Single-cell RNA sequencing identifies phenotypically, functionally, and anatomically distinct stromal niche populations in human bone marrow

Hongzhe Li[1], Sandro Bräunig[1], Stefan Lang[1], Stefan Scheding[1,2]

[1]Division of Molecular Hematology and Stem Cell Center, Lund University, Lund, Sweden; [2]Department of Hematology, Skåne University Hospital, Lund, Sweden

[*]Corresponding author: Stefan Scheding, MD, Molecular Hematology, Lund Stem Cell Center, University of Lund, BMC B12, Klinikgatan 26, 22184 Lund, Sweden, Phone: +46-46-2223331, email: stefan.scheding@med.lu.se

Summary

Hematopoiesis is regulated by the bone marrow (BM) stroma. However, cellular identities and functions of the different BM stromal elements in humans remain poorly defined. Based on single-cell RNA sequencing, we systematically characterized the BM stromal compartment which led to the identification of six transcriptionally and functionally distinct stromal cell populations. Stromal cell differentiation hierarchy was recapitulated based on RNA velocity analysis, in vitro proliferation capacities and differentiation potentials. Potential key factors that govern the transition from stem and progenitor cells to fate-committed cells were identified. Cell-cell communication and in situ localization analysis demonstrated distinct hematopoietic stromal cell niches in specific BM locations, which used either the CXCL12 or SPP1 axis as the major hematopoiesis-regulating mechanism. These findings provide the basis for a comprehensive understanding of the cellular complexity of the human BM microenvironment and the intricate stroma-hematopoiesis crosstalk mechanisms, thus refining our current view on hematopoietic niche organization.

Keywords

Hematopoietic microenvironment, human bone marrow, stromal stem/progenitor cells, single-cell RNA sequencing, hematopoiesis, BM stromal cells, differentiation hierarchy, hematopoietic-stromal communication, CXCL12, SPP1

Introduction

Bone marrow (BM) is the principal site of hematopoiesis in adult humans. In the BM, hematopoietic stem cells (HSCs) and their progenies are contained in specialized microenvironments which regulate HSC maintenance and differentiation.

Our understanding of the BM hematopoietic microenvironment (HME) has evolved considerably over the past decade through a number of landmark studies that have identified the cellular identity, anatomy, and functions of different murine HME components (Baccin et al., 2020; Baryawno et al., 2019; Tikhonova et al., 2019; Wolock et al., 2019). Skeletal and stromal cell populations, endothelial cells and other non-hematopoietic cells but also differentiated hematopoietic cells have been described as important HME elements in mice (Casanova-Acebes et al., 2013; Chan et al., 2018; Ding et al., 2012; Mendez-Ferrer et al., 2008; Urbieta et al., 2010; Yamazaki et al., 2011; Zhao et al., 2014).

In contrast, important HME elements in human BM remain poorly defined, which is due to the fact that the identity and function of the stromal stem/progenitor cells has been difficult to investigate (Li et al., 2016). Nevertheless, a number of recent studies have provided first important insights into the complexity of the human BM and the potential diverse functional roles of BM stromal cells (Chan et al., 2018; de Jong et al., 2021; Triana et al., 2021; Wang et al., 2021).

Based on single-cell RNA sequencing (scRNAseq) technology we herein investigated the human non-hematopoietic BM cell compartment aiming to resolve the composition of the human HME at the highest possible resolution, to identify potential novel marrow stromal subsets and cellular hierarchies as well as to establish functional relationships between stromal and hematopoietic elements.

Results

Single-cell RNAseq identifies distinct cell populations in the human bone marrow microenvironment

We explored the cellular composition of the normal BM stroma by scRNAseq profiling of the non-hematopoietic cell-containing $CD45^{low/-}CD235a^{-}$ population from healthy donors (Fig. 1A). The well-known low frequency of the BM stromal stem/progenitor cells (around 0.001-0.01% of BM cells) (Pittenger et al., 1999) imposed challenges to detect CXCL12-expressing stromal cells in the $CD45^{low/-}CD235a^{-}$ population (Fig S1A). In order to also capture and be able to analyze extremely rare stromal cell subsets, we therefore also sorted $CD45^{low/-}CD235a^{-}CD271^{+}$ cells, which are highly enriched for CXCL12-expressing BM stromal stem and progenitor cells (Fig. 1A, S1B) (Li et al., 2014; Tormin et al., 2011). As $CD45^{low/-}CD235a^{-}CD271^{+}$ cells represent a subset of $CD45^{low/-}CD235a^{-}$ cells, we combined the two datasets which allowed for a detailed analysis of the structural and developmental organization of the BM stroma at the highest possible resolution.

In total, our dataset was comprised of 25067 cells (median of 856 genes and 2937 UMIs per cell, Fig. S1C) which formed 42 clusters corresponding to distinct cell types and differentiation stages, respectively (Fig. 1B, S1D-F, Table S1-2). As expected (Baryawno et al., 2019; Boulais et al., 2018), scRNAseq analysis of $CD45^{low/-}CD235a^{-}$ cells captured several hematopoietic cell types including a cluster of hematopoietic stem/progenitor cells (HSPCs) enriched for CD34 expression and other HSPC markers (PROM1, CRHBP, AVP, MLLT3, FAM30A, GATA1) (Fig. S1G, and Table S1-2). Importantly, non-hematopoietic cells including stromal cells as well as non-stromal cells, such as endothelial cells (cluster 28), KRT5-enriched basal cells (cluster 0), and a cluster of cells enriched for neuronal markers expression (cluster 39) could be clearly identified (Fig. 1C, 1D and S1H). These clusters were detected in both $CD45^{low/-}CD235a^{-}$ and $CD45^{low/-}CD235a^{-}CD271^{+}$ cells (Fig. 1D-E). Of note, following exclusion of hematopoietic cells and the other non-stromal cell types, we identified

nine stromal clusters (clusters 3, 5, 6, 8, 16, 23, 29, 37, 38, Fig. 1B and 1C), which only became detectable after $CD45^{low/-}CD235a^{-}CD271^{+}$ cells were included in the analysis. These findings thus validated our $CD271^{+}$ cell enrichment strategy and proved that a high-resolution transcriptomic analysis of the thus far poorly defined human BM stromal compartment would not have been achieved without $CD45^{low/-}CD235a^{-}CD271^{+}$ enrichment.

Cellular heterogeneity of stromal progenitors

The nine stromal progenitor populations shared a unique gene expression profile, including both well-established stromal markers as well as novel markers (Table S1 and S2) (Li et al., 2014). These markers clearly distinguished stromal cells from the other non-hematopoietic cell types as demonstrated by the differential expression of several stromal genes (Fig. 1D, and S1H).

The stromal clusters were then annotated based on their expression of previously reported BM stromal markers (Churchman et al., 2012; Li et al., 2014), which allowed us to discriminate six different cell types, i.e multipotent stromal stem cells (MSSCs), adipo-primed progenitors, balanced progenitors, pre-osteoblasts, osteochondrogenic progenitors (OCs), and pre-fibroblasts (Fig. 2A, B, S2A).

Cluster 3 was annotated as MSSCs with multi-lineage differentiation capacity based on its high expression of stromal markers (CXCL12, LEPR, DCN, PTGDS) and genes indicating both adipogenic (CEBPD, LPL, PLIN1, ADIPOQ, CCL2, PPARG) and osteogenic (GAS6, FBN1, ALPL, RUNX1, SPARCL1, CDH11) differentiation capacities (Fig. 2C and 2D). Annotation of cluster 3 as MSSCs was further confirmed by velocity analysis as shown in Figure 3A. Cells in cluster 5 expressed high levels of adipogenic differentiation markers as well as a group of stress-related transcription factors such as FOS, FOSB, JUNB and EGR1, which have been shown to mark BM adipogenic progenitors (Fig. 2B-C and S2B) (Ambrosi

et al., 2017). Osteogenic differentiation markers were also expressed but at considerably lower levels compared to adipogenic genes (Fig. 2D). Thus, cluster 5 was termed as adipo-primed progenitors. Cluster 16 on the other hand, showed similar expression levels of adipogenic and osteogenic markers and was therefore annotated as balanced progenitors (Fig. 2A-D). We observed a gradual increase of more mature osteogenic markers such as RUNX1, CDH11, EBF1, and EBF3 from cluster 5 and 16 to cluster 38 (Fig. 2D). Accordingly, cluster 38 was denoted as pre-osteoblasts. Cluster 29 and 23 were characterized by prominent expression of osteochondrogenic markers including BGLAP, CHAD, SPP1, RUNX2, CDH11, CDH2, and SPARC (Fig. 2B, 2E, S2A). Hence, these two clusters were assigned as osteochondrogenic progenitors (OCs). While both adipogenic and osteogenic marker expression was detected in clusters 6, 37 and 8, they also expressed several hematopoiesis-related markers such as SRGN, CD52, CD37, CD48, and PTPRC (Fig. 2B and S2A, 2C). Furthermore, S100 genes, which encode a group of calcium-binding cytosolic proteins, were expressed almost exclusively in cluster 8 (Fig. 2B and 2F), and the previously reported murine fibroblast marker, S100A4 (Baryawno et al., 2019), was highly expressed in all three clusters. Based on this expression pattern, clusters 6, 37, and 8 were subsequently annotated as pre-fibroblasts.

Inferred trajectories of stromal cell gene expression reconstruct the temporal sequence of transcriptomic events during stromal cell differentiation

To resolve the differentiation dynamics of stromal subsets and infer the directionality of individual cells during differentiation, RNA velocity analysis was performed using the scVelo toolkit based on the changes in the spliced/unspliced ratio of mRNA counts (Bergen et al., 2020). Figures 3A and 3A' illustrate the transcriptional dynamics as indicated by streamlines with arrows pointing from the MSSC cluster to different progenitor clusters.

RNA velocity analysis predicted that the MSSC cluster has two main developmental directions (Fig. 3A') with the first originating from MSSCs (cluster 3) and moving towards adipo-primed progenitors (cluster 5), followed by OCs (cluster 29) and terminating at the pre-osteoblast cluster (cluster 38). The second main differentiation path, which also originated from MSSCs (cluster 3), was directed towards balanced progenitors and ended at either the pre-osteoblasts or the pre-fibroblasts stage. A small fraction within adipo-primed progenitors also showed development potential towards balanced progenitors and pre-osteoblasts (Fig. 3A and 3A'). Latent time analysis, which represents the cell's internal clock and approximates the real time experienced by cells as they differentiate, also identified MSSCs as the root cell population and placed pre-osteoblasts close to the terminal state of differentiation (Fig. 3B).

To provide a more precise approximation of the main differentiation stages, we performed latent time and pseudotime analysis on only those clusters that formed a continuum of cellular differentiation states in UMAP, i.e. MSSCs, adipo-primed progenitors, balanced progenitors, OCs (cluster 29), pre-osteoblasts and one of the pre-fibroblast clusters (cluster 6) (Fig. 3C). Thereby, disturbances of the data analysis by larger gaps in cellular state transitions between the clusters within the continuum and the outlying OCs (cluster 23) and pre-fibroblasts (cluster 8 and 37) were avoided (Fig. 3A). The results of these analyses using two different methods which are based on different mathematical algorithms further corroborated our initial conclusion that the MSSC cluster was at the apex of a differentiation hierarchy with downstream differentiation paths into adipo-primed, balanced-, OCs and pre-osteoblasts (Fig. S3A-B). Furthermore, generally high velocity confidences indicated that RNA velocity calculations were reliable (Fig. S3C).

