A systematic dissection of human primary osteoblasts in vivo at single-cell resolution

Yun Gong1**, Junxiao Yang2**, Xiaohua Li3, Cui Zhou3, Yu Chen3, Zun Wang6, Xiang Qiu5, Ying Liu3, Huixi Zhang3, Jonathan Greenbaum1, Liang Cheng7, Yihe Hu2, Jie Xie2, Xuecheng Yang2, Yusheng Li2, Yuntong Bai8, Yu-Ping Wang8, Yiping Chen9, Li-Jun Tan3, Hui Shen1, Hong-Mei Xiao4,5, Hong-Wen Deng1,5

1Tulane Center for Biomedical Informatics and Genomics, Deming Department of Medicine, School of Medicine, Tulane University, New Orleans, LA 70112, USA
2Department of Orthopedics, Xiangya Hospital, Central South University, Changsha 410008, China
3Laboratory of Molecular and Statistical Genetics, College of Life Sciences, Hunan Normal University, Changsha 410081, China
4Center of Reproductive Health, System Biology and Data Information, Institute of Reproductive and Stem Cell Engineering, School of Basic Medical Science, Central South University, Changsha 410081, China
5School of Basic Medical Science, Central South University, Changsha 410008, China
6Xiangya Nursing School, Central South University, Changsha 410013, China
7Department of Orthopedics and National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha 410008, China
8Tulane Center for Bioinformatics and Genomics, Department of Biomedical Engineering, Tulane University, New Orleans, LA 70112, USA
3Department of Cell and Molecular Biology, School of Science and Engineering, Tulane University, New Orleans, LA 70112, USA
*Equal contribution

Correspondence to: Hong-Mei Xiao, Hong-Wen Deng; email: hmxiao@csu.edu.cn, hdeng2@tulane.edu
Keywords: single-cell RNA sequencing, osteoblasts, cellular heterogeneity, bone formation
Received: March 2, 2021 Accepted: June 19, 2021 Published: August 24, 2021

Copyright: © 2021 Gong et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Human osteoblasts are multifunctional bone cells, which play essential roles in bone formation, angiogenesis regulation, as well as maintenance of hematopoiesis. However, the categorization of primary osteoblast subtypes in vivo in humans has not yet been achieved. Here, we used single-cell RNA sequencing (scRNA-seq) to perform a systematic cellular taxonomy dissection of freshly isolated human osteoblasts from one 31-year-old male with osteoarthritis and osteopenia after hip replacement. Based on the gene expression patterns and cell lineage reconstruction, we identified three distinct cell clusters including preosteoblasts, mature osteoblasts, and an undetermined rare osteoblast subpopulation. This novel subtype was found to be the major source of the nuclear receptor subfamily 4 group A member 1 and 2 (NR4A1 and NR4A2) in primary osteoblasts, and the expression of NR4A1 was confirmed by immunofluorescence staining on mouse osteoblasts in vivo. Trajectory inference analysis suggested that the undetermined cluster, together with the preosteoblasts, are involved in the regulation of osteoblastogenesis and also give rise to mature osteoblasts. Investigation of the biological processes and signaling pathways enriched in each subpopulation revealed that in addition to bone formation,
preosteoblasts and undetermined osteoblasts may also regulate both angiogenesis and hemopoiesis. Finally, we demonstrated that there are systematic differences between the transcriptional profiles of human and mouse osteoblasts, highlighting the necessity for studying bone physiological processes in humans rather than solely relying on mouse models. Our findings provide novel insights into the cellular heterogeneity and potential biological functions of human primary osteoblasts at the single-cell level.

INTRODUCTION

Osteoblasts, which account for 4-6% of resident cells in bone, are derived from multipotential skeletal stem cells (SSCs) through the activation of signaling pathways regulated by osterix (OSX) and Runx-related transcription factor 2 (Runx-2) [1]. In addition, recent studies have demonstrated that periosteal cells [2] and growth plate chondrocytes [3] are also significant sources of osteoblasts. Osteoblasts are primarily known for their bone-building functions, including deposition of calcium phosphate crystals (e.g. hydroxyapatite), production of bone matrix constituents (e.g. type I collagen), as well as their ability to secrete a number of important proteins for bone metabolic processes such as integrin-binding sialoprotein (IBSP), secreted phosphoprotein 1 (SPP1), and bone gamma-carboxyglutamic acid-containing protein (BGLAP) [4]. In general, osteoblasts play a crucial role in the mineralization of the bone matrix [5].

Many osteoblasts ultimately differentiate into osteocytes, which become embedded in the bone matrix to form a communication network for the regulation of bone formation and resorption. The osteoblasts and osteocytes regulate the differentiation of osteoclasts, which are primarily involved in bone resorption activities [6]. For instance, osteoblasts can promote osteoclast proliferation by producing macrophage-colony stimulating factor (M-CSF) [7]. Osteoblasts can also produce the receptor activator of nuclear factor kappa-B ligand (RANKL) and osteoprotegerin (OPG) to modulate osteoclast proliferation through the RANKL/RANK/OPG pathway [8]. On the other hand, osteoclasts may also regulate bone formation by osteoblasts [9]. The complex dynamics between the major bone cells control the delicate balance of bone formation/resorption that is critical for maintaining bone health.

While osteoblasts are typically associated with bone remodeling processes, previous studies have demonstrated that they also have the ability to interact with immune cells via the regulation of hematopoietic stem cell (HSC) niches. Specifically, primary osteoblast lineage cells synthesize granulocyte colony-stimulating factor (G-CSF), granulocyte macrophage CSF (GM-CSF), IL-1, IL-6, lymphotoxin, TGFβ, TNFα, leukemia inhibitory factor (LIF), and stem cell factor (SCF); all of which play crucial roles in hemopoiesis [10]. Additionally, osteoblasts are responsible for modulating angiogenesis, which supports the high degree of vascularization in the bone marrow environment to provide sufficient oxygen for bone metabolism [11].

Cellular heterogeneity is an essential characteristic of various cell populations. Although every cell shares almost the same genome, each cell acquires a unique identity and thus specific functional capabilities through molecular coding across the DNA, RNA, and protein conversions [12]. Therefore, even for the same known classical cell types, cells may be further classified into distinct subpopulations due to systematic differences in their gene expression profiles. Emerging evidence from in vitro studies in mice has revealed notable cell-to-cell heterogeneity within the osteoblast cell population. For instance, one study performed the soft agarose cloning technique on rat osteoblastic cells and detected diverse gene expression patterns in osteoblast cells at different stages of cellular differentiation [13]. Another study showed that the expression levels of osteoblastic specific markers including osteopontin, bone sialoprotein, and osteocalcin, varied in the mature osteoblasts of mice with different cellular morphology, suggesting that even these terminally differentiated osteoblasts were composed of multiple subgroups instead of a single unique cell group [14]. While these early efforts revealed the existence of osteoblast heterogeneity, the functional differences between distinct osteoblast subtypes were not well characterized.

