Single Cell Cortical Bone Transcriptomics Defines Novel Osteolineage Gene Sets Altered in Chronic Kidney Disease

Rafiou Agoro[1], Intawat Nookaew[2], Megan L. Noonan[1], Yamil G. Marambio[1], Sheng Liu[1], Wennan Chang[1,4], Hongyu Gao[1,3], Lainey M. Hibbard[1], Corinne E. Metzger[5], Daniel Horan[5], William R. Thompson[5], Xiaoling Xuei[1], Yunlong Liu[1,3], Chi Zhang[1,4], Alexander G. Robling[5], Lynda F. Bonewald[5,6], Jun Wan[1] and Kenneth E. White[1,7,*]

[1]Department of Medical and Molecular Genetics, Indianapolis, IN 46202, USA
[2]Department of Biomedical Informatics, University of Arkansas for Medical Sciences, AR 72205, USA
[3]Center for Medical Genomics, Indiana University School of Medicine, Indianapolis, IN 46202, USA
[4]Department of Electrical and Computer Engineering, Purdue University, Indianapolis, IN 46202, USA
[5]Department of Anatomy, Cell Biology and Physiology, Indianapolis, IN 46202, USA
[6]Indiana Center for Musculoskeletal Health, Indiana University, Indianapolis, IN 46202, USA
[7]Department of Medicine/Nephrology; Indiana University School of Medicine, Indianapolis, IN 46202

Running title: Osteoblast and osteocyte phenotypes at the single cell level

Keywords: Bone, scRNAseq, Osteoblast, Osteocyte, Chronic Kidney Disease

*Corresponding author information:
Kenneth E. White, Ph.D.
Indiana University School of Medicine
Department of Medical and Molecular Genetics
(Mail) 975 West Walnut St., IB130
(Office) 635 Barnhill Dr., MS5010
Indianapolis, IN 46202
317-278-1775 p
317-274-2293
kenewhit@iu.edu

Abstract

Due to a lack of spatial-temporal resolution at the single cell level, the etiologies of the bone dysfunction caused by diseases such as normal aging, osteoporosis, and the metabolic bone disease associated with chronic kidney disease (CKD) remain largely unknown. To this end, flow cytometry and scRNAseq were performed on long bone cells from Sost-cre/Ai9^{+} mice, and pure osteolineage transcriptomes were identified, including novel osteocyte-specific gene sets. Clustering analysis isolated osteoblast precursors that expressed *Tnc*, *Mmp13,* and *Spp1,* and a mature osteoblast population defined by *Smpd3*, *Col1a1*, and *Col11a1*. Osteocytes were demarcated by *Cd109*, *Ptprz1, Ramp1, Bambi, Adamts14*, *Spns2, Bmp2*, *Wasl,* and *Phex*. We validated our *in vivo* scRNAseq using integrative *in vitro* promoter occupancy via ATACseq coupled with transcriptomic analyses of a conditional, temporally differentiated MSC cell line. Further, trajectory analyses predicted osteoblast-to-osteocyte transitions via defined pathways associated with a distinct metabolic shift as determined by single-cell flux estimation analysis (scFEA). Using the adenine mouse model of CKD, at a time point prior to major skeletal alterations, we found that gene expression within all stages of the osteolineage was disturbed. In sum, distinct populations of osteoblasts/osteocytes were defined at the single cell level. Using this roadmap of gene assembly, we demonstrated unrealized molecular defects across multiple bone cell populations in a mouse model of CKD, and our collective results suggest a potentially earlier and more broad bone pathology in this disease than previously recognized.

Introduction

Osteocytes are derived from osteoblasts through a dynamic, spatio-temporal process regulated within cortical bone [1]. The exposure of these cells to the bloodstream through dendrites allows osteocytes to play key endocrine functions, including sending and receiving signals to vascularized organs such as the kidney, among others. A complete identification of the osteocyte and osteocyte precursor transcriptomes is a critical need towards understanding the molecular pathways of osteolineage differentiation, as well as to specify the contributions of osteoblast/osteocyte genes to musculoskeletal disease.

Common bone pathologies are associated with osteocyte dysfunction, including aging [2], osteoporosis [3], and chronic kidney disease (CKD) [4]. CKD is a worldwide public health problem with an estimated 5-10 million lives lost each year [5]. Endocrine changes due to loss of renal function lead to increased bone production of fibroblast growth factor-23 (FGF23), which acts on the kidney to reduce circulating 1,25$(OH)_2$ vitamin D (1,25D) concentrations. These actions in turn cause hypocalcemia-mediated compensatory elevations in PTH and subsequent bone loss due to increased osteoclastic activity [6]. Indeed, fracture is the most important clinical outcome of the CKD bone disorder, with an estimation that CKD patient hip and spine fracture rates range between 2- to 4- fold greater than the general population [7], markedly increasing patient morbidity and mortality [8].