In addition to the identification of stromal differentiation directions and hierarchies, scVelo allowed us to detect potential key driver genes that govern cellular state transitions. Using

likelihood-based computation, we instructed the program to identify the top 300 genes whose transcriptional dynamics correlated with cellular state transitions along the inferred latent time in the stromal continuum (Fig. 3C, Table S3). A group of genes, including CSTA, B2M, CXCL12, and COL1A2 was found to drive the gradual and subtle cellular state transitions within the MSSC cluster (Fig. 3C). Known adipogenic genes such as CEBPD (Cao et al., 1991) together with LEPR, CTSC, LMNA, KLF6 and EGR1 were identified to play important roles in the transition from MSSCs to adipo-primed progenitors (Fig. 3C). As exemplified by CEBPD, EGR1, and LMNA (Fig. S3D-E), gene transcriptional activity and expression level gradually increased from the MSSCs to adipo-primed progenitors. Furthermore, both novel drivers (NFKBIA, RBM3, TUBA1A, and DDX24), as well as reported genes such as ANXA1 (Headland et al., 2015), were identified to serve as critical genes in the transition from adipo-primed progenitors to OCs (Fig. 3C and S3F). VCAN expression has been demonstrated to correlate with bone formation in rats (Nakamura et al., 2005). Accordingly, VCAN was identified as one of the crucial genes for the transition from MSSCs to balanced progenitors and even to the pre-osteoblasts. Other potential novel driver genes along this path included CHL1 and TXNIP (Fig. S3G) and NEAT1 and EIF5 (not shown). These genes demonstrated elevated expression levels in balanced progenitors and pre-osteoblasts.

In summary, RNA velocity analysis of the transcriptional dynamics allowed us to recapitulate stromal cell differentiation pathways and identify key factors, including putative driver genes that govern cell fate commitment.

Surface molecule profiling identifies marker combinations for the prospective isolation of functionally distinct stromal progenitor populations

Based on the expression patterns of stromal markers, osteochondrogenic genes, and fibroblastic markers, the nine stromal subsets could be further assigned to three main groups (Fig. 2 and S4A). Group A cells (clusters 3, 5, 16, 38), which included MSSCs, adipo-primed progenitors, balanced progenitors, and pre-osteoblasts, showed high expression of stromal markers, adipogenic and osteogenic differentiation genes. Group B cells (clusters 29 and 23), annotated as OCs (Fig. 2), demonstrated exclusive and specific expressions of osteochondro-lineage markers. Group C cells (clusters 6, 37, and 8), assigned as pre-fibroblasts (Fig. 2), exhibited characteristic expression of a group of S100 genes, such as S100A4, S100A6, S100A8, and S100A9. These clear differences in gene expression between the groups pointed to potential functional differences. We therefore went on to identify suitable cell surface markers and marker combinations, respectively, that would allow for the prospective isolation and functional investigation of the three groups of cells.

As shown in figure 4A, surface marker gene expression differed between groups A, B, and C. Group B cells showed exclusive expression of NCAM1 and CD9, which were considerably lower expressed by group A and C cells (Fig. 4A). Group C cells expressed several surface molecules which are conventionally considered as hematopoietic markers (CD52, CD37, PTPRC, and CD48). Thus, lack of expression of NCAM1, CD9 and group C-specific antigens could be used to identify group A cells, which could furthermore be separated into different subpopulations based on the differential expression of CD81.

Based on this, we designed fluorescence-activated cell sorting (FACS) strategies for the isolation of group A ($CD45^{low/-}CD235a^{-}CD71^{-}CD271^{+}NCAM1^{-}CD52^{-}$), group B ($CD45^{low/-}CD235a^{-}CD71^{-}CD271^{+}NCAM1^{+}CD52^{-}$) and group C ($CD45^{low/-}CD235a^{-}CD71^{-}CD271^{+}NCAM1^{-}CD52^{+}$) cells (Fig. 4B). CD71 (TFRC) was included as an additional exclusion marker to remove residual CD71-expressing erythroid cells which were present in the $CD45^{low/-}CD235a^{-}$ population (Fig. S1C). Group A cells ($CD45^{low/-}CD235a^{-}CD71^{-}$

CD271^{+}NCAM1^{-}CD52^{-}) were further separated into four subpopulations (A1-A4) based on CD81 expression, which corresponded to MSSCs (A1: CD81^{++}), adipo-primed progenitors (A2: CD81^{+}), balanced progenitors (A3: CD81$^{+/-}$), and pre-osteoblasts (A4: CD81^{-}) (Fig 4B). Functional analysis of sorted cells showed that the MSSC population (A1: CD45$^{low/-}$CD235a^{-}CD71^{-}CD271^{+}NCAM1^{-}CD52^{-}CD81^{++}) had the highest potential to form fibroblastic colonies (colony-forming units, fibroblast: CFU-F) and multi-lineage differentiation capacity toward osteoblasts, adipocytes, and chondrocytes (Fig. 4C-D, S4B). Although the CFU-F frequency of adipo-primed progenitor cells (A2: CD45$^{low/-}$CD235a^{-}CD71^{-}CD271^{+}NCAM1^{-}CD52^{-}CD81^{+}) was lower compared to MSSCs, they still gave rise to a significant fraction of colonies and showed full in vitro differentiation capacities (Fig. 4C-D, S4B).

Colony-forming capacities of balanced progenitors (A3: CD45$^{low/-}$CD235a^{-}CD71^{-}CD271^{+}NCAM1^{-}CD52^{-}CD81^{+-}), pre-osteoblasts (A4: CD45$^{low/-}$CD235a^{-}CD71^{-}CD271^{+}NCAM1^{-}CD52^{-}CD81^{-}) and OCs (B: CD45$^{low/-}$CD235a^{-}CD71^{-}CD271^{+}NCAM1^{+}CD52^{-}) were considerably lower in comparison with MSSCs and adipo-primed progenitors (Fig. 4C and S4B). Pre-fibroblasts (group C) showed the lowest potential to form fibroblastic colonies (Fig. 4C and S4B). Furthermore, the adipogenic potential of pre-fibroblasts was markedly reduced and the pre-fibroblast-derived chondrocyte pellets were smaller compared to the other groups while osteogenic differentiation was normal (Fig. 4D), thus indicating a limited differentiation potential of this population. OCs (group B) demonstrated an enhanced capacity to differentiate toward osteoblasts and chondrocytes, however, at the expense of adipogenic potential (Fig. 4D). These findings suggest a skewed differentiation potential of OCs, which is consistent with the lower expression of adipogenic markers and elevated expression of osteochondrogenic markers (Fig. 2C and 2E). The in vitro differentiation capacities of balanced progenitors (A3) and pre-osteoblasts (A4) could not be assayed due to insufficient proliferation upon culture expansion.

Taken together, the data-driven FACS gating strategy allowed us to isolate distinct stromal subsets for functional investigation. Both MSSCs and adipo-primed progenitors demonstrated considerably higher in vitro colony-forming capacities in comparison with balanced progenitors, OCs, pre-osteoblasts, and pre-fibroblasts. While MSSCs and adipo-primed progenitors exhibited full multi-differentiation potentials, pre-fibroblasts had only limited differentiation potentials. OCs showed enhanced osteoblastic and chondrogenic differentiation capacities and compromised adipogenic potential, which is consistent with their differentiation marker expression profiles (Fig. 2).

In-situ localization of stromal cell subpopulations

Based on the markers identified above, we used CD271 (NGFR), CD81, NCAM1, and CD45 to identify stromal cells ($CD271^{+}$) and to investigate the in-situ localization of MSSCs ($CD271^{+}CD81^{++}$) as well as OCs ($CD271^{+}CD56^{+}$) using bone marrow biopsies from hematologically normal donors. As shown in Figure 4E and S4C, CD271-expressing cells were found in perivascular regions, endosteal regions, and throughout the stroma. We observed that perivascular cells surrounding the capillary endothelium and larger vessels primarily expressed CD271 alone. The bone-lining cells proximal to the surface of trabecular bone showed expression of both CD271 and NCAM1, confirming the osteochondrogenic nature of OCs. In contrast to these CD271/NCAM1 double-positive cells, MSSCs ($CD271^{+}CD81^{++}$) were mainly localized in the adipocyte-enriched regions and in close proximity and encircling adipocytes, respectively. Co-staining with CD45 confirmed that perivascular, periadipocytic, and bone-lining cells were CD45 negative. These results demonstrated the distinct anatomical localizations of MSSCs and OCs, which thus confirms and extends our previous findings on differently localized putative hematopoietic niche cells

(Tormin et al., 2011) and point to possible divergent physiological functions of the different stromal subsets.

Cell-cell interactions in the human bone marrow microenvironment

It has been proposed that different niches exist for different types of hematopoietic stem and progenitor cells. We therefore went on to study the interactions between the different human BM stromal cell populations identified herein and hematopoietic cells based on ligand-receptor (LR) expression using CellPhoneDB (Efremova et al., 2020) complemented by analysis of additional published LR pairs for erythroid cells (Kleven et al., 2018).

As shown in Figure 5A and S5A, we identified a wide range of interactions between stromal cells and hematopoietic cells. Generally, cells belonging to the three stromal groups (A-C) had multiple interactions with a large number of hematopoietic cell types, including HSPCs and more mature hematopoietic cells (Fig. 5A and S5A). Interestingly, only a few LR pairs were identified between stromal cells and plasma cells (Fig. 5A).

To study the functional relationships between stromal and hematopoietic cells in more detail and to identify subset-specific interactions, we plotted selected ligand-receptor pairs that are known to be relevant for hematopoietic-stromal crosstalk (Fig. 5B). Selected LR pairs, including essential cytokines and established hematopoiesis-supporting molecules were analyzed for effects from stromal cells on hematopoiesis and vice versa, including stromal subset-specific interactions, intra-, and inter-stromal cell interactions, and other potentially important interactions (Fig. 5B, S5B, S7B).

<u>Stromal cell regulation of hematopoiesis</u>

Regarding the regulatory and supportive effects of stromal cells on hematopoiesis, we found that group A and C stromal subsets interacted with essentially all hematopoietic cell types

through similar pathways. CXCL12 was highly expressed by group A and C stromal cells, especially the MSSC population (Fig. 6A), and CXCL12-CXCR4 crosstalk was detected in almost all hematopoietic clusters except for erythroid progenitors (Fig. 5B, S5B, and 6A). Furthermore, large dot sizes and high statistical power indicated that CXCL12-mediated interactions represent one of the major interaction mechanisms between these stromal cells and hematopoietic cells (Fig. 5B, S5B). In contrast, CXCL12 interactions with CXCR3 and DPP4 were less significant (Fig. 5B and 6A).

KITLG (SCF)-involving stroma-hematopoiesis interactions were also identified. However, these interactions were not as strong as those involving CXCL12 (Fig. S6A). Interestingly, MDK, a neurite growth factor expressed by stromal cells, showed interactions with the HSPCs, lymphoid lineage progenitors (B, T and NK), and monocytes through LRP1 or SORL1 (Fig. S6B), which has thus far only been described in mouse fetal liver (Gao et al., 2021). Furthermore, other cytokines known to be important for the maintenance of $CD34^+$ cells include FLT3-ligand (FLT3LG) and thrombopoietin (THPO). However, whereas expression of FLT3 was detected in several hematopoietic cell types including HSPCs, dendritic cells and monocytes (Fig. S6C), THPO and MPL expression were barely detected (Fig. S6D). JAG1 expression was highly enriched in stromal cells while NOTCH1, 2 and 4 expressions were detected in hematopoietic cells, suggesting that activation of Notch signaling is involved in stromal-hematopoietic interplay (Fig. S6E). All the three stromal groups were furthermore found to interact with T cells through the IL7-IL7R RL combination, indicating a T cell supporting mechanism by stromal cells (Fig. 5B, 6B, S5B).

In addition to soluble factor-mediated interactions, we also identified direct cell-to-cell communication between group A and C stromal cells and hematopoietic cells mediated by VCAM1 (Fig. 6C), FN1, PLAUR and TNC. These results indicated that group A and C cells

can support hematopoietic cells via direct cell-to-cell contact as well as hematopoiesis-supporting cytokines.

In contrast to the strong contribution of CXCL12-mediated interactions between group A and C stromal cells and hematopoietic cells, SPP1-mediated interactions were predominant in the crosstalk between group B stromal cells (OCs) and hematopoietic cells (except for megakaryocytes), indicating a stromal group-specific interaction (Fig. 5B and S5B). Interacting partners of SPP1-expressing OCs were erythroid cells, lymphoid progenitors (B, T, NK progenitors) and CD34-enriched HSPC cluster which expressed SPP1 binding partners such as CD44 and PTGER4 (Fig. 5B, 6D and S5B). SPP1-involving integrin alpha 4 beta 1 (a4b1) and alpha 9 beta 1 (a9b1) complexes could also play important roles in the interaction of OCs with a number of hematopoietic cells. Consistent with this, FACS analysis demonstrated that SPP1 expression was higher in OCs in comparison with MSSCs (Fig. S6F). Moreover, in situ staining demonstrated that SPP1-expressing cells were exclusively colocalized endosteally with NCAM1^{+} OCs (Fig. 6E), pointing to a localization-specific regulation.