The recent development of the state-of-the-art single-cell RNA sequencing (scRNA-seq) technology is expected to provide the most powerful approach to study the nature and characteristics of cell-to-cell heterogeneity. Compared with the conventional bulk RNA-seq approaches, scRNA-seq can reveal complex and rare cell populations, track the trajectories of distinct cell lineages in development, and identify novel regulatory relationships between genes [15]. Accumulating evidence from a few early scRNA-seq studies in vivo in mice has demonstrated the existence
of several osteoblast subtypes. Baryawno et al. [16] identified pre- and mature osteoblast subpopulations based on their transcriptional profiles, while Tikhonova et al. [17] classified the osteoblasts into three subgroups and revealed significant changes in subgroup proportions under hematopoietic stress conditions induced by chemotherapy treatment. Although our understanding of osteoblast heterogeneity has evolved substantially based on these studies, the cell subtype characteristics of primary osteoblasts in vivo in humans has not been successfully explored. Studying the cellular heterogeneity in mice, or even the cultured cells from humans (though not yet existing), although useful, may not be ideal for studying human disease etiology. This is because the cell identity may vary between mice and humans, and cell culturing may systematically alter the gene expression of the studied cells [18].

In this study, we successfully performed the first unbiased examination of the in vivo cellular landscape of freshly isolated osteoblasts via scRNA-seq from one 31-year-old male with osteoarthritis and osteopenia. We identified three distinct cell subtypes along with their differentiation relationships based on the transcriptional profiling of 5,329 osteoblast cells from the femur head of one human subject. We then compared the most differentially expressed genes (DEGs) of each cluster with known cell characterizing markers. Further, we identified distinct functional characteristics of each cell subpopulation suggesting their involvement in bone metabolism, angiogenesis modulation, as well as hematopoietic stem cell (HSC) niche regulation. The discovery of osteoblast subtypes is well beyond the scope of current gene expression studies for bone health and represents an important and necessary step to provide novel insights into bone physiological processes at the refined single-cell level.

RESULTS

Osteoblasts identification

We applied an established protocol for in vivo human osteoblast isolation [19] to obtain the alkaline phosphatase (ALPL)\textsuperscript{high}/CD45/34/31\textsuperscript{low} cells from the femur head-derived bone tissue of one human subject (31-year-old Chinese male) with osteoarthritis and osteopenia through fluorescence-activated cell sorting (FACS). Several studies [20, 21] have demonstrated that this isolation protocol can successfully recover a high proportion of osteoblasts based on the elevated expression levels of osteoblastic markers via quantitative polymerase chain reaction (qPCR) or bulk RNA sequencing. In total, 9,801 single cells were encapsulated for cDNA synthesis and barcoded using the 10x Genomics Chromium system, followed by library construction and sequencing (Figure 1A and Supplementary Figure 1A [22]). After quality control (Supplementary Figure 1B [22]), we obtained 8,557 cells, with an average of 2,365 and median of 2,260 genes detected per cell (Supplementary Table 1 [22]). A recently developed dimension reduction technique for scRNA-seq analysis, uniform manifold approximation and projection (UMAP) [23], was applied to project the gene expression profiles on a two-dimensional panel for visualization of cellular heterogeneity (Supplementary Figure 1C [22]). Compared with t-SNE [24], another commonly used dimension reduction method in single cell analysis, UMAP provides the useful and intuitively pleasing feature that it preserves more of the global structure [23], which is important to identify the systematic differences of gene expression patterns between different clusters through the inter-cluster distances. UMAP has already been successfully and now commonly used for scRNA-seq data dimension reduction in recent scRNA-seq studies [25–27]. After clustering the cells into six distinct subsets (C1-C6) by the k-nearest neighbor algorithm [28], we used pairwise differential expression analysis for comparing each individual cluster against all the others to identify the DEGs of each subtype (Supplementary Figure 1D [22]).

While ALPL was enriched in clusters C1, C2 and C5, the osteoblast specific marker, RUNX2 (encoded by RUNX2) [29], was only enriched in clusters C1 and C2, suggesting the presence of contamination by other cell types during the cell isolation process (Supplementary Figure 1D [22]). Based on the identified marker genes and gene ontology (GO) enrichment analysis of the differentially expressed genes in each cluster (Supplementary Figure 1D–1F [22]), the contaminant cells were classified as: 1) two nucleated erythrocyte groups (C3 and C4, expressing hemoglobin coding genes HBB and HBA1) [30]; 2) one smooth muscle cell group (C5, expressing transgelin coding gene, TAGLN [31], and CNN1, which is specific to differentiated mature smooth muscle cells) [32]; and 3) one neutrophil group (C6, expressing neutrophil related genes S100A8 and ELANE) [33, 34]. These findings suggest that the protocol for human osteoblast isolation based on ALPL\textsuperscript{high}/CD45/34/31\textsuperscript{low} may not be sufficient for specifically isolating osteoblasts alone, since ALPL is also highly enriched in other mesenchymal cell-derived cells such as vascular smooth muscle cells [35]. Notably, by comparing the expression profiles between osteoblasts and contaminant cells, we found that the average fold change (avg FC) of alpha-1 type I collagen (COL1A1, encoded by COL1A1) was large (avg FC =4.497, Supplementary Table 2 [22]), suggesting that positive selection based on the combination of ALPL and COL1A1 would be a more
**appropriate choice for human primary osteoblast isolation.** While several studies [17, 36] in mice have used COL1A1 for osteoblast sorting, few (if any) studies have adopted this marker for human primary osteoblast isolation.

**scRNA-seq identifies multiple cell subtypes in human osteoblasts**

To investigate the cellular diversity within osteoblasts, we extracted the clusters C1 and C2 with high

---

**Figure 1. scRNA-seq analysis of human osteoblasts.** (A) Study overview. (B) Three osteoblast clusters. UMAP visualization of 5,329 osteoblasts, colored by clustering. (C) Proportion of three osteoblast clusters. Colored by clustering. (D) Cluster signature genes. Violin plots showing the log-transformed normalized expression levels of the two most significant marker genes in clusters O1, O2, and O3, respectively. (E) Log-normalized expression of adipocyte and chondrocyte biomarkers in osteoblast clusters. (F) Log-normalized expression of LEPR mesenchymal cell related markers in osteoblast clusters. (G) Log-normalized expression of OPG, RANKL, IL6, and IL7 in three clusters.
expression levels of osteoblast related markers, i.e., RUNX2 and COL1A1 (Supplementary Figure 1D [22]).

