Hanoun M, et al. PDGFRalpha and CD51 mark human nestin+ sphere-forming mesenchymal stem cells capable of hematopoietic progenitor cell expansion. *J Exp Med*. 2013;210:1351–1367.
5. Mendez-Ferrer S, Michurina TV, Ferraro F, et al. Mesenchymal and haematopoietic stem cells form a unique bone marrow niche. *Nature*. 2010;466:829–834.
6. Schepers K, Pietras EM, Reynaud D, et al. Myeloproliferative neoplasia remodels the endosteal bone marrow niche into a self-reinforcing leukemic niche. *Cell Stem Cell*. 2013;13:285–299.
7. Zhou BO, Yue R, Murphy MM, et al. Leptin-receptor-expressing mesenchymal stromal cells represent the main source of bone formed by adult bone marrow. *Cell Stem Cell*. 2014;15:154–168.
8. Kenswil KJG, Pisterzi P, Sanchez-Duffhues G, et al. Endothelium-derived stromal cells contribute to hematopoietic bone marrow niche formation. *Cell Stem Cell*. 2021;28:653–670.e11.
9. Zambetti NA, Ping Z, Chen S, et al. Mesenchymal inflammation drives genotoxic stress in hematopoietic stem cells and predicts disease evolution in human pre-leukemia. *Cell Stem Cell*. 2016;19:613–627.
10. Satija R, Farrell JA, Gennert D, et al. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol*. 2015;33:495–502.
11. Herman JS, Sagar, Grün D. FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat Methods*. 2018;15:379–386.
12. Grün D. Revealing dynamics of gene expression variability in cell state space. *Nat Methods*. 2020;17:45–49.
13. Moon KR, van Dijk D, Wang Z, et al. Visualizing structure and transitions in high-dimensional biological data. *Nat Biotechnol*. 2019;37:1482–1492.
14. Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*. 2016;352:189–196.
15. Coskun S, Chao H, Vasavada H, et al. Development of the fetal bone marrow niche and regulation of HSC quiescence and homing ability by emerging osteolineage cells. *Cell Rep*. 2014;9:581–590.
16. Sunkara V, Heinz GA, Heinrich FF, et al. Combining segmental bulk- and single-cell RNA-sequencing to define the chondrocyte gene expression signature in the murine knee joint. *Osteoarthritis Cartilage*. 2021;29:905–914.
17. Gori F, Schipani E, Demay MB. Fibromodulin is expressed by both chondrocytes and osteoblasts during fetal bone development. *J Cell Biochem*. 2001;82:46–57.
18. Breitbach M, Kimura K, Luis TC, et al. In vivo labeling by CD73 marks multipotent stromal cells and highlights endothelial heterogeneity in the bone Marrow Niche. *Cell Stem Cell*. 2018;22:262–276.e7.
19. Kimura K, Breitbach M, Schildberg FA, et al. Bone marrow CD73(+) mesenchymal stem cells display increased stemness in vitro and promote fracture healing in vivo. *Bone Rep*. 2021;15:101133.
20. Ding L, Morrison SJ. Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature*. 2013;495:231–235.
21. Ding L, Saunders TL, Enikolopov G, et al. Endothelial and perivascular cells maintain haematopoietic stem cells. *Nature*. 2012;481:457–462.
22. Baccin C, Al-Sabah J, Velten L, et al. Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. *Nat Cell Biol*. 2020;22:38–48.
23. Liu Y, Chen Q, Jeong HW, et al. A specialized bone marrow microenvironment for fetal haematopoiesis. *Nat Commun*. 2022;13:1327.
24. Zhong L, Yao L, Tower RJ, et al. Single cell transcriptomics identifies a unique adipose lineage cell population that regulates bone marrow environment. *Elife*. 2020;9:e54695.
25. Grun D, Muraro MJ, Boisset JC, et al. De Novo prediction of stem cell identity using single-cell transcriptome data. *Cell Stem Cell*. 2016;19:266–277.
26. Wu Q, Kawahara M, Kono T. Synergistic role of Igf2 and Dlk1 in fetal liver development and hematopoiesis in bi-maternal mice. *J Reprod Dev*. 2008;54:177–182.
27. Taipaleenmaki H, Harkness L, Chen L, et al. The crosstalk between transforming growth factor-beta1 and delta like-1 mediates early chondrogenesis during embryonic endochondral ossification. *Stem Cells*. 2012;30:304–313.
28. Chen L, Jiang W, Huang J, et al. Insulin-like growth factor 2 (IGF-2) potentiates BMP-9-induced osteogenic differentiation and bone formation. *J Bone Miner Res*. 2010;25:2447–2459.
29. Ohta S, Muramatsu H, Senda T, et al. Midkine is expressed during repair of bone fracture and promotes chondrogenesis. *J Bone Miner Res*. 1999;14:1132–1144.
30. Sato Y, Takita H, Ohata N, et al. Pleiotrophin regulates bone morphogenetic protein (BMP)-induced ectopic osteogenesis. *J Biochem*. 2002;131:877–886.
31. Morrison SJ, Scadden DT. The bone marrow niche for haematopoietic stem cells. *Nature*. 2014;505:327–334.
32. Benedito R, Roca C, Sörensen I, et al. The notch ligands Dll4 and Jagged1 have opposing effects on angiogenesis. *Cell*. 2009;137:1124–1135.
33. Poulos MG, Guo P, Kofler NM, et al. Endothelial Jagged-1 is necessary for homeostatic and regenerative hematopoiesis. *Cell Rep*. 2013;4:1022–1034.
34. Brulois K, Rajaraman A, Szade A, et al. A molecular map of murine lymph node blood vascular endothelium at single cell resolution. *Nat Commun*. 2020;11:3798.
35. Hooper SB, Polglase GR, Roehr CC. Cardiopulmonary changes with aeration of the newborn lung. *Paediatr Respir Rev*. 2015;16:147–150.
36. Kusumbe AP, Ramasamy SK, Adams RH. Coupling of angiogenesis and osteogenesis by a specific vessel subtype in bone. *Nature*. 2014;507:323–328.
37. Edwards JR, Williams K, Kindblom LG, et al. Lymphatics and bone. *Hum Pathol*. 2008;39:49–55.
38. Helbling PM, Piñeiro-Yáñez E, Gerosa R, et al. Global transcriptomic profiling of the bone marrow stromal microenvironment during postnatal development, aging, and inflammation. *Cell Rep*. 2019;29:3313–3330.e4.
39. Goedhart M, Cornelissen AS, Kuijk C, et al. Interferon-gamma impairs maintenance and alters hematopoietic support of bone marrow mesenchymal stromal cells. *Stem Cells Dev*. 2018;27:579–589.
40. Janowska-Wieczorek A, Marquez-Curtis LA, Shirvaikar N, et al. The role of complement in the trafficking of hematopoietic stem/progenitor cells. *Transfusion*. 2012;52:2706–2716.
41. Wilson A, Shehadeh LA, Yu H, et al. Age-related molecular genetic changes of murine bone marrow mesenchymal stem cells. *BMC Genomics*. 2010;11:229.
42. Baryawno N, Przybylski D, Kowalczyk MS, et al. A cellular taxonomy of the bone marrow stroma in homeostasis and leukemia. *Cell*. 2019;177:1915–1932.e16.
43. Mizoguchi T, Pinho S, Ahmed J, et al. Osterix marks distinct waves of primitive and definitive stromal progenitors during bone marrow development. *Dev Cell*. 2014;29:340–349.