Taken together, these data demonstrated that the most predominant interactions were CXCL12- and SPP1- mediated interactions for group A/C stromal cells and OCs, respectively, indicating that regulation of hematopoiesis is stromal group and location specific.

<u>Erythroid cell-specific interactions</u>

Unlike the majority of the hematopoietic cells in our dataset, the interplay between erythroid cells and group A and C stromal cells was not dependent on CXCL12-involving interactions as indicated by lacking expression of CXCL12 receptors in erythroid clusters (Fig. 6A). CellPhoneDB analysis demonstrated that while group A and C stromal cells interacted with the erythroid cells mainly through VCAM1-, FN1-, TNC- and PLAUR- involving complexes,

the interactions between OCs (group B) and erythroid cells were mainly mediated by SPP1-CD44 (Fig. 5B, 6D and S5B). In addition to that, we found that TF (transferrin) was exclusively expressed by all stromal clusters and its corresponding receptors, TFRC and TFR2 were highly expressed by erythroid progenitors (Fig. 6F).

Hematopoietic regulation of stromal cells

Aside from the supportive and regulatory role of stromal cells for hematopoiesis, it is well described that regulatory signals originating from hematopoietic cell types can affect stromal cells (Baksh et al., 2005). Accordingly, we found that the crucial stromal cell growth factors PDGFs (PDGFA, PDGFB and PDGFC) were highly expressed by megakaryocytes while their corresponding receptors (PDGFRA and PDGFRB) were predominantly expressed by stromal cells (Fig. 6G), suggesting interactions through these axes (Fig. 5B, S5B). Our analysis also predicted that hematopoietic cells control stromal cells via the TGFB1-EGFR and TGFB1-TGFBR1/2/3 axes (Fig. 5B, S5B, S6G). Furthermore, Oncostatin M (OSM) expression was found in various hematopoietic cell types, indicating a possible regulation of stromal cells through its receptors, OSMR and LIFR (Fig. 5B, S5B, S6H).

Together, these data indicate that the regulatory effects between stromal cells and hematopoietic cells are bidirectional, which is consistent with the increased fibroblastic colony size in the presence of hematopoietic mononuclear cells in CFU-F assays (Fig. S6I).

Inter- and intra-stromal signaling

As shown in Figure S7A, CFU-F frequencies were significantly reduced when stromal cells were cultured under single-cell conditions in comparison with bulk-cultured cells, suggesting the presence of inter- and intra-stromal signaling among stromal cells. Indeed, multiple stimulatory pathways were operative between different stromal subsets as well as within the

same stromal cell population (Fig. S7B). Among them, FGF receptor-mediated signaling and various collagen-involving pathways were identified that reflected interactions between different stromal cells (Fig. S7C). In addition, CXCL12-, MDK-, GRN-, NRP1- and BST2-mediated interactions were also found to significantly contribute to the regulatory mechanisms of stromal cells (Fig. S7B-D).

In summary, our data demonstrated all three stromal groups can interact with a wide range of hematopoietic cells. Group A and C cells communicate with hematopoietic cells mainly through CXCL12-, VCAM1-, FN- and MDK-involving pathways while group B cells interact mainly via SPP1-mediated pathways. On the other side, stroma cells are regulated by hematopoietic cells through PDGF-, OSM-, and TGFB1-mediated signaling. Additionally, multiple intra- and inter-stromal pathways were identified. Our in silico analysis of scRNAseq data of both stromal and hematopoietic clusters thus predicted the presence of a complicated interaction network in human bone marrow microenvironment.

Discussion

The BM microenvironment plays a critical role in regulating hematopoiesis. Recent landmark studies have illustrated the molecular complexity of the murine BM microenvironment (Baccin et al., 2020; Baryawno et al., 2019; Tikhonova et al., 2019). However, the exact definition of the cell populations that form the BM stroma in humans remains elusive. In this study, based on single-cell RNA sequencing technology, we provide a comprehensive description of the non-hematopoietic cellular components of this important organ including their developmental and functional roles in hematopoiesis.

Due to the low frequencies of the stromal stem/progenitor cells, it has been difficult to investigate the potential heterogeneity of the human BM stromal cell compartment even when employing single cell approaches (Triana et al., 2021). In this study, we therefore used not

only BM stroma-enriched $CD45^{low/-}CD235a^{-}$ cells but added also sorted $CD271^{+}$ cells, which are highly and exclusively enriched for stromal stem/progenitor cells (Tormin et al., 2011). This approach ensured that no stromal cells were excluded from our analysis while stromal cell resolution was substantially increased.

Following exclusion of the hematopoietic clusters, we identified primarily nine stromal clusters which could further be assigned to six stromal cell types, i.e. MSSCs, adipo-primed progenitors, balanced progenitors, OCs, pre-osteoblasts, and pre-fibroblasts. Of note, corresponding murine stromal cell types have been reported for most of these cell types (MSSCs, adipo-primed progenitors, OCs, pre-osteoblasts, and pre-fibroblasts), whereas the cluster annotated as balanced progenitors has not been described neither in mice nor in humans (Baryawno et al., 2019; Tikhonova et al., 2019). Thus, this analysis allowed us to dissect the human bone marrow stromal cells at a thus far unreached molecular level, to solve the stromal cell heterogeneity and to generate a detailed transcriptional fingerprint of distinct stromal populations.

Furthermore, we were able to define and functionally validate the developmental relations and hierarchies of the different human BM stromal populations. We identified MSSCs as the most primitive population, which is in accordance with published murine and human stromal stem cell studies (Baryawno et al., 2019; Wolock et al., 2019; Zhou et al., 2014). Pre-fibroblasts, pre-osteoblasts, and OCs were placed downstream of MSSCs in the developmental hierarchy, thus confirming the reliability of the velocity analyses. By using a likelihood-based computation, we also detected potential key driver genes that govern the transition between different cellular states. Identified genes included both novel and established transcription factors and certainly, these data provide the basis for future important mechanistic studies of the dynamic processes associated with cellular state transition and fate choices.

Additionally, we were able to provide a phenotypical definition that allowed for the faithful identification and prospective isolation of molecularly defined stromal stem and progenitors. These cells have thus far remained incompletely characterized and our data thus complement and extend reported phenotype definitions of MSCs from hematopoietically active bone marrow and skeletal tissues (Chan et al., 2018; Delorme et al., 2008; Gronthos et al., 2007; Jones et al., 2002; Li et al., 2016; Quirici et al., 2002; Sacchetti et al., 2007).

In mice, elegant studies using genetic approaches involving in-vivo labelling of cell types and ablation of candidate niche cells and niche factors have provided detailed insight into the functional roles of stromal cells in regulating hematopoiesis (Pinho and Frenette, 2019). Comparable studies in human are impossible to perform and analysis of human stromal cells function has been limited to in-vitro models. We therefore investigated cellular interactions based on ligand-receptor interactions and found a broad and complicated crosstalk network, which is likely to reflect – at least in part – the in-situ situation as data were generated with directly isolated cells. We found that stromal cells generally showed hematopoiesis supporting potential, which is consistent with previously published in-vitro coculture data (Li et al., 2014; Li et al., 2020) and which we confirmed in co-cultures with BM $CD34^+$ cells and MSSC- or OC-derived stromal feeder cells (Fig. S7E). Furthermore, whereas all stromal cell groups showed the potential to interact with almost all hematopoietic cells, cell-cell interaction mechanisms differed considerably between different groups of stromal cell groups, i.e. between group A/C cells which included the MSSCs and group B cells (OCs). Whereas the former interacted with hematopoiesis through cytokines such as CXCL12, IL7, KITLG, MDK, and TF and adhesion molecules mediated direct cell-to-cell contact (e.g., VCAM1), we identified SPP1-mediated interactions as the main pathways for the latter. These results thus indicated that hematopoietic cells were maintained by different stromal populations through diverse but nevertheless stromal cell-specific pathways. Taking furthermore into

account that SPP1-expressing cells were located endosteally and that CD271 single positive cells and MSSCs were localized in the perivascular regions and stromal regions, respectively, suggested that different stromal cells provide specialized niches for hematopoietic cells in different locations. Our data thus confirm and extend previous reports describing that endosteal and perivascular niches exist in human BM (Pinho et al., 2013; Tormin et al., 2011), and conceptually support a model where different hematopoietic cell types are differentially regulated by distinct niche milieus which are composed of specific cytokine-producing stromal cells.

Limitations of Study

In this study we used human BM aspiration samples in which bone-lining or bone-attaching cells are most likely underrepresented. Our experiments could therefore ideally be complemented by studies on fresh bone marrow biopsy samples from healthy individuals, which however were not available to us, and which might be difficult to acquire because of ethical considerations. The potential key genes that govern the stromal cells fate commitment were predicted using likelihood-based computation and the stroma-hematopoiesis crosstalk mechanisms were inferred based on ligand-receptor interactions. The identified key genes and interactions will need to be validated in future experiments, for example using in vitro engineered human BM models or in vivo humanized BM ossicles (Dupard et al., 2020). Despite these limitations, we believe that our results provide the basis for a better understanding of the cellular complexity of the human BM microenvironment and put a new light on the current concept that specialized niches exist for distinct types of hematopoietic stem and progenitor cells (Morrison and Scadden, 2014).

Materials and Methods

Human bone marrow mononuclear cell isolation

Human bone marrow (BM) cells were collected at the Hematology Department, Skåne University Hospital Lund, Sweden, from consenting healthy donors by aspiration from the iliac crest as described previously (Li et al., 2014). Bone marrow mononuclear cells (BM-MNC) were isolated by density gradient centrifugation (LSM 1077 Lymphocyte, PAA, Pasching, Austria). The use of human samples was approved by the Regional Ethics Review Board in Lund, Sweden.

Flow cytometry and fluorescence activated cell sorting (FACS)

For sorting of BM-MNCs for scRNAseq analysis, freshly isolated BM-MNCs were incubated in blocking buffer [DPBS w/o Ca2+, Mg2+, 3.3 mg/ml human normal immunoglobulin (Gammanorm, Octapharm, Stockholm, Sweden), 1% FBS (Invitrogen)], followed by staining with monoclonal antibodies against CD271, CD235a, and CD45. Sorting gates were set according to the corresponding fluorescence-minus-one (FMO) controls and cells were sorted on a FACS Aria II or Aria III (BD Bioscience, Erembodegem, Belgium). Dead cells were excluded by 7-Amino-actinomycin (7-AAD, Sigma) staining and doublets were excluded by gating on FSC-H versus FSC-W and SSC-H versus SSC-W. Cells were directly sorted into PBS with 0.04% BSA for scRNAseq analysis. Sorting of BM-MNCs for CFU-F or stromal culture was performed using DAPI for dead cell exclusion and cells were directly sorted into stromal cell culture medium [StemMACS MSC Expansion Media, human (Miltenyi Biotec, Bergisch Gladbach, Germany)].

CFU-F (colony-forming unit, fibroblast) assay and generation of cultured stromal cells (cSCs)

Unsorted and FACS-sorted primary BM-MNC were cultured at plating densities of 5×10^4 cells/cm^2 cells and 10-50 cells/cm^2, respectively. Colonies (≥40 cells) were counted after 14 days (1% Crystal Violet, Sigma). Assays were set up in duplicates or triplicates. For single-cell CFU-F assays, cells were sorted into 96-well plates, cultured in stromal cell culture medium, and colonies were counted after 3 weeks. Thereafter, cells were harvested and culture-expanded for use in subsequent experiments. Medium was changed weekly and cultured stromal cells were passaged at 80% confluency after trypsinization (0.05% trypsin/EDTA, Invitrogen, Carlsbad, USA).