To ensure the high quality of the osteoblasts data for downstream analysis, we further removed osteoblasts with more than 5% of the transcripts attributed to mitochondrial genes, following the quality control criterion applied in several scRNA-seq studies in bone field [37–39]. 5,329 high quality sequenced osteoblasts were retained for downstream analysis. We noticed that the biomarkers of other mesenchymal-derived cells such as adipocytes and chondrocytes were not enriched in the remaining cells [38] (Figure 1E). Further, although the remaining cells highly expressed the leptin receptor positive (LEPR+) mesenchymal cell related markers EBF transcription factor 3 (encoded by EBF3), C-X-C motif chemokine ligand 12 (encoded by CXCL12), and platelet derived growth factor receptor alpha (encoded by PDGFRα) [40] (Figure 1F), the high enrichment of osteoblast related markers (e.g., COL1A1, ALPL, RUNX2, etc.) suggests that these cells are actually osteoblasts rather than LEPR+ mesenchymal cells or CXCL12-abundant reticular (CAR) cells [41]. We identified three cell subtypes of osteoblasts (Figure 1B–1D) based on their transcriptional profiles and by comparing classical osteoblastic markers [42] with three published scRNA-seq datasets of mouse osteoblasts [16, 17, 41]. The cell subtypes were annotated as: 1) O1 (65.77%), pre-osteoblasts, with relatively high expression levels of LEPR+ mesenchymal cell markers LEPR (encoded by LEPR) [43] and vascular cell adhesion molecule 1 (encoded by VCAM1) [44]; 2) O2 (25.39%), mature osteoblasts, which showed highest expression levels of COL1A1 (COL1A1) and osteogenesis-associated genes including BGLAP (also known as osteocalcin, encoded by BGLAP), SPP1 (also known as osteopontin, encoded by SPP1), IBSP (encoded by IBSP) and interferon induced transmembrane protein 5 (encoded by IFITM5) [45, 46]; 3) O3 (8.84%), undetermined osteoblasts, which not only expressed several LEPR+ mesenchymal cell-associated genes (e.g., LEPR and VCAM1) but are also distinguished from the other subtypes by distinctively expressing high levels of nuclear receptor subfamily 4 group A member 1 and 2 (encoded by NR4A1 and NR4A2). Although pre- and mature osteoblasts are vague concepts largely suggested by in vitro studies rather than discrete in vivo cell types, we use these terms to better describe the specific characteristics of clusters O1 and O2. Notably, few osteoblasts in this cell population expressed OPG and RANKL (Figure 1G). This is consistent with the result proposed by Tat et al. [47] that the expression levels of OPG and RANKL are significantly reduced in the osteoblasts from human osteoarthritic bone. In addition, we also noticed the high expression of the inflammation markers IL6 and IL7 in cluster O3 and O1, respectively (Figure 1G). Previous studies have proposed that IL6 is one of several pro-inflammatory cytokines present in individuals with confirmed clinical diagnosis of osteoarthritis [48, 49].

We further examined the transcriptional profiles of the three identified osteoblast subtypes and found that in addition to the cell markers, some bone development regulators were highly enriched in the pre-osteoblast and mature osteoblast clusters. For instance, pre-osteoblasts expressed insulin-like growth factor-binding protein 2 and 4 (IGFBP2 and IGFBP4, encoded by IGFBP2 and IGFBP4) (Figure 2A). Previous studies have considered IGFBP2 as a stimulator for osteoblast differentiation through positive regulation of the AMP-activated protein kinase (AMPK) [50], and demonstrated that IGFBP4 could stimulate adult skeletal growth in males [51]. Meanwhile, tenascin, a glycoprotein encoded by TNC that modulates osteoblast mineralization via matrix vesicles [52], was highly enriched in mature osteoblasts. On the other hand, the cluster O3 showed high expression level of osteomodulin (OMD) (Figure 2A). It has been proposed that OMD induces endochondral ossification through PI3K signaling, which is an essential process for long bone formation [53].

In addition to the known osteoblast subtypes (pre-osteoblast and mature osteoblast), we also identified one potential novel osteoblast subtype (cluster O3), which is a major source for orphan nuclear receptors coding genes NR4A1 and NR4A2. Although previous studies have reported the presence of NR4A1 in mice osteoblasts induced by parathyroid hormone (PTH) in vitro [54], the enrichment of this gene in osteoblasts in vivo has not yet been proposed or identified earlier in mice or in humans. Several studies have reported that NR4A receptors have an essential impact on bone metabolism by regulating the expression of several osteoblastic marker genes such as SPP1, BGLAP, COL1A1, and ALPL [55–57]. Additionally, the overexpression of NR4A2 led to increased expression of ALPL, COL1A1, and BGLAP in osteoblasts [55–57]. The immunofluorescence on mouse femur illustrated the co-staining of the osteoblast marker ALPL with the biomarker NR4A1 on the bone surface, hereby supporting the expression of NR4A1 in osteoblasts in vivo (Figure 2F).

While all the osteoblast clusters highly expressed the osteoblastogenic gene RUNX2, they may affect bone development from different aspects. To further investigate the distinct biological functions related to osteogenesis among the three clusters of osteoblasts, we performed GO enrichment analysis based on the DEGs (avg_FC > 1.200 and p_val_adj < 0.05, Supplementary Table 3 [22]) in each cluster. All the clusters were enriched in GO terms related to bone development including “extracellular matrix organization” and
“extracellular structure organization” with a relatively high gene ratio (Figures 2B–2E). The pre-osteoblasts and mature osteoblasts are involved in the extracellular matrix (ECM) formation through production of different types of collagen, such as types 3, 14, and 18 in pre-osteoblasts and types 1, 12, and 13 in mature osteoblasts (Supplementary Tables 4, 5 [22]). The mature osteoblasts were uniquely enriched for “bone mineralization”, while mature and undetermined osteoblasts were both enriched for “osteoblast differentiation” (Figure 2C–2E and Supplementary Tables 5, 6 [22]). Additionally, the undetermined osteoblasts were uniquely enriched for “regulation of osteoblast differentiation” and “positive regulation of osteoblast differentiation” (Figure 2D). Surprisingly, few GO terms related to osteoblast differentiation are enriched in the pre-osteoblasts, suggesting that the undetermined and mature osteoblasts may modulate the osteoblast differentiation processes to a larger extent than pre-osteoblasts.

Dynamic gene expression patterns reveal the differentiation relationship between different osteoblast subtypes

In order to reveal the differentiation dynamics of the osteoblast cell population, we reconstructed the
developmental trajectory of the three identified clusters of osteoblasts. Since all the three clusters expressed the LEPR+ mesenchymal cell related markers EBF3, CXCL12, and PDGFRA (Figure 1F), we suggested that all osteoblasts are in the LEPR cells derived osteoblastic lineage. All cells were contained within one cellular lineage without any bifurcations (Figure 3A), suggesting that only one terminal subtype exists in the osteoblastic population. While pre-osteoblasts and undetermined osteoblasts were highly enriched in the early stages of pseudotime, mature osteoblasts were distributed in the terminal stages of the osteoblastic lineage (Figure 3A, 3C). To strengthen the trajectory inference, we also assessed the transcriptional continuum of the cell lineage. Since the scRNA-seq data is difficult to model with a parametric curve (e.g., least squares regression, etc.) due to the high sparsity and variability, we used the loess curves to reveal the trends of gene expression along with the pseudotime. The results showed that the expression of the LEPR+ mesenchymal cell associated genes LEPR and VCAM1 [58] as well as cluster O3 specific markers, NR4A1 and NR4A2, decreased over pseudotime, while the expression levels of osteogenic markers such as RUNX2, BGLAP, SPARC [59], and COL1A1 were the highest at the end stages of pseudotime (Figure 3B). This result is consistent with the findings from other studies that have examined the transcriptional continuum in osteoblastic differentiation [42]. The analysis of the gene expression profiles of mouse osteoblasts cultured in vitro showed that the expression levels of NR4A1 rapidly increased from day 5 to day 7 whereas the expression of ALPL remained relatively stable throughout the osteoblast development (Figure 3D), further supporting the conclusion that the undetermined osteoblasts are involved in the early stages of the osteoblastic lineage.

Figure 3. Trajectory inference of human osteoblasts. (A) Reconstructed cell differentiation trajectory of human osteoblasts. The upper-right trajectory plot in the square indicates the direction of pseudotime. (B) Expression levels (log-normalized) of indicated genes in the three osteoblast subtypes with respect to their pseudotime coordinates. The x-axis indicates the pseudotime, while the y-axis represents the log-normalized gene expression levels. The color corresponding to the three different osteoblast subsets. The loess regression was applied to fit the relationship between pseudotime and expression level. (C) Cell distribution based on the pseudotime coordinates. The x-axis is the pseudotime and the y-axis represents the osteoblast subtypes. (D) Expression levels of ALPL and NR4A1 in mouse osteoblasts in vitro at day 0, 5 and 7, respectively. N.S., not significant, *p-adjusted ≤ 0.05, **p-adjusted ≤ 0.01, ***p-adjusted ≤ 0.005.
**Pre-osteoblasts regulate angiogenesis through multiple signaling pathways**

Although evidence has shown that osteoblasts can regulate angiogenesis [11], few studies have explored this relationship at the single-cell level. Based on functional enrichment of the highly expressed genes, we found that the GO term “regulation of angiogenesis” was enriched in pre-osteoblasts with a gene ratio of 0.10 but not in other clusters (Figure 4A). We further investigated the enriched genes in pre-osteoblasts and found that 22 genes are involved in angiogenesis (Supplementary Table 4 [22]). In particular, the pre-osteoblasts showed significantly higher expression levels of *SFRP1*, *MDK*, and *THBS1* compared with the other clusters (Figure 4B). It has been demonstrated that Secreted Frizzled-related protein-1 (sFRP1, encoded by *SFRP1*), a modulator of the Wnt/Fz pathway, can modify mesenchymal cell capacities enabling mesenchymal cells to increase vessel maturation [60]. Midkine (MDK, encoded by *MDK*) is an enhancer of angiogenesis, and MDK expression has been shown to be positively correlated with vascular density in bladder tumors [61]. Thrombospondin 1 (THBS1, encoded by *THBS1*) is considered to be a potent endogenous inhibitor of angiogenesis through antagonization of vascular endothelial growth factor (VEGF) activity [62].

Next, we investigated the signaling pathways enriched in pre-osteoblasts that are related to angiogenesis.

![Figure 4. Potential functions of pre- and undetermined osteoblasts in angiogenesis and hemopoiesis. (A) Angiogenesis and hematopoiesis modulation related GO terms enriched in the three clusters. The x-axis represents the clusters and the y-axis is the GO terms related to the angiogenesis and hematopoiesis regulation. The size of the dot indicates the gene ratio and the color indicates the adjusted p-values. (B) Angiogenesis associated genes enriched in cluster O1. X-axis represents the three clusters and y-axis reflects log-normalized gene expression levels. Stars indicate the significance levels of the gene expression difference between two clusters by Wilcoxon signed-rank test. N.S., not significant, *p*-adjusted ≤ 0.05, **p*-adjusted ≤ 0.01, *** p*-adjusted ≤ 0.005. (C) Gene expression pattern in enriched pathways. Squares show enriched DEGs in the corresponding terms (rows). Color indicates the expression value of the DEGs (average logFC). (D) Hemopoiesis factors enriched in clusters O1 and O3.](image-url)
Kyoto encyclopedia of genes and genomes (KEGG) pathway analysis revealed three highly enriched pathways including PI3K-Akt, phospholipase D, and relaxin signaling pathways, which are known to be implicated in angiogenesis modulation [63–65] (Figure 4C). Several genes related to the “regulation of angiogenesis” GO term were enriched in these pathways (Supplementary Table 4 [22], Figure 4C) including COL4A2, TGFBR2, and IL7. These findings suggest that pre-osteoblasts may potentially regulate angiogenesis to a larger extent than other osteoblast subtypes.

**Osteoblast populations modulate the development of HSC niche**

Osteoblasts are part of the stromal cell support system that provides critical regulators of hematopoiesis [10]. It has been reported that osteoblasts lining the bone surface can generate an extracellular environment supporting non-skeletal hematopoietic cells [66]. To identify the specific roles of each osteoblast subpopulation in hematopoiesis regulation, we first examined the expression patterns of hematopoiesis related factors in all three osteoblast subtypes. We found that while both pre-osteoblasts and undetermined osteoblasts highly expressed insulin-like growth factor 1 (IGF-1, encoded by *IGF1*), which is a growth factor that has been shown to be responsible for burst-like growth of early erythroid progenitor cells *in vitro* [67], the pre-osteoblasts may also play an important role in modulating macrophage development due to the high expression of M-CSF (encoded by *CSF1*, Figure 4D) [68]. Further, the high enrichment of the stem cell factor (SCF, encoded by *KITLG*, Figure 4D) in pre-osteoblasts suggests that these cells may be crucial for preventing HSC apoptosis [69]. Since Aguila et al. [70] proposed that the overexpression of human interleukin 7 (*IL-7*, encoded by *IL7*) increases the development of B cells in female mice, it is plausible that the pre-osteoblasts, with high expression of IL7 (Figure 4D), may affect the development of B cells in humans.