***In vitro* differentiation assays**

Cultured BM mesenchymal stromal cells (cMSCs) were differentiated towards the adipogenic, osteoblastic, and chondrogenic lineage as described previously (Li et al., 2014). Briefly, cells were cultured for 14 days in AdipoDiff medium (Miltenyi) and cells were stained with Oil Red O (Sigma). For osteogenic differentiation, cells were cultured in osteogenesis induction medium for 21 days and calcium depositions were detected by Alizarin Red staining (Sigma). Osteogenesis induction medium: standard MSC medium supplemented with 0.05 mM L-ascorbic-acid-2-phosphate (Wako Chemicals, Neuss, Germany), 0.1 µM dexamethasone and 10 mM β-glycerophosphate (both from Sigma).

Chondrogenic differentiation was induced by culturing cell pellets ($2.5\text{x}10^5$ cells/pellet) for 28 days in chondrogenesis induction medium. Chondrogenesis induction medium: DMEM-high glucose supplemented with 0.1 µM dexamethasone, 1 mM sodium pyruvate, 0.35 mM L-proline (all from Sigma), 0.17 mM ascorbic acid, 1% ITS+ culture supplements (BD Biosciences) and 0.01 µg/ml TGF-ß3 (R&D Systems). Pellets were paraformaldehyde (PFA)-fixed and frozen in O.C.T. Compound (Sakura, Zoeterwoude, Netherlands). Cryosections were stained with Alcian Bue and nuclei were stained with Nuclear Fast Red (both from

Sigma). Sections were analyzed with a Nikon Eclipse TS100 microscope equipped with a Nikon DS-L3 digital camera.

Antibodies

The following antibodies were used for FACS analysis and cell sorting: CD235a-PE-Cy5 [clone GA-R2 (HIR2)], CD71- PE-Cy5 (clone M-A712), CD45-APC-Cy7 (clone 2D1), CD52-Alexa Fluor (AF) 488 (clone 4C8), CD56 (NCAM1)- Brilliant Violet 605 (clone Leu-19, NKH1), CD81-PE (JS-81) (all BD Bioscience), LEPR-PE (clone 52263, R&D Systems) CD271-APC (clone ME20.4-1.H4, Miltenyi).

Immunofluorescence staining

Human bone marrow paraffin sections (5µm) were deparaffinized and rehydrated following standard protocols. Following antigen retrieval with antigen retrieval solution (DAKO, S1699) and blocking with goat serum (Invitrogen, #31873), slides were stained sequentially scanned, and restained following fluorescence bleaching of primary coupled antibodies in PBS with 50 mM Na_2CO_3 and 3% H_2O_2. An OlympusVS120 slide scanner was used for scanning. Antibodies used were CD45-AlexaFlour 647 (Bio-Rad, MCA87A647T), CD56-AF647 (Biolegend, 318313), mouse anti-human CD271 (R&D Systems, MAB367), and rabbit anti-human CD81 (Novus Biologicals, NBP1-77039) with secondary goat anti-mouse AF647 (Jackson ImmunoResearch, 115-605-166) and goat anti-rabbit AF488 (Jackson ImmunoResearch, 111-545-003) with DAPI (Sigma, D956410MG) as a nuclear counterstain, respectively.

For SPP1 staining, after deparaffinization, rehydration, and antigen retrieval, slides were permeabilized with 0.1% Triton X100 (Sigma Ultra, T9284) solution at 37°C for 15 minutes, blocked using donkey serum (Jackson Immuno Research, 017-000-121), and stained

overnight at 4°C with goat anti-human SPP1 antibody (R&D Systems, AF1433-SP), followed by secondary staining with donkey anti-goat AF594 (Jackson ImmunoResearch, 706-585-147) and staining with CD56-AF647 plus DAPI. Slides were then scanned, bleached, and incubated with rabbit anti-human CD45 (Sigma-Aldrich, HPA000440) and mouse anti-human CD271 overnight, followed by secondary antibody staining with goat anti-mouse AF647 and goat anti-rabbit AF488 and scanned again.

Immunofluorescence image processing

All image data were analyzed and processed using arivis Vision4D. The sequential regions were aligned and overlayed based on nuclei positions. Sequentially used fluorophores were assigned pseudo colors. The obtained 5 color stacks were cropped to an overlapping region and a maximum intensity projection of the three-layer z-stack was generated.

Co-culture of BM CD34^{+} cells with stromal cells

FACS-sorted stromal cells (OC: CD45$^{low/-}$CD235a^{-}CD71^{-}CD271^{+}CD52^{-}NCAM1^{+}; MSSC: CD45$^{low/-}$CD235a^{-}CD71^{-}CD271^{+}CD52^{-}NCAM1^{-}CD81^{++}) were culture expanded and plated as adherent feeder cells into 96-well plates at 10,000 cells per well and cultured in StemMACS MSC Expansion medium for one day. Then, medium was removed and 5,000 BM CD34^{+} cells were added in serum-free expansion medium (SFEM, STEMCELL Technologies) supplemented with or without stem cell factor, thrombopoietin, and FLT-3 ligand (all at 25 ng/ml). Expanded cells were harvested, counted and the expression of CD34 and CD90 were analyzed by flow cytometry after seven days of co-culture.

Single-cell RNA sequencing and data analysis

<u>Single-cell RNA sequencing</u>

Single-cell RNA sequencing (scRNAseq) was performed by the Single-Cell Genomics Platform at the Center for Translational Genomics at Lund University. FACS-sorted BM-MNC populations analyzed were CD45 $^{low/-}$CD235a^{-} cells from 4 donors (2 males, 2 females, ages 19, 22, 52, and 53 years) and CD45 $^{low/-}$CD235a^{-}CD271^{+} from 5 donors (3 males, 2 females, ages 21, 25, 32, 58 and 61 years). Donors younger than 35 years or older than 50 were considered as young donors and old donors, respectively. FACS-sorted BM-MNC cells were encapsulated into emulsion droplets as single cells using the Chromium Controller (10X Genomics). scRNAseq libraries were constructed using the Chromium Single Cell Gene Expression 3' v3 Reagent Kit according to the manufacturer's instructions. Reverse transcription and library preparation was performed on a C1000 Touch Thermal Cycler (Bio-Rad). Amplified cDNA and final libraries concentrations were measured with a Qubit 4 Fluorometer (Invitrogen) using the dsDNA High Sensitivity Assay (Invitrogen). cDNA and library traces were evaluated on a TapeStation (Agilent) using High Sensitivity D5000 and D1000 Screen Tapes (Agilent). Individual libraries were diluted to 1.5 nM and pooled for sequencing. Pools were sequenced on a NovaSeq 6000 (Illumina) aiming for a sequencing depth of 50.000 reads per cell.

Pre-processing of scRNA-seq data, cell filtering and normalization

scRNAseq data of both CD45 $^{low/-}$CD235a^{-} and CD45 $^{low/-}$CD235a^{-}CD271^{+} populations were processed and combined for further data analysis. Sequencing data were demultiplexed and UMI-collapsed using the CellRanger toolkit (version 3.1, 10X Genomics). Reads were mapped against the human GRCh38 genome paired with the gencode.v31 gene level data using the kallisto/bustools package (version 0.46.0 resp. 0.39.3). Mitochondrial and ribosomal reads were excluded and cells with less than 1000

(diploid) detected UMIs were removed from the analysis. Expression was normalized using the scanpy.pp.downsample_counts function to 1000 UMIs.

Dimensionality reduction

Dimensionality reduction was performed using gene expression data for a subset of 3000 variable genes with the scanpy.pp.highly_variable_genes function. UMAP was generated with the default setting of Scanpy version 1.6.0 using the top 50 principal components (PC).

Clustering and visualization

Graph-based clustering of the PCA reduced data was performed using the Louvain method implemented in scanpy. The resolution parameter was set to 11, which resulted in 103 preliminary clusters that were merged based on the correlation of mean expression for each cluster. Clusters with correlations greater than 0.95 were merged, resulting in the final clusters displayed in Figure 1B. Cell type annotation was performed based on the published marker genes expression (Table S1) (Baryawno et al., 2019).

Differential expression of gene signature

For each cluster, we used the scanpy.tl.rank_genes_groups function with default values to identify genes that had a significant fold change of higher than 2 in comparison to the other clusters.

RNA velocity analysis

Bone marrow stromal cell dynamics were analyzed using the scVelo package (version 0.2.2) in Scanpy 1.6.0 (Bergen et al., 2020a; Svensson and Pachter, 2018). The data were processed using default parameters following preprocessing as described in scVelo package. Analysis of

cellular trajectory inference by RNA velocity was performed using dynamical modeling, which is a generalization of the original RNA velocity method and which allows for characterization of multiple transcriptional states. Latent time analysis, pseudotime analysis and velocity confidence were computed using default parameters. Extrapolated states were then projected on the UMAP embedding produced during the initial analysis step. The clustering and visualization were repeated using the same parameters as above for the selected stromal clusters forming the continuum.

For velocity analysis, the samples were pre-processed using functions for detection of minimum number of counts, filtering and normalization using scv.pp.filter_and_normalise and followed by scv.pp.moments function. The gene specific velocities were then calculated using scv.tl.velocity with mode set to stochastic, and visualized using scv.pl.velocity_embedding function. In addition, we used scv.tl.latent_time function to infer a shared latent time from splicing dynamics and plotted the genes along a time axis sorted by expression along dynamics using scv.pl.heatmap function. Pseudotime trajectory and velocity confidence were computed using velocity_pseudotime, and scv.tl.velocity_confidence functions.

Cell-cell communication analysis

Cell-cell interaction analysis was conducted with CellPhoneDB2 (Efremova et al., 2020), which is a Python-based analytical tool for calculating the interaction between ligands and receptors between different cell populations. The mean and p value of each ligand-receptor interaction between different clusters were calculated by CellPhoneDB2. To illustrate the strength of specific pathways between different clusters, several common pathways with high

mean values were selected for visualization. Dot plots were generated with R studio (Version 1.3.1093).

Data availability

The scRNA-seq matrix data generated in this study have been deposited in the GEO database (GSE190965).

Acknowledgements

This work was supported by funds from the StemTherapy Program, the Swedish Cancer Foundation, the Swedish Childhood Cancer Foundation, the Swedish Bloodcancer Association (Blodcancerförbundet), Foundation Siv-Inger and Per-Erik Anderssons minnesfond, John Persson Foundation, ALF (Government Public Health Grant), and the Skåne County Council Research Foundation. The authors would like to thank Helene Larsson and Anna Jonasson for their help to collect bone marrow samples and the Lund Stem Cell Center FACS facility personnel for technical assistance. We would also like to thank Dr. Paul Bourgine for critically reading the manuscript.

Author contributions

H.L. designed and performed experiments, analyzed data and wrote the manuscript. SB designed and performed in-situ staining experiments, analyzed data, and co-wrote corresponding parts of the manuscript. SL analyzed the scRAN-seq data and co-wrote corresponding parts of the manuscript. S.S. supervised the study, designed the experiments, and wrote the manuscript.

Declaration of interest

The authors have no conflicts of interest to disclose.

References

Ambrosi, T.H., Scialdone, A., Graja, A., Gohlke, S., Jank, A.M., Bocian, C., Woelk, L., Fan, H., Logan, D.W., Schurmann, A.*, et al.* (2017). Adipocyte Accumulation in the Bone Marrow during Obesity and Aging Impairs Stem Cell-Based Hematopoietic and Bone Regeneration. Cell Stem Cell *20*, 771-784 e776.

Baccin, C., Al-Sabah, J., Velten, L., Helbling, P.M., Grunschlager, F., Hernandez-Malmierca, P., Nombela-Arrieta, C., Steinmetz, L.M., Trumpp, A., and Haas, S. (2020). Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. Nat Cell Biol *22*, 38-48.

Baksh, D., Davies, J.E., and Zandstra, P.W. (2005). Soluble factor cross-talk between human bone marrow-derived hematopoietic and mesenchymal cells enhances in vitro CFU-F and CFU-O growth and reveals heterogeneity in the mesenchymal progenitor cell compartment. Blood *106*, 3012-3019.

Baryawno, N., Przybylski, D., Kowalczyk, M.S., Kfoury, Y., Severe, N., Gustafsson, K., Kokkaliaris, K.D., Mercier, F., Tabaka, M., Hofree, M.*, et al.* (2019). A Cellular Taxonomy of the Bone Marrow Stroma in Homeostasis and Leukemia. Cell *177*, 1915-1932 e1916.

Bergen, V., Lange, M., Peidli, S., Wolf, F.A., and Theis, F.J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol *38*, 1408-1414.

Boulais, P.E., Mizoguchi, T., Zimmerman, S., Nakahara, F., Vivie, J., Mar, J.C., van Oudenaarden, A., and Frenette, P.S. (2018). The Majority of CD45(-) Ter119(-) CD31(-) Bone Marrow Cell Fraction Is of Hematopoietic Origin and Contains Erythroid and Lymphoid Progenitors. Immunity *49*, 627-639 e626.

Cao, Z., Umek, R.M., and McKnight, S.L. (1991). Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. Genes Dev *5*, 1538-1552.

Casanova-Acebes, M., Pitaval, C., Weiss, L.A., Nombela-Arrieta, C., Chevre, R., N, A.G., Kunisaki, Y., Zhang, D., van Rooijen, N., Silberstein, L.E.*, et al.* (2013). Rhythmic modulation of the hematopoietic niche through neutrophil clearance. Cell *153*, 1025-1035.

Chan, C.K.F., Gulati, G.S., Sinha, R., Tompkins, J.V., Lopez, M., Carter, A.C., Ransom, R.C., Reinisch, A., Wearda, T., Murphy, M.*, et al.* (2018). Identification of the Human Skeletal Stem Cell. Cell *175*, 43-56 e21.

Churchman, S.M., Ponchel, F., Boxall, S.A., Cuthbert, R., Kouroupis, D., Roshdy, T., Giannoudis, P.V., Emery, P., McGonagle, D., and Jones, E.A. (2012). Transcriptional profile of native CD271+ multipotential stromal cells: evidence for multiple fates, with prominent osteogenic and Wnt pathway signaling activity. Arthritis Rheum *64*, 2632-2643.

de Jong, M.M.E., Kellermayer, Z., Papazian, N., Tahri, S., Hofste Op Bruinink, D., Hoogenboezem, R., Sanders, M.A., van de Woestijne, P.C., Bos, P.K., Khandanpour, C.*, et al.* (2021). The multiple myeloma microenvironment is defined by an inflammatory stromal cell landscape. Nat Immunol *22*, 769-780.

Delorme, B., Ringe, J., Gallay, N., Le Vern, Y., Kerboeuf, D., Jorgensen, C., Rosset, P., Sensebe, L., Layrolle, P., Haupl, T.*, et al.* (2008). Specific plasma membrane protein phenotype of

culture-amplified and native human bone marrow mesenchymal stem cells. Blood *111*, 2631-2635.
Ding, L., Saunders, T.L., Enikolopov, G., and Morrison, S.J. (2012). Endothelial and perivascular cells maintain haematopoietic stem cells. Nature *481*, 457-462.
Dupard, S.J., Grigoryan, A., Farhat, S., Coutu, D.L., and Bourgine, P.E. (2020). Development of Humanized Ossicles: Bridging the Hematopoietic Gap. Trends Mol Med *26*, 552-569.
Efremova, M., Vento-Tormo, M., Teichmann, S.A., and Vento-Tormo, R. (2020). CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc *15*, 1484-1506.
Gao, S., Shi, Q., Zhang, Y., Liang, G., Kang, Z., Huang, B., Ma, D., Wang, L., Jiao, J., Fang, X.*, et al.* (2021). Identification of HSC/MPP expansion units in fetal liver by single-cell spatiotemporal transcriptomics. Cell Res.
Gronthos, S., Fitter, S., Diamond, P., Simmons, P.J., Itescu, S., and Zannettino, A.C. (2007). A novel monoclonal antibody (STRO-3) identifies an isoform of tissue nonspecific alkaline phosphatase expressed by multipotent bone marrow stromal stem cells. Stem Cells Dev *16*, 953-963.
Headland, S.E., Jones, H.R., Norling, L.V., Kim, A., Souza, P.R., Corsiero, E., Gil, C.D., Nerviani, A., Dell'Accio, F., Pitzalis, C.*, et al.* (2015). Neutrophil-derived microvesicles enter cartilage and protect the joint in inflammatory arthritis. Sci Transl Med *7*, 315ra190.
Jones, E.A., Kinsey, S.E., English, A., Jones, R.A., Straszynski, L., Meredith, D.M., Markham, A.F., Jack, A., Emery, P., and McGonagle, D. (2002). Isolation and characterization of bone marrow multipotential mesenchymal progenitor cells. Arthritis Rheum *46*, 3349-3360.
Kleven, M.D., Jue, S., and Enns, C.A. (2018). Transferrin Receptors TfR1 and TfR2 Bind Transferrin through Differing Mechanisms. Biochemistry *57*, 1552-1559.
Li, H., Ghazanfari, R., Zacharaki, D., Ditzel, N., Isern, J., Ekblom, M., Mendez-Ferrer, S., Kassem, M., and Scheding, S. (2014). Low/negative expression of PDGFR-alpha identifies the candidate primary mesenchymal stromal cells in adult human bone marrow. Stem Cell Reports *3*, 965-974.
Li, H., Ghazanfari, R., Zacharaki, D., Lim, H.C., and Scheding, S. (2016). Isolation and characterization of primary bone marrow mesenchymal stromal cells. Ann N Y Acad Sci *1370*, 109-118.
Li, H., Lim, H.C., Zacharaki, D., Xian, X., Kenswil, K.J.G., Braunig, S., Raaijmakers, M., Woods, N.B., Hansson, J., and Scheding, S. (2020). Early growth response 1 regulates hematopoietic support and proliferation in human primary bone marrow stromal cells. Haematologica *105*, 1206-1215.
Mendez-Ferrer, S., Lucas, D., Battista, M., and Frenette, P.S. (2008). Haematopoietic stem cell release is regulated by circadian oscillations. Nature *452*, 442-447.
Morrison, S.J., and Scadden, D.T. (2014). The bone marrow niche for haematopoietic stem cells. Nature *505*, 327-334.
Nakamura, M., Sone, S., Takahashi, I., Mizoguchi, I., Echigo, S., and Sasano, Y. (2005). Expression of versican and ADAMTS1, 4, and 5 during bone development in the rat mandible and hind limb. J Histochem Cytochem *53*, 1553-1562.
Pinho, S., and Frenette, P.S. (2019). Haematopoietic stem cell activity and interactions with the niche. Nat Rev Mol Cell Biol *20*, 303-320.
Pinho, S., Lacombe, J., Hanoun, M., Mizoguchi, T., Bruns, I., Kunisaki, Y., and Frenette, P.S. (2013). PDGFRalpha and CD51 mark human nestin+ sphere-forming mesenchymal stem cells capable of hematopoietic progenitor cell expansion. J Exp Med *210*, 1351-1367.

Pittenger, M.F., Mackay, A.M., Beck, S.C., Jaiswal, R.K., Douglas, R., Mosca, J.D., Moorman, M.A., Simonetti, D.W., Craig, S., and Marshak, D.R. (1999). Multilineage potential of adult human mesenchymal stem cells. Science *284*, 143-147.
Quirici, N., Soligo, D., Bossolasco, P., Servida, F., Lumini, C., and Deliliers, G.L. (2002). Isolation of bone marrow mesenchymal stem cells by anti-nerve growth factor receptor antibodies. ExpHematol *30*, 783-791.
Sacchetti, B., Funari, A., Michienzi, S., Di Cesare, S., Piersanti, S., Saggio, I., Tagliafico, E., Ferrari, S., Robey, P.G., Riminucci, M.*, et al.* (2007). Self-renewing osteoprogenitors in bone marrow sinusoids can organize a hematopoietic microenvironment. Cell *131*, 324-336.
Tikhonova, A.N., Dolgalev, I., Hu, H., Sivaraj, K.K., Hoxha, E., Cuesta-Dominguez, A., Pinho, S., Akhmetzyanova, I., Gao, J., Witkowski, M.*, et al.* (2019). The bone marrow microenvironment at single-cell resolution. Nature *569*, 222-228.
Tormin, A., Li, O., Brune, J.C., Walsh, S., Schutz, B., Ehinger, M., Ditzel, N., Kassem, M., and Scheding, S. (2011). CD146 expression on primary nonhematopoietic bone marrow stem cells is correlated with in situ localization. Blood *117*, 5067-5077.
Triana, S., Vonficht, D., Jopp-Saile, L., Raffel, S., Lutz, R., Leonce, D., Antes, M., Hernandez-Malmierca, P., Ordonez-Rueda, D., Ramasz, B.*, et al.* (2021). Single-cell proteo-genomic reference maps of the hematopoietic system enable the purification and massive profiling of precisely defined cell states. Nat Immunol *22*, 1577-1589.
Urbieta, M., Barao, I., Jones, M., Jurecic, R., Panoskaltsis-Mortari, A., Blazar, B.R., Murphy, W.J., and Levy, R.B. (2010). Hematopoietic progenitor cell regulation by CD4+CD25+ T cells. Blood *115*, 4934-4943.
Wang, Z., Li, X., Yang, J., Gong, Y., Zhang, H., Qiu, X., Liu, Y., Zhou, C., Chen, Y., Greenbaum, J.*, et al.* (2021). Single-cell RNA sequencing deconvolutes the in vivo heterogeneity of human bone marrow-derived mesenchymal stem cells. Int J Biol Sci *17*, 4192-4206.
Wolock, S.L., Krishnan, I., Tenen, D.E., Matkins, V., Camacho, V., Patel, S., Agarwal, P., Bhatia, R., Tenen, D.G., Klein, A.M.*, et al.* (2019). Mapping Distinct Bone Marrow Niche Populations and Their Differentiation Paths. Cell Rep *28*, 302-311 e305.
Yamazaki, S., Ema, H., Karlsson, G., Yamaguchi, T., Miyoshi, H., Shioda, S., Taketo, M.M., Karlsson, S., Iwama, A., and Nakauchi, H. (2011). Nonmyelinating Schwann cells maintain hematopoietic stem cell hibernation in the bone marrow niche. Cell *147*, 1146-1158.
Zhao, M., Perry, J.M., Marshall, H., Venkatraman, A., Qian, P., He, X.C., Ahamed, J., and Li, L. (2014). Megakaryocytes maintain homeostatic quiescence and promote post-injury regeneration of hematopoietic stem cells. Nat Med *20*, 1321-1326.
Zhou, B.O., Yue, R., Murphy, M.M., Peyer, J.G., and Morrison, S.J. (2014). Leptin-receptor-expressing mesenchymal stromal cells represent the main source of bone formed by adult bone marrow. Cell Stem Cell *15*, 154-168.

Figure legends

Figure 1. Single-cell atlas of human bone marrow CD45$^{low/-}$CD235a^{-} cells. (A) Overview of the study design including gating strategies for isolation of human bone marrow CD45$^{low/-}$CD235a^{-} and CD45$^{low/-}$CD235a^{-}CD271^{+} cells. (B) Uniform Manifold Approximation and Projection (UMAP) display of single-cell transcriptomic data of human bone marrow CD45$^{low/-}$CD235a^{-} cells containing enriched CD45$^{low/-}$CD235a^{-}CD271^{+} cells to allow for a detailed analysis of rare stromal cell subpopulations. Data are shown for a total of n=25067 cells (5704 CD45$^{low/-}$CD235a^{-} cells and 19363 CD45$^{low/-}$CD235a^{-}CD271^{+} cells) from a total of nine healthy donors. Color legend indicate cluster numbers and annotations. Basal, basal cluster; B prog., B cell progenitors; SC, stromal cells; PC, plasma cells; HSPC, hematopoietic stem and progenitor cells; Mk, megakaryocytes; DC, dendritic cells; GC, granulocytes; Ery., erythroid cells; NK, natural killer cells; EC, endothelial cells; Neuronal, neuronal cluster. (C) UMAP display of non-hematopoietic clusters (n=9686 cells). The circled area in the overview (top left corner) indicates the non-hematopoietic clusters selected from (B). (D) Heatmap of representative differentially expressed genes for each of the non-hematopoietic clusters in (C). For stromal marker identification, the top 100 significant differentially expressed genes from each stromal cluster were selected to identify overlapping genes. Three genes shared by all nine clusters were identified in this comparison (NNMT, IFITM3, and DCN). Nine more genes were identified when OC clusters 23 and 29 (identified as OCs in Fig. 2A) were removed from the comparison, as OCs showed a considerably different gene expression profiles as compared with other stromal clusters. Cluster numbers and corresponding cell types are indicated under the heatmap. The scale bar indicates gene expression levels. EC, endothelial cells. A blow-up heatmap for clusters 38, 29, 23, 6, and 37 is shown under the main heatmap for better visualization.