We also found that the undetermined osteoblasts represent a major source of C-C motif chemokine ligand 2 (CCL-2, encoded by *CCL2*) (Figure 4D), which is a critical factor for monocyte recruitment in acute inflammatory response [71]. Additionally, the undetermined osteoblasts expressed interleukin 6 (*IL-6*, encoded by *IL6*), which displays a strong synergy with IL-7 to stimulate naïve CD8+ T cells [72]. Therefore, the undetermined osteoblasts may coordinate with pre-osteoblasts to support the hematopoietic niche. However, few hematopoietic factors were enriched in mature osteoblasts, suggesting that they may have limited impact on the hematopoietic system.

We investigated the predicted functions associated with hematopoiesis regulation enriched in the pre-osteoblasts and undetermined osteoblasts. Notably, while the GO term “positive regulation of hemopoiesis” was enriched in pre-osteoblasts, the highly expressed genes of the undetermined cluster were related to the negative regulation of hemopoiesis (Figure 4A). Apart from the growth factors mentioned above, pre-osteoblasts also had a significantly higher expression level of growth arrest-specific 6 (GAS6, encoded by *GAS6*) (Figure 4D), which can induce natural killer cell development via its positive regulatory effect on fibromyalgia syndrome-like tyrosine kinase 3 (FLT3) signaling in CD34+ hematopoietic progenitor cells (HPCs) [73]. Meanwhile, the undetermined subgroup had significantly higher expression levels of zinc-finger protein 36 (ZFP36, encoded by *ZFP36*) (Figure 4D), which can inhibit the erythroid differentiation [74]. These findings suggest that these two osteoblast subgroups may play different roles in hemopoiesis regulation, and that the relative proportions and functional levels of these subtypes within the osteoblast cell population may be crucial for maintaining the homeostasis of hematopoiesis.

**Transcriptional divergence of human and mouse osteoblasts**

Due to the limited *in vivo* transcriptomics studies in mouse osteoblasts at different developmental stages, we first compared the expression profiles of osteoblasts acquired from humans *in vivo* and the osteoblasts from mice cultured *in vitro*. The *in vitro* mouse osteoblast data demonstrate the gene expression levels at three different time points (Day 5, 14, and 21), which is considered as the reference for us to identify the divergence of transcriptional profiles between human and mouse osteoblasts in multiple development stages. Several fundamental differences were observed between humans and mice. For instance, all three human osteoblast clusters *in vivo* expressed secretory leukocyte protease inhibitor (encoded by *SLPI*) (Figure 2A), a serine protease inhibitor that promotes cell migration and proliferation while also suppressing the inflammatory response [75]. This conflicts with the previous *in vitro* findings in mice [76], which demonstrated the enrichment of SLPI in mature but not in pre-osteoblasts. Additionally, we found that the human pre-osteoblasts isolated *in vivo* showed a significantly higher expression level of SLPI compared with mature osteoblasts (Figure 2A). In further contrast, the mature mouse osteoblasts cultured *in vitro* were highly enriched for myristoylated alanine rich protein kinase C (PKC) substrate (MARCKS, encoded by *MARCKS*), while this gene was highly expressed by the pre-osteoblasts but not the mature osteoblasts in humans.
in vivo (Figure 2A). It has been shown that MARCKS is the PKC-δ effector which modulates cathepsin K secretion and bone resorption in osteoclasts [77]. Remarkably, we also found that CPE, encoding carboxypeptidase E, was only enriched in mature osteoblasts in vivo in humans (Figure 2A), which is inconsistent with the previous mouse study in vitro finding that the expression level of CPE decreases throughout osteoblast development [76]. While CPE plays an important role in RANKL-induced osteoclast differentiation [78], it is plausible that a reduction in the proportion of mature osteoblasts could attenuate osteoclast proliferation.

To better understand the systematic differences in osteoblastic transcriptional profiles between human and mouse, we then integrated our sequencing data with the publicly available data from a previous scRNA-seq study [17] on mouse osteoblasts in vivo (Figure 5A). After correcting for the potential batch effects between independent experiments using canonical correlation analysis (CCA) [28], a moderate positive correlation (R = 0.53, p-value < 2.23e-16) was observed between the human and mouse osteoblastic transcriptomic profiles (Figure 5B). This suggests that while some common characteristics exist between human and mouse osteoblasts, there are considerable differences in the gene expression profiles at the single-cell level.

To further investigate the shared and distinct features among human and mouse osteoblasts, we applied unbiased clustering analysis after dimension reduction

Figure 5. Integrated cross-species analysis between human and mouse osteoblasts in vivo. (A) UMAP visualization of human and mouse osteoblast integration. The upper-left plot represents the three osteoblast clusters in mice, the upper-right plot indicates the three osteoblast subtypes in humans. The bottom plot represents the five clusters after the integration, colored by clusters. The middle-right is the integrated data colored by different species. (B) Correlation of gene expression among different osteoblast datasets after CCA integration. Each dot represents an individual gene. Axes measure the average gene expression levels (logFC) in the indicated subject. Correlations were measured by Spearman correlation coefficients (R) (p < 0.01). (C) Log-normalized expression of NR4A1 in human and mouse integrated data. (D) Proportion of the human and mouse osteoblasts in each cluster after CCA integration, colored by different species. The x-axis represents five different clusters and y-axis indicates the proportion.
by UMAP [23]. The result showed that human osteoblast cluster O2 overlapped with mouse osteoblast subtype M2, which represents the mature osteoblasts in mice [17] (Figure 5A, 5C), indicating that the overall gene expression patterns of human and mouse mature osteoblasts are highly similar. Furthermore, we also found strong correlations in the expression patterns between small subgroups in human pre-osteoblasts (cluster O1) and the mouse osteoblast subtype M1 [17] (Figure 5A, 5C). In contrast, the human osteoblast cluster O3 did not overlap with any osteoblast subtypes in mice (Figure 5A). However, we did observe that NR4A1 was expressed in cluster O1 and M1 and M3 (Figure 5B). Since few human osteoblasts in cluster O1 expressed NR4A1 (Figure 1D), mouse osteoblast cluster M1 is the major source of the NR4A1 in the cluster O1 and M1. Therefore, although NR4A1 is expressed by both human and mouse osteoblasts, heterogeneity still exists between human and mouse NR4A1+ osteoblasts. These findings suggest that the transcriptional profiles of human and mouse osteoblasts may demonstrate systematic differences in early cellular developmental stages but share similar features at the terminal stage of osteoblastic development lineage. In addition, the comparison between our data and osteoblasts from mice cultured in vitro [76] revealed that several genes (e.g., SLPI, MARCKS, CPE etc.), which are not correlated with osteoarthritis pathogenesis, show fundamental differences of expression levels between human and mouse osteoblasts. This suggests that the systematic variations may still exist between healthy human and mouse osteoblasts. Therefore, studies based on osteoblasts acquired from mouse models may introduce notable bias when inferring the biological characteristics of human osteoblasts, especially for those that have not yet reached a mature state.

DISCUSSION

While it is now well-appreciated that human osteoblast heterogeneity may be distinguished by differentiation stages [13], the full spectrum of cells that comprise the osteoblast population, especially in vivo in humans, has remained elusive. In this study, for the first time, we classified the freshly isolated primary osteoblasts from human bone (without any in vitro culturing) into three subpopulations based on systematic differences in gene expression profiles and revealed their distinct functional roles in bone metabolism as well as in the regulation of angiogenesis and hematopoiesis. Further, in contrast to the results proposed by Tarkkonen et al. [79], our results demonstrate systematic differences in the transcription profiles between humans and mice, emphasizing the importance of in vivo studies in humans. Here, we highlight some of the key findings. First, our results indicate that the osteoblast isolation technique typically used in the field [19], FACS isolation based on ALPL, was not sufficient for specific isolation of osteoblasts since the isolated cell population included approximately 40% contamination by other cell types. By comparing the transcriptomic profiles between osteoblasts and contamination cells, we hypothesized that the combination of ALPL and COL1A1 would reduce contamination in osteoblast selection. In addition to known osteoblast subtypes, pre- and mature osteoblasts, we also identified one rare osteoblast subpopulation which highly expressed NR4A1 and NR4A2. Based on the gene expression patterns and the inferred osteoblastic lineage trajectory, we found that: 1) Although both pre-osteoblasts (LEPR<sup>high</sup>/VCAM1<sup>high</sup>) and undetermined osteoblasts (NR4A1<sup>high</sup>/NR4A2<sup>high</sup>) are in the differentiation lineage ordering, pre-osteoblasts are primarily responsible for ECM organization during bone formation processes as well as inducing hematopoiesis and modulating angiogenesis, while undetermined osteoblasts are mainly involved in the regulation of osteoblastogenesis and inhibition of hematopoiesis; 2) mature osteoblasts (IBSP<sup>high</sup>/BGALP<sup>high</sup>) arise at the terminal stages of cellular differentiation and are crucial for bone mineralization. We used pre- and mature osteoblasts to annotate clusters O1 and O2, which are vague concepts largely suggested from in vitro literature rather than in vivo studies, mainly based on the expression levels of the gene markers related to specific differentiation stages in the osteoblastic lineage (e.g., LEPR in early stages, BGLAP in terminal stages, etc.). Consistent with the transcriptional characteristics of these two clusters, the result of the pseudotime analysis showed that while cluster O1 cells are enriched in the early stages of the osteoblastic lineage, cluster O2 cells are distributed in the terminal stages.

Despite the novelty of this scRNA-seq study in freshly isolated human osteoblasts, an important limitation is that all the cells were collected from the femur head of one 31-year-old Chinese male subject with osteoarthritis and osteopenia. This might introduce some bias in profiling the gene expression patterns of osteoblasts. However, since no study has demonstrated the disease-associated changes in cell compositions of human osteoblasts at the single-cell resolution, there is no evidence to suggest the existence of the undetermined osteoblasts is due to the bias stemming from the patient’s condition. Further, if the existence of cluster O3 is due to the osteoarthritis, specific markers NR4A1 and NR4A2, would have important roles in osteoarthritis pathogenesis. While previous studies have reported NR4A1 [80] and NR4A2 [81] to be regulators of inflammatory responses in chondrocytes, no study has proposed the relationship between osteoarthritis and
NR4A1 produced by osteoblasts. In addition, scRNA-seq studies in both normal mouse osteoblasts in vivo performed by Tikhonova et al. [17] and transcriptional profiles of mouse osteoblasts in vitro [76] showed the presence of NR4A1 in normal osteoblasts. Thus, we speculate that cluster O3 may also exist in the normal osteoblastic lineage.

Another limitation is that, due to limited cell numbers, the differentiation relationships between pre-osteoblasts and undetermined osteoblasts have not been clarified here. Additionally, substantial technical differences between the cross-species comparison might exist which might be difficult to be normalized with CCA method, though there is no evidence supporting this. Future studies based on a larger sample size are needed to uncover the disease associated changes in cell subtype compositions, as this will have significant implications for the development of novel therapeutics. Despite these potential limitations, our results provide the first necessary and valuable insights into the cellular heterogeneity of osteoblasts along with a comprehensive and systematic understanding of the cell-intrinsic and cell-specific mechanisms which may underlie bone metabolism and associated disorders.

MATERIALS AND METHODS

Study population

The clinical study was approved by the Medical Ethics Committee of Central South University and written informed consent was obtained from the study participant. The study population included one 31-year-old male with osteoarthritis and osteopenia (BMD T-score: 0.6 at lumbar vertebrae, -1.1 at total hip), who underwent hip replacement at the Xiangya Hospital of Central South University. The subject was screened with a detailed questionnaire, medical history, physical examination, and measured for bone mineral density (BMD) before surgery. Subject was excluded from the study if he had preexisting chronic conditions which may influence bone metabolism including diabetes mellitus, renal failure, liver failure, hematologic diseases, disorders of the thyroid/parathyroid, malabsorption syndrome, malignant tumors, and previous pathological fractures [82]. The femur head was collected from the patient during hip replacement surgery. The specimen was shortly stored in 4°C and transferred to the laboratory within 30 mins, where it was then processed immediately after delivery.

Mice

2 months old female C57BL/6J wild-type mice were purchased from Jackson Laboratory (Bar Harbor, ME, USA). All mice were housed in pathogen-free conditions and fed with autoclaved food. All experimental procedures were approved by the Ethics Committee of Xiangya Hospital of Central South University.

BMD measurement

BMD (g/cm²) was measured using DXA fan-beam bone densitometer (Hologic QDR 4500A, Hologic, Inc., Bedford, MA, USA) at the lumbar spine (L1–L4) and the total hip (femoral neck and trochanter) as described previously by our group [83, 84]. According to the World Health Organization definition [85] and the BMD reference established for Chinese population [86], subject with a BMD of 2.5 SDs lower than the peak mean of the same gender (T-score ≤ −2.5) were determined to be osteoporotic, while subject with -2.5 < T-score < -1 are classified as having osteopenia and subject with T-score > -1.0 are considered healthy.