Figure 2. scRNAseq reveals distinct gene expression patterns in different bone marrow stromal cell populations. (A) UMAP display of the nine stromal subsets (n=7069 cells). Color legends indicate stromal cluster numbers and annotations. MSSC, multipotent stromal stem cells; OC, osteochondrogenic progenitors. (B) Single cell heatmap of representative differentially expressed genes in each cluster shown in (A). A blow-up heatmap for clusters 38, 29, 23, 6, and 37 is shown in Figure S2A for better visualization. (C-F) Stacked violin plots of adipogenic- (C), osteogenic- (D), osteochondrogenic- (E), and fibroblastic (F) markers in different stromal clusters. Dashed lines separate MSSCs, adipo-primed progenitors, balanced progenitors, and pre-osteoblast (left; cluster 3, 5, 16 and 38), OCs (middle; cluster 29, 23), and pre-fibroblasts (right; cluster 6, 37, 8).

Figure 3. RNA velocity analysis reconstructs the temporal sequence of transcriptomic events of stromal cells. (A and A') Single cell velocities of the nine stromal clusters visualized as streamlines in a UMAP. Black arrows indicate direction and thickness indicates speed along the stromal cell development trajectory. (A') Colored thick arrows indicate the main directions of stromal cell developmental paths. Colors correspond with the differentiation destination. Red: adipo-primed progenitors; yellow: balanced progenitors; light green: OCs; dark green: pre-osteoblasts; blue: pre-fibroblasts. (B) UMAP display of stromal cells colored by inferred latent time. Inferred latent time is represented by a color scale from 0 (the earliest latent time) to 1 (the latest latent time). (C) Left: UMAP display of stromal cell clusters that form a continuum of different cellular states (n=6376 cells). Colors correspond to different stromal clusters as in Figure 2A. Right: Heatmap constructed by the top 300 likelihood-ranked genes demonstrates gene expression dynamics along latent time. Colors on top of the heatmap correspond with cluster colors in UMAP (left). Key genes are

highlighted by different colors on the right. Gene name colors correspond to different developmental stage transitions. Red: genes responsible for MSSC intra-cluster transition; orange: genes responsible for the transition to adipo-primed progenitors; light green: genes responsible for the transition to OCs; dark green: genes responsible for the transition to balanced progenitors and pre-osteoblasts.

Figure 4. Stromal cell isolation based on data-driven gating strategies distinguishes cell subsets with different colony-forming capacities and differentiation capacities. (A) Dot plot of surface marker gene expression in different stromal cell clusters. Cluster numbers and corresponding stromal cell groups are indicated on the y-axis legend. Dot sizes represent the percentage of cells expressing a certain gene in each cluster and dot colors represent the scaled average expression of that gene. (B) FACS plots illustrating the gating strategy for the isolation of different stromal subsets. The displayed cell populations are indicated on top of the plot. Following exclusion of doublets, dead cells, CD45- and CD235a-expressing cells, $CD45^{low/-}CD235a^{-}CD71^{-}CD271^{+}$ cells (left panel) were gated based on CD52 and NCAM1 expression (middle panel). The resulting three populations were labelled as A-C (middle panel), corresponding to the stromal cell groups in (A) (A, $CD52^{-}NCAM1^{-}$; B, $CD52^{-}NCAM1^{+}$; C, $CD52^{+}NCAM1^{-}$). $CD45^{-}CD235a^{-}CD71^{-}CD271^{+}CD52^{-}NCAM1^{-}$ (group A) cells were further divided based on CD81 expression and four populations were identified (A1-A4) (right panel). A1, $CD81^{++}$; A2, $CD81^{+}$; A3; $CD81^{+/-}$; A4, $CD81^{-}$. (C) CFU-F frequencies of sorted stromal cell populations as shown in (B). Data are presented as individual data (dots) and median (horizontal lines) from independent experiments (n=3-6). Symbol colors and x-axis labels correspond to the cell population colors in (B). *: $p<0.05$; ****: $p <0.0001$ (Kruskal-Wallis test). (D) In vitro differentiation capacity of sorted stromal cell populations (as indicated in B) towards the adipogenic, osteoblastic, and chondrogenic lineage. Non-

induction controls are shown in the left panel. Scale bars represent 200 µm. A representative set of pictures from a total of three independent experiments is shown. (F) Formalin-fixed, paraffin-embedded (FFPE) human BM slides were sequentially stained for DAPI (blue), CD45 (yellow), CD81 (green), CD271 (pink), and NCAM1 (cyan) and scanned with the OlympusVS120 slide scanner. Different staining combinations are shown as indicated under each picture to provide better visualization of individual staining obtained from the same FFPE slide. Scale bars represent 50 µm. Red arrows: $CD271^{+}CD81^{++}$ cells; white arrows: arteriolar walls; white lines: bone lining regions. Bone (b), adipocytes (a), and capillaries (*) are indicated.

Figure 5. Cell-cell interaction between stromal cells and hematopoietic cells in human bone marrow. (A) Circos plot visualization of cell-cell interaction between bone marrow stromal cells and different hematopoietic cell types. Colored lines connect stromal clusters with hematopoietic clusters based on LR expression analysis. Line colors indicate cell type and line thicknesses correspond to the number of interacting LR pairs. Cluster numbers and corresponding cell types and stromal groups are indicated. Hematopoietic cell type abbreviations are listed in the table. (B) Overview of selected ligand-receptor pairs between MSSCs and OCs, respectively, and different hematopoietic cell types. The y-axis label indicates the pair of interacting cell clusters (‘cluster X_cluster Y’ indicates cluster X interaction with cluster Y by ligand-receptor expression). The x-axis indicates the interacting receptor/ligand (R/L) or ligand/receptor pairs, respectively (‘molecule L_molecule R’, molecule L interacts with molecule R). The means of the average expression level of the interacting molecules are indicated by circle size (adjacent lower scale bar). P values are indicated by circle color and correspond to the upper scale bar. Cluster numbers of the

hematopoietic cell types are listed in the table. The dashed line separates MSSC from OC interactions.

Figure 6. Expression of selected ligand and receptor pairs involved in cell-cell interaction between stromal cells and hematopoietic cells in human bone marrow. (A-D, F-G) UMAP (as in Fig. 1B) illustration of the normalized expression of selected ligand and receptor gene pairs. Red dashed lines in figures A, B and D mark the erythroid clusters. The blow-up shown in (G) is to better visualize PDGFA- and PDGFB-expressing cells. (E) Formalin-fixed, paraffin-embedded (FFPE) human BM slides were sequentially stained for DAPI (blue), CD45 (yellow), SPP1 (red), CD271 (pink), and NCAM1 (cyan) and scanned with an OlympusVS120 slide scanner. Lower panel: Single staining for each marker is shown as indicated under each image. Scale bars represent 50 μm. White lines indicate bone lining regions. b, Bone.

Figure 1

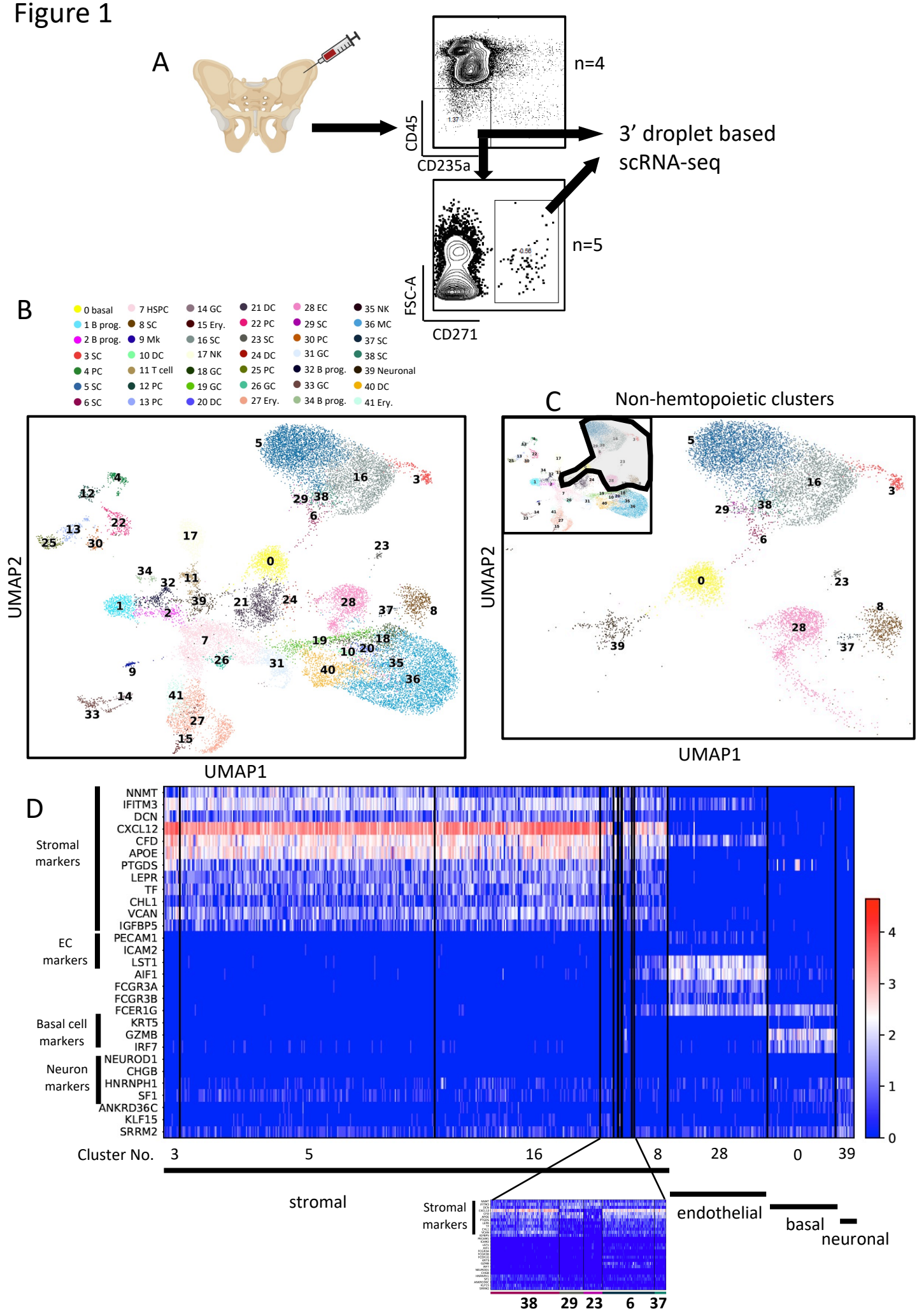

A
n=4
CD45
CD235a
3' droplet based
scRNA-seq
n=5
FSC-A
CD271
B
0 basal
1 B prog.
2 B prog.
3 SC
4 PC
5 SC
6 SC
7 HSPC
8 SC
9 Mk
10 DC
11 T cell
12 PC
13 PC
14 GC
15 Ery.
16 SC
17 NK
18 GC
19 GC
20 DC
21 DC
22 PC
23 SC
24 DC
25 PC
26 GC
27 Ery.
28 EC
29 SC
30 PC
31 GC
32 B prog.
33 GC
34 B prog.
35 NK
36 MC
37 SC
38 SC
39 Neuronal
40 DC
41 Ery.
UMAP2
UMAP1
C
Non-hemtopoietic clusters
UMAP2
UMAP1
D
Stromal markers
NNMT
IFITM3
DCN
CXCL12
CFD
APOE
PTGDS
LEPR
TF
CHL1
VCAN
IGFBP5
EC markers
PECAM1
ICAM2
LST1
AIF1
FCGR3A
FCGR3B
FCER1G
Basal cell markers
KRT5
GZMB
IRF7
Neuron markers
NEUROD1
CHGB
HNRNPH1
SF1
ANKRD36C
KLF15
SRRM2
Cluster No.
3
5
16
8
28
0
39
stromal
endothelial
basal
neuronal
Stromal markers
38
29
23
6
37