Isolation of osteoblasts

Osteoblasts were extracted from the human femur head based on the widely used dissociation protocols [19] with a few minor adjustments. Briefly, bone tissue samples were chopped into small fragments and washed twice with phosphate-buffered saline (PBS). These fragments, containing a mixture of cortical and trabecular bone, were then incubated with a highly purified, endotoxin-free type II collagenase (1 mg/ml, Cat: A004174-0001, Sangon Biotech, Shanghai, China) at 37.0°C for 30 mins and 60 mins in the first and second round of digestion respectively. After the incubation, the solution was filtered through a 70 μm filter and incubated with red blood cell lysis buffer (Cat: R1010, Solarbio Science and Technology CO., Beijing, China) for 5 mins. The collected cells were washed twice with PBS.

Fluorescence-activated cell sorting (FACS) enrichment for osteoblasts

Before FACS, collected cells were incubated with human CD31/34/45-PE (Cat:303106, Cat:343606, Cat:304008, BioLegend, San Diego, CA, USA), human ALPL-APC (Cat: FAB1448A, R&D Systems, Minneapolis, MN, USA), and 7-AAD (488 nm) antibody (R&D Systems, Minneapolis, MN, USA). FACS was performed on a BD FACS Aria II sorter (Becton-Dickinson, Franklin Lakes, NJ, USA), where dead cells and debris were excluded by forward scatter (FSC) versus side scatter (SSC), and live cells were further enriched by negative selection of 7-AAD. ALP+/CD45/34/31- cells were collected as osteoblasts (Supplementary Figure 1) [19]. The percentage of live cells was 80% after filtering the dead cells. Gating schemes were established with fluorescence-minus-one (FMO) controls.
Single-cell RNA-seq (scRNA-seq) library preparation and sequencing

scRNA-seq libraries were prepared using Single Cell 3' Library Gel Bead Kit V3 following the manufacturer’s guidelines (https://support.10xgenomics.com/single-cell-gene-expression/library-prep/doc/user-guide-chromium-single-cell-3-reagent-kits-user-guide-v3-chemistry). Single cell 3’ Libraries contain the P5 and P7 primers used in Illumina bridge amplification PCR. The 10x Barcode and Read 1 (primer site for sequencing read 1) were added to the molecules during the gel bead-in emulsions reverse transcription (GEM-RT) incubation. The P5 primer, Read 2 (primer site for sequencing read 2), Sample Index and P7 primer were added during library construction. The protocol is designed to support library construction from a wide range of cDNA amplification yields spanning from 2 ng to 2 μg without modification. All constructed single-cell RNA-seq libraries were sequenced on the Illumina Novaseq6000 platform with a sequencing depth of at least 100,000 reads per cell for a 150bp paired end run.

Pre-processing of single-cell RNA-seq data

We demultiplexed the cellular barcodes and aligned reads to the human transcriptome (GRCh38/hg38) using Cell Ranger 3.0 (https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger). To create Cell Ranger-compatible reference genomes, the references were rebuilt according to instructions from 10x Genomics (https://www.10xgenomics.com), which performs alignment, filtering, barcode counting and unique molecular identifier (UMI) counting. Next, a digital gene expression matrix (gene counts versus cells) was generated and converted to a Seurat object by the Seurat R package [28]. For quality control, we removed the cells with <200 genes or >5,000 genes detected, or cells where >15% of the transcripts were attributed to mitochondrial genes based on the quality control metrics (Supplementary Figure 1B) and criteria used by previous scRNA-seq studies [87–89]. We normalized the filtered gene expression matrix by the NormalizeData function in Seurat R package, in which the number of UMIs of each gene was divided by the sum of the total UMIs per cell, multiplied by 10,000, and then transformed to log-scale.

Dimension reduction and osteoblastic clusters identification

For data visualization and classification, we projected the normalized gene expression matrix on a two-dimensional panel. The 2,000 genes with the highest dispersion (variance/mean) were selected for principal component analysis (PCA). The first 18 principal components (number of PCs was chosen based on standard deviations of the principal components, corresponding to the plateau region of "elbow plot") were used for uniform manifold approximation and projection (UMAP) [23] dimension reduction. After data visualization, we applied an unbiased graph-based clustering method [90] with the resolution of 0.1 for clustering analysis. To identify the differentially expressed genes (DEGs) in each cluster, we used the Wilcoxon Rank-Sum test to find the genes showing significantly higher levels of expression (false discovery rate (FDR) < 0.05) in a specific cluster compared to the other clusters. After osteoblast identification, we further removed osteoblasts where >5% of the transcripts were attributed to mitochondrial genes. 2,000 genes with the highest dispersion and the first 18 principal components of osteoblast data were chosen for the UMAP dimension reduction. We then applied the resolution of 0.15 for clustering analysis to reveal three clusters of osteoblasts.

Pathway enrichment and trajectory inference analysis

To investigate the biological processes and signaling pathways associated with each subtype, we performed gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) enrichment analysis for the genes identified as important cluster DEGs (avg_FC >1.2000 and p_val_adj < 0.05) by using the clusterProfiler R package [91]. We then used the Monocle 2 v2.8.0 [92] R package to reconstruct the single-cell developmental trajectories in pseudo-time order. The principle of this analysis is to determine the pattern of the dynamic process experienced by the cell population and to order the cells along their developmental trajectory based on differences in the expression profiles of highly variable genes. We followed the official tutorial of Monocle 2 v2.8.0 to perform the pseudotime analysis. Briefly, 1,000 with the highest dispersion were selected as the highly variable genes. Based on the highly variable genes, we then used “reduceDimension” function for the dimension reduction and “orderCells” function with default parameters for the cell ordering.

Cross-species scRNA-seq data integration

One previous scRNA-seq dataset of mouse osteoblasts was acquired from GEO database under the accession numbers of GSE108891 [17]. After acquiring the expression matrix of the osteoblasts in mice, we clustered them into three subtypes (M1, M2, and M3), following the analysis pipeline proposed by Tikhonova.
et al. [17]. Next, we integrated the scRNA-seq data of mouse and human osteoblasts by canonical correlation analysis (CCA) using Seurat R package [93]. The biological variance of transcriptional profiles across humans and mice was then evaluated based on the Spearman correlation of average gene expression between each of the datasets.

**Bone sectioning, immunostaining, and confocal imaging**

Freshly dissected femur from female C57BL/6 wild-type mouse was fixed in 4% paraformaldehyde overnight followed by decalcification in 10% EDTA for 1 week. Samples were cut in 5-µm-thick longitudinally oriented sections. After deparaffinize and antigen retrieval, sections were blocked in PBS with 5% bovine serum albumin (BSA) for 1 hour and then stained overnight with goat-anti-Alpl (R&D: AF2910-SP, 10 µg/mL). Rabbit-anti-Nur77 (Proteintech:12235-1-AP, 1:2000) was used as secondary antibodies (from Invitrogen, 1:400). Slides were mounted with anti-fade prolong gold (Invitrogen) and images were acquired with a Zeiss LSM780 confocal microscope.