Figure 2

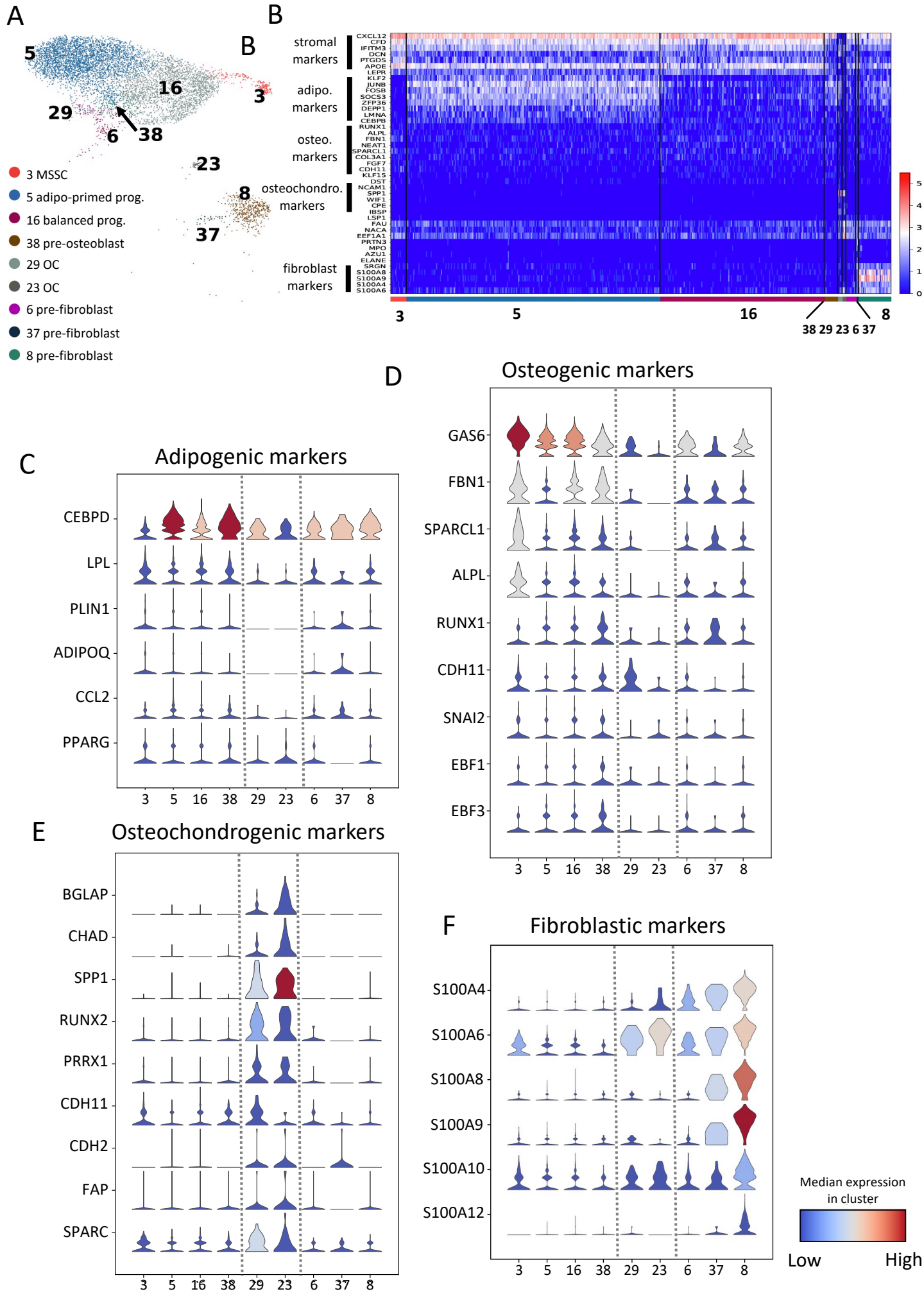

A
5
16
3
29
6
38
23
8
37
3 MSSC
5 adipo-primed prog.
16 balanced prog.
38 pre-osteoblast
29 OC
23 OC
6 pre-fibroblast
37 pre-fibroblast
8 pre-fibroblast
B
stromal markers
adipo. markers
osteo. markers
osteochondro. markers
fibroblast markers
CXCL12
CFD
IFITM3
DCN
PTGDS
APOE
LEPR
KLF2
JUNB
FOSB
SOCS3
ZFP36
DEPP1
LMNA
CEBPB
RUNX1
ALPL
FBN1
NEAT1
SPARCL1
COL3A1
FGF7
CDH11
KLF15
DST
NCAM1
SPP1
WIF1
CPE
IBSP
LSP1
FAU
NACA
EEF1A1
PRTN3
MPO
AZU1
ELANE
SRGN
S100A8
S100A9
S100A4
S100A6
3
5
16
38 29 23 6 37
8
C
Adipogenic markers
CEBPD
LPL
PLIN1
ADIPOQ
CCL2
PPARG
3 5 16 38 29 23 6 37 8
D
Osteogenic markers
GAS6
FBN1
SPARCL1
ALPL
RUNX1
CDH11
SNAI2
EBF1
EBF3
3 5 16 38 29 23 6 37 8
E
Osteochondrogenic markers
BGLAP
CHAD
SPP1
RUNX2
PRRX1
CDH11
CDH2
FAP
SPARC
3 5 16 38 29 23 6 37 8
F
Fibroblastic markers
S100A4
S100A6
S100A8
S100A9
S100A10
S100A12
3 5 16 38 29 23 6 37 8
Median expression in cluster
Low
High

Figure 3

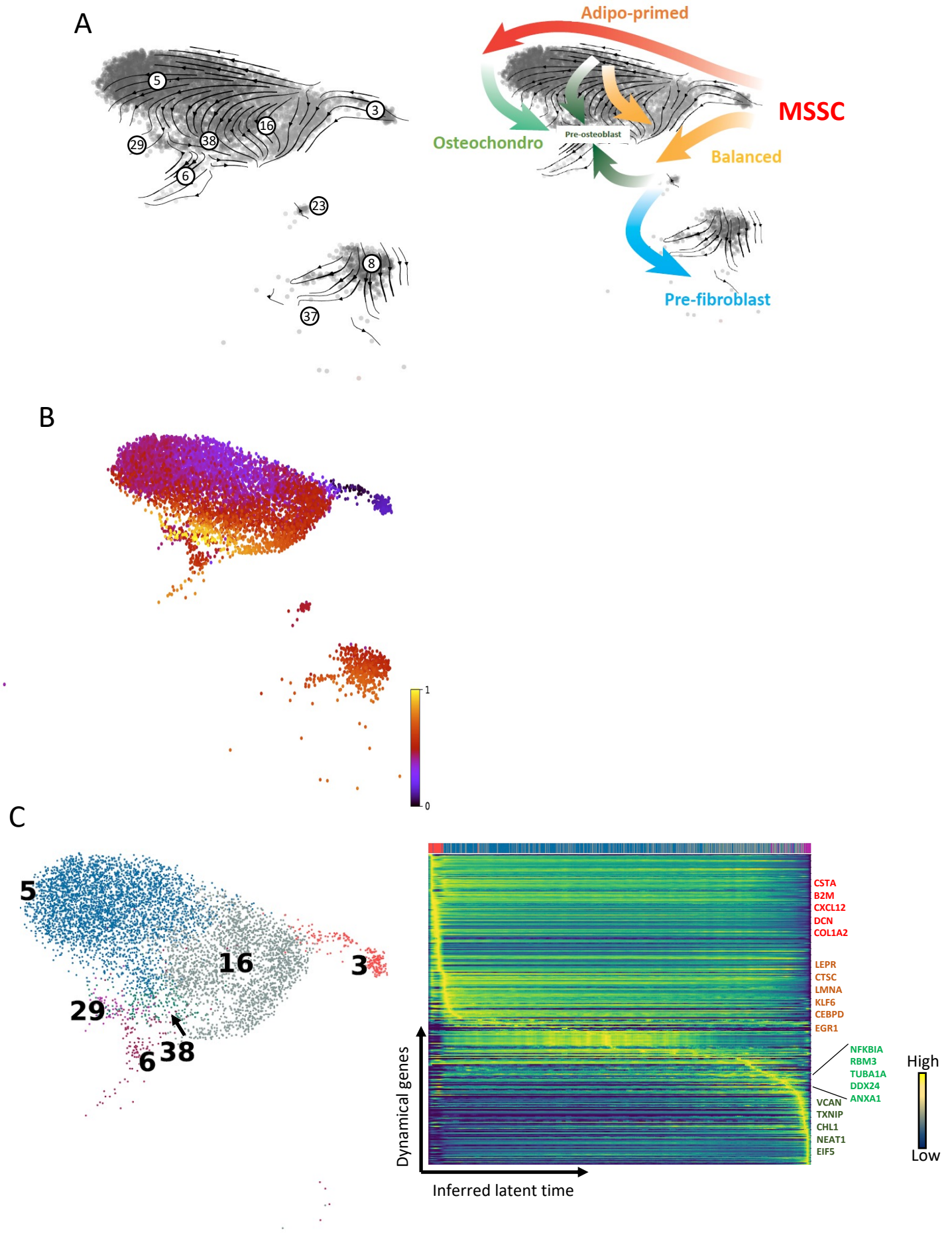

A
5
3
16
29
38
6
23
8
37
Adipo-primed
MSSC
Osteochondro
Pre-osteoblast
Balanced
Pre-fibroblast
B
1
0
C
5
16
3
29
38
6
Dynamical genes
Inferred latent time
CSTA
B2M
CXCL12
DCN
COL1A2
LEPR
CTSC
LMNA
KLF6
CEBPD
EGR1
NFKBIA
RBM3
TUBA1A
DDX24
ANXA1
VCAN
TXNIP
CHL1
NEAT1
EIF5
High
Low

Figure 4

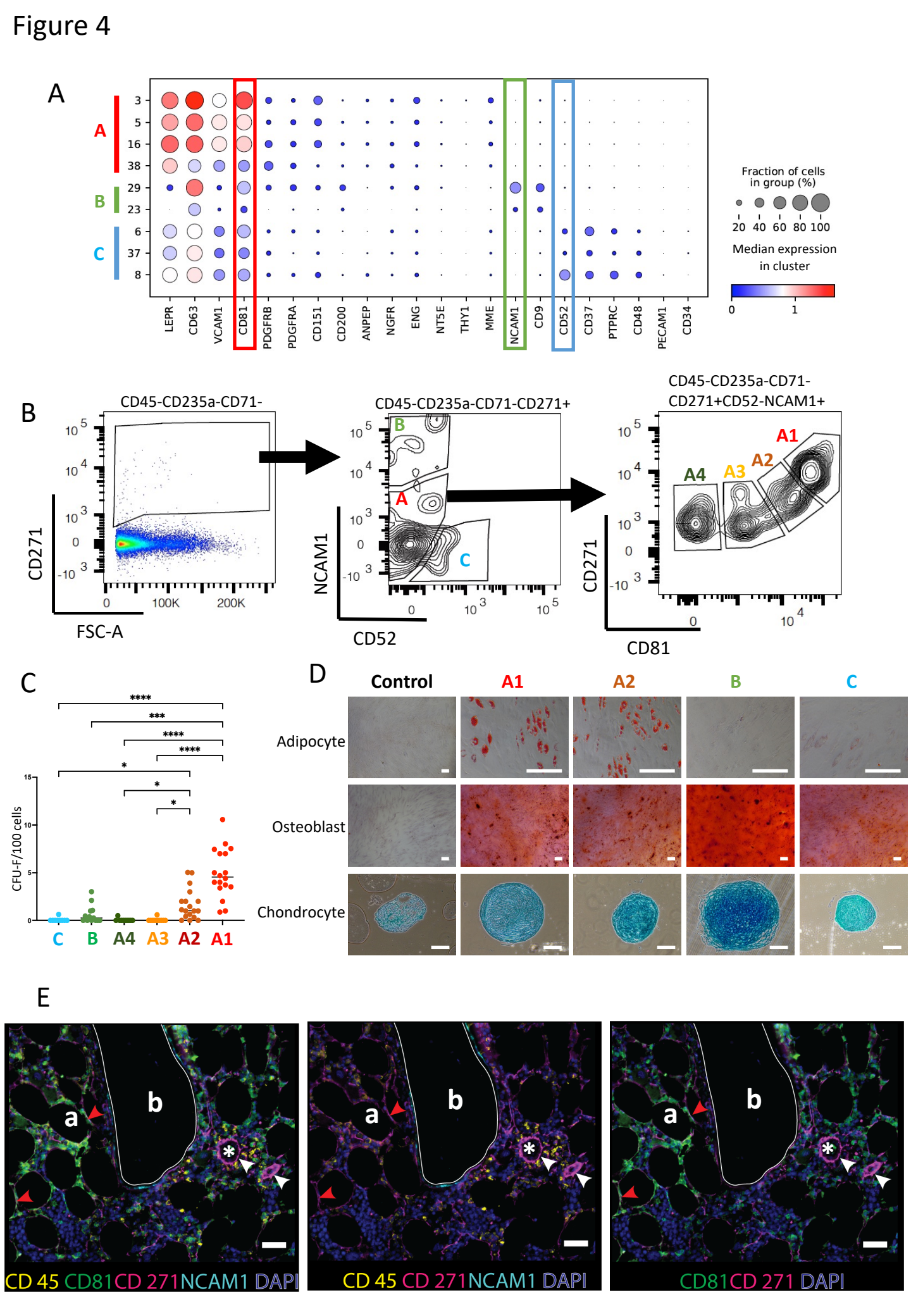
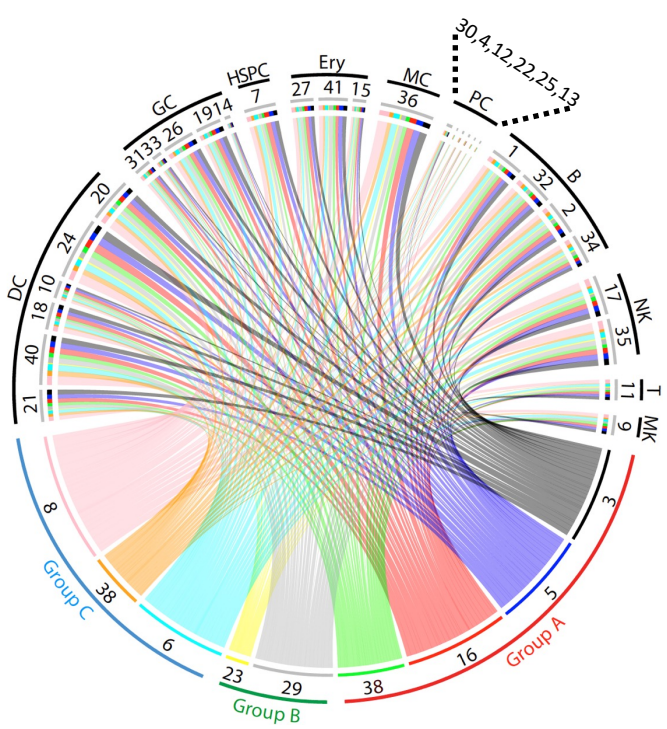

A
A
B
C
3
5
16
38
29
23
6
37
8
LEPR
CD63
VCAM1
CD81
PDGFRB
PDGFRA
CD151
CD200
ANPEP
NGFR
ENG
NT5E
THY1
MME
NCAM1
CD9
CD52
CD37
PTPRC
CD48
PECAM1
CD34
Fraction of cells in group (%)
20 40 60 80 100
Median expression in cluster
0
1
B
CD45-CD235a-CD71-
CD271
FSC-A
CD45-CD235a-CD71-CD271+
NCAM1
CD52
CD45-CD235a-CD71-CD271+CD52-NCAM1+
A4
A3
A2
A1
CD271
CD81
C
CFU-F/100 cells
C
B
A4
A3
A2
A1
D
Control
A1
A2
B
C
Adipocyte
Osteoblast
Chondrocyte
E
a
b
CD 45 CD81 CD 271 NCAM1 DAPI
CD 45 CD 271 NCAM1 DAPI
CD81 CD 271 DAPI

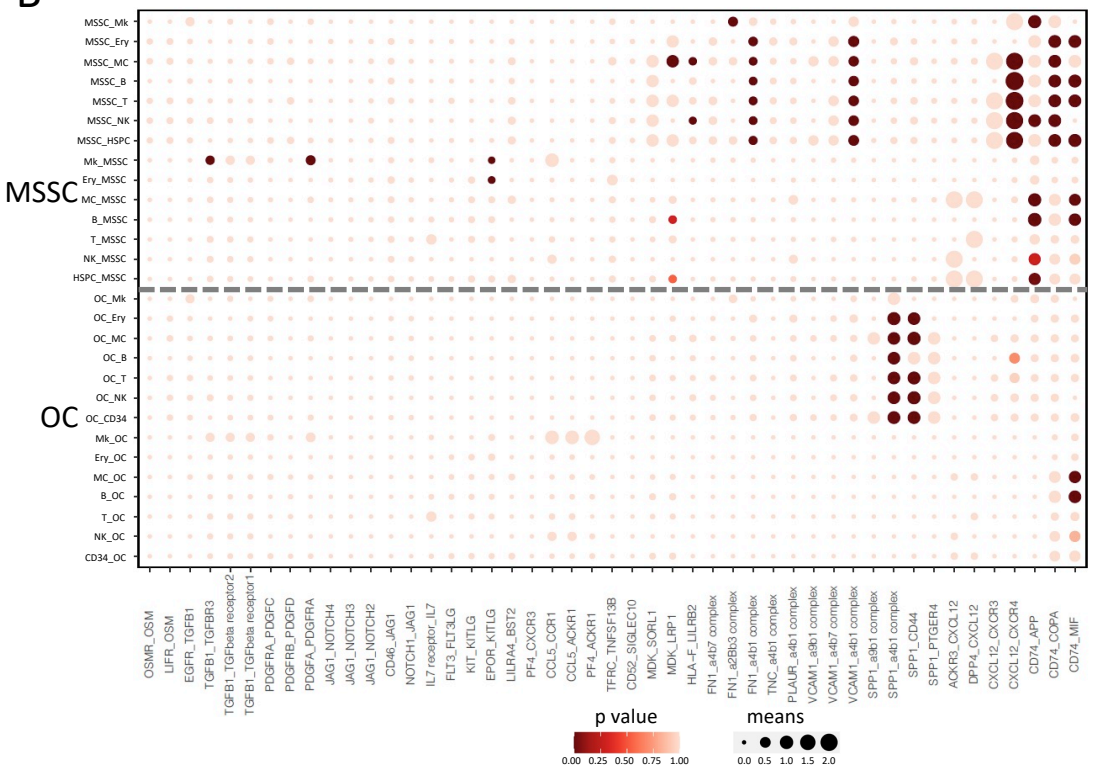
Figure 5

A



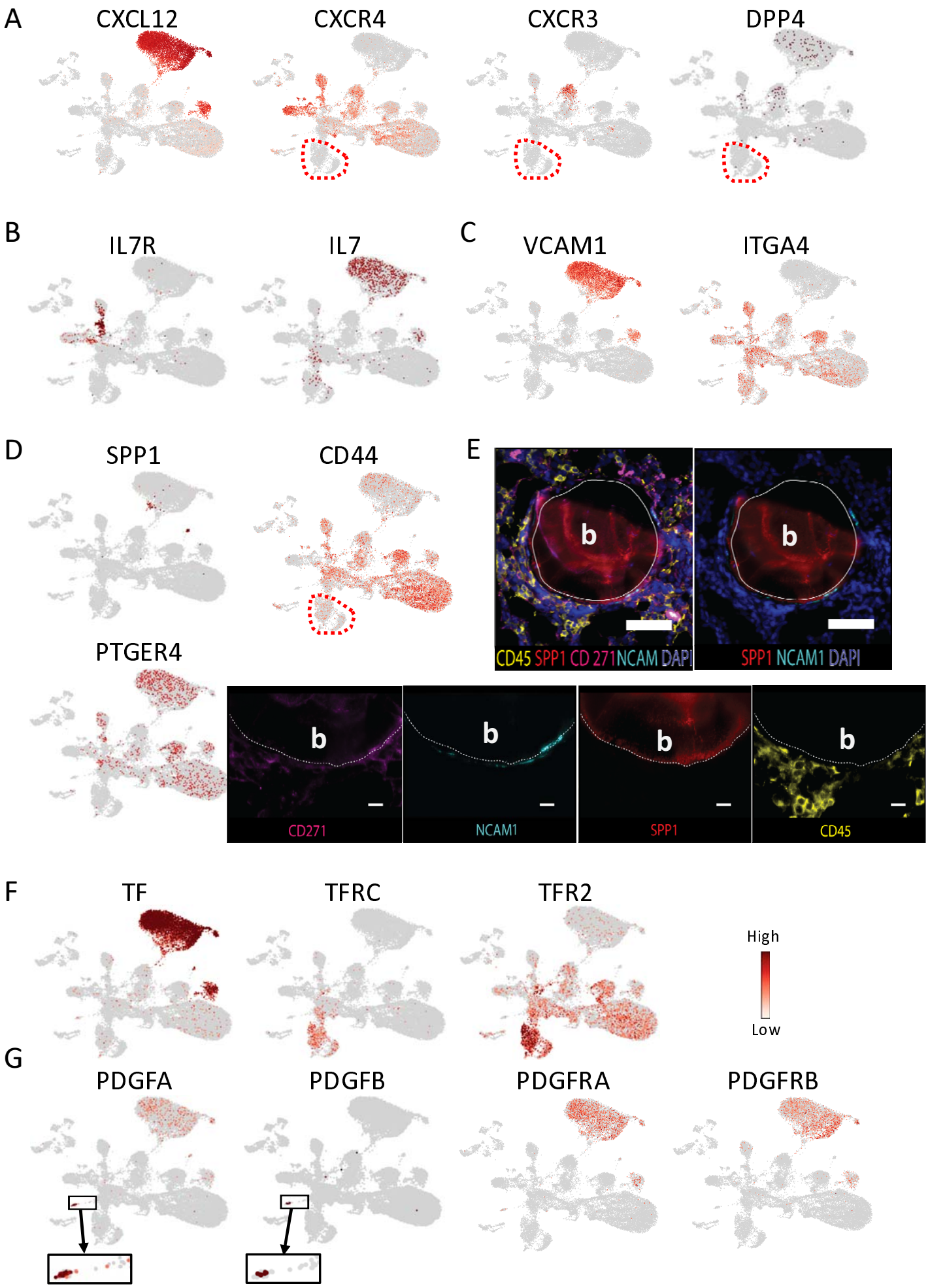
Abbreviation	Cell Type
GC	granulocytes
DC	Dendritic cells
Mk	megakaryocytes
Ery	erythroid cells
MC	monocytes
B	B cells
T	T cells
NK	natural killer cells
PC	plasma cells
HSPC	Hematopoietic stem and progenitor cells

B



Cell Type	Cluster No.
MSSC	3
OC	23
Mk	9
Ery	41
MC	36
B	32
T	11
NK	35
HSPC	7

Figure 6


A
CXCL12
CXCR4
CXCR3
DPP4
B
IL7R
IL7
C
VCAM1
ITGA4
D
SPP1
CD44
PTGER4
E
b
CD45 SPP1 CD 271NCAM DAPI
SPP1 NCAM1 DAPI
CD271
NCAM1
SPP1
CD45
F
TF
TFRC
TFR2
High
Low
G
PDGFA
PDGFB
PDGFRA
PDGFRB