**Osteogenesis induction**

Murine mesenchymal stem cells (MSCs) (1.0 × 10^4 per well) from STEMCELL (Seattle, WA, United States) were plated in 48-well plates and cultured in MesenCult™ basal expansion medium with 10% 10x Supplement (Stemcell) for 72 h. Next, the cells were rinsed with PBS and the medium was replaced with osteogenic differentiation medium (Stemcell). MSCs cultured in expansion medium were served as the negative control. Half of the medium was changed every 3 days, and cells were harvested at 0, 5, and 7 days after induction.

**qRT-PCR analysis**

Total RNA was extracted using an RNA extraction kit (Qiagen, Hilden, Germany). and cDNA was synthesized from 1 µg of total RNA by using the Revert Aid First Strand cDNA synthesis kit (Thermo). Then, the cDNA was amplified with iTaq™ Universal SYBR® Green Supermix (BioRad, Hercules, CA). in an ABI PRISM® 7900HT System (Applied Biosystems, Foster City, USA). The relative standard curve method (2−ΔΔCT) was used to determine the relative gene expression and GAPDH was used as a housekeeping gene for internal normalization. The PCR primers used in this study were as follows:

GAPDH: forward, 5’-CACCATGGAGAAGGCCGGG-3’
reverse, 5’-GACG- GACACATTGGGGGT AG-3’

ALPL: forward, 5’-CCAACCTTTTIGGCCAGA GA-3’
reverse, 5’-GGCTACATTGGGTGTGGAGC TTTT-3’

NR4A1: forward, 5’-AGGGTGTGTGTGACATAGG A-3’
reverse, 5’-CCGCCATCTTTTCCTGTAC G-3’

**Data availability**

The scRNA-seq data for primary osteoblasts from one human sample can be accessed with accession number under GSE147390 in GEO database. One previous scRNA-seq data of mice osteoblasts used in this study can be accessed with accession number of GSE108891.

**AUTHOR CONTRIBUTIONS**

YG wrote the main manuscript text and conducted major analysis; JY and LC collected the human sample and corresponding clinical information; XL, YC, and CZ performed the experiments; JG, HS, YPW, YPC and HWD did language proofreading; ZW, LJT, and YB prepared supplementary information and validated the results; the study was conceived, designed, initiated, directed, and supervised by HS, HMX, and HWD. All authors participated in the discussions of the project and reviewed and/or revised the manuscript.

**CONFLICTS OF INTEREST**

The authors declare that they have no conflicts of interest.

**FUNDING**

This research was benefitted by grants from the National Institutes of Health (R01AR069055, U19AG055373, P20GM109036, R01AG061917), National Natural Science Foundation of China (Grant No. 81902277), National Key R&D Program of China (Grant No. 2017YFC1001100), Natural Science Foundation of Hunan Province (S2019JJQNJJ2093), Changsha Science and Technology project (kq1907153), Central South University (Grant Nos. 164990007, 2018zzts886), and Xiangya Clinical Big Data Project (xyyysdsj9).

**REFERENCES**


47. Kwan Tat S, Pelletier JP, LaJeunesse D, Fahmi H, Lavigne M, Martel-Pelletier J. The differential expression of osteoprotegerin (OPG) and receptor activator of nuclear factor kappaB ligand (RANKL) in human osteoarthritic subchondral bone osteoblasts is an indicator of the metabolic state of these disease cells. Clin Exp Rheumatol. 2008; 26:295–304. PMID:18565252


72. Gagnon J, Ramanathan S, Leblanc C, Cloutier A, McDonald PP, Ilungumar S. IL-6, in synergy with IL-7 or IL-15, stimulates TCR-independent proliferation and...
https://doi.org/10.4049/jimmunol.180.12.7958
PMID:18523259

https://doi.org/10.1002/eji.201243116
PMID:2372894

https://doi.org/10.1091/mbc.E10-01-0040
PMID:20702587

https://doi.org/10.1128/IAI.73.3.1271-1274.2005
PMID:15731023

https://doi.org/10.1073/pnas.012025199
PMID:11904358

https://doi.org/10.1002/jbmr.1701
PMID:22806935

https://doi.org/10.14348/molcells.2014.0179
PMID:25220258

https://doi.org/10.1016/j.gene.2017.05.028
PMID:28502869

https://doi.org/10.3389/fcell.2020.00158
PMID:32258036

PMID:25717285

https://doi.org/10.1002/jbmr.169
PMID:20578216

https://doi.org/10.1007/s00198-005-1930-4
PMID:16025191

https://doi.org/10.1016/j.bone.2018.08.021
PMID:30172743

https://doi.org/10.1002/jbmr.5650090802
PMID:7976495

https://doi.org/10.1007/s00198-003-1517-x
PMID:14598027

https://doi.org/10.1038/s41467-020-20598-y
PMID:33441552


Supplementary Figure 1. Osteoblasts isolation and identification. (A) Representative flow cytometry of osteoblasts isolation. The x-axis represents the expression level of ALPL and the y-axis is the expression level of CD45/31/34. Dots represent the cells. Cells in Q3 panel (downer-right) are isolated for downstream analysis. (B) Number of genes per cell and the percentage of the transcripts attributed to mitochondrial genes per cell. Dash line represents the thresholds for quality control. (C) UMAP dimension reduction of isolated cells, colored by different clusters. (D) Expression of known cell markers of each isolated subpopulations embedded on UMAP dimension reduction map. The first three markers (ALPL, RUNX2 and COL1A1) in the first row are the osteogenesis markers. HBB and HBA1 are markers of two nucleated erythrocyte clusters C2 and C3. The last four genes are markers for cluster C4 and C5, respectively. (E) GO enrichment for all clusters. The size of dot indicates the gene ratio. The color indicates the adjusted p-value for enrichment analysis. (F) Heatmap of gene expression profile of isolated cells, based on the relative gene expression level of top 10 most-significant markers for each cluster. X-axis represents different osteoblast clusters, while y-axis indicates the 10 most-significant markers in each cell cluster.
Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1–6.

Supplementary Table 1. The number of genes/number of UMIs detected per cell.

Supplementary Table 2. DEGs in osteoblasts (clusters C1 and C2) compared with other contamination cells (clusters C3, C4, C5, and C6).

Supplementary Table 3. DEGs with avg.FC > 1.200 and p_val_adj < 0.05 in each osteoblast subtype compared with other subtypes.

Supplementary Table 4. Enriched GO terms for osteoblast subtypes O1, O2, and O3, respectively.

Supplementary Table 5. Enriched GO terms for osteoblast subtypes O1, O2, and O3, respectively.

Supplementary Table 6. Enriched GO terms for osteoblast subtypes O1, O2, and O3, respectively.