Several studies have provided initial insight into novel bone cell bioactivity using the analysis of bulk sequencing of cortical bone [9] and single cell analyses of combined cortical bone and calvaria cells [10]. These approaches have led to an increased understanding of transitional states of the skeletal cell populations in diseases typically associated with progressive loss of bone mass, including osteoporosis and aging [11]. Additionally, the previous studies identified novel osteoblast and osteocyte genes, providing key insight into new roles of these loci for a potentially broader set of bone diseases. However, the full spectrum of osteoblast and osteocyte

genes in bone cell populations specifically affected during CKD remain understudied, hampering the ability to target cell lineages towards improved bone health and patient outcomes. Thus, the nature of osteolineage cells affected during CKD, as well as the molecular mechanisms and spatial transcriptional reprogramming that contribute to bone loss remain to be determined.

Herein, we developed a successful workflow to enrich cortical bone osteoblast and osteocyte cell populations using flow cytometry followed by single cell RNAseq (scRNAseq). Our transcriptomic data detected two osteoblast populations, as well as novel osteocyte-specific genes, and also revealed cell-specific metabolic profiles. This approach identified previously unrecognized transcriptional reprogramming prior to major bone ultrastructural changes in a mouse model of CKD. Our collective results suggest that the endocrine bone disease arising from CKD affects multiple cell populations in parallel and thus emerging treatments for CKD may be more effective by targeting a broader set of cell lineages.

Methods

Sost-CreERT2-Cre+/Ai9 mice: Animal protocols and studies were approved by the Indiana University Institutional Animal Care Committee and conform to NIH guidelines [12]. Male and female Sost-CreERT2-Cre+ mice were crossed with the 'Ai9' mouse in which the recombination event causes activation of tdTomato fluorescent protein to generate Sost-CreERT2 Cre+/Ai9 [13]. Both males and females were included in the scRNAseq study. For the sequencing experiments, with the goal to isolate cells in their native state without confounding effects on bone, due to the presence of basal Cre activation in Sost-CreERT2-Cre+ mice [13], six Sost-CreERT2/Ai9 mice at eight weeks of age were directly sacrificed without tamoxifen injection, and long bones (femurs, tibiae, and humeri) were collected and pooled, followed by a sequential bone digestion (see below).

114 **Preparation of mouse long bones for osteoblast/osteocyte isolation:** Isolated long bones
115 were placed in 100 mm petri dishes containing αMEM with 10% penicillin and streptomycin. Any
116 remaining muscle and connective tissue from the bones were removed and the periosteum was
117 scraped away using a scalpel. The bones were then washed, epiphyses removed, and the marrow
118 flushed using a 27G syringe. The flushed bones were then cut in half lengthwise and into 1 to 2
119 mm lengths using a scalpel and briefly washed with Hank's Balanced Salt Solution (HBSS;
120 Hyclone).

121 **Bone digestion and isolation of osteoblasts and osteocytes:** The long bone pieces were
122 digested as previously described [14]. In brief, bones were alternately immersed in warmed
123 collagenase type IA (2 mg/mL; Sigma) and EDTA (5 mM), then incubated at 37°C for 25 min on
124 a rotating shaker (~200 rpm). A total of 9 digestions was performed. The extracted bone cells
125 were then centrifuged at 600 rpm for 15 minutes at 4°C, washed 3 times in αMEM supplemented
126 with 10% FBS and the pellet resuspended in αMEM supplemented with 10% FBS. TdTomato^{+}
127 cells were then sorted using Fluorescence-activated cell sorting (FACS) Aria Fusion (BD
128 Biosciences) with standard Phycoerythrin (PE) Texas Red gating. After cell isolation, cell viability
129 was assessed using Trypan Blue staining; more than 90% of sorted cells were viable and these
130 cells were immediately processed for scRNAseq.

131 **CKD mouse models:** Male and female C57BL/6J mice were purchased from the Jackson
132 Laboratory and housed at least one week in the Indiana University School of Medicine Laboratory
133 Animal Resource Center (LARC) for acclimatization prior to the start of experiments, according to
134 previous protocols [15]. The adenine diet model was used to induce CKD in male and female
135 mice. At 8 weeks of age, mice were fed a control casein-based diet (0.9% phosphate and 0.6%
136 calcium, TD.150303 Envigo) for 2 weeks, or CKD-causing diet with 0.2% added adenine
137 (TD.160020; Envigo) for 2 or 4 weeks. Diets and water were provided ad libitum.

138 **Osteocyte and osteoblast cell counting in cortical and trabecular bone.** Femurs from mice
139 fed with the casein control or adenine-CKD diets for 2 or 4 weeks were fixed in 4%
140 paraformaldehyde for 24 hours. The samples were then decalcified in a solution that contains
141 1.2% neutral buffered formalin and 10% EDTA for 12-days at 4°C on a rocker platform. The
142 decalcified solution was changed three times over the decalcification period. The femurs were
143 then washed in water and stored in 70% ethanol at 4°C. For histological analysis, the femurs were
144 dehydrated overnight and embedded in paraffin. Sections of 5 µm of each sample were stained
145 with hematoxylin eosin (H&E) for light microscopy. For counting osteocytes in cortical bone, a
146 blinded experiment was conducted and an area of cortical bone of approximately ~550 mm^2 in
147 the midshaft was analyzed for cell numbers followed with a normalization to bone surface.
148 Analyses were performed using BIOQUANT imaging (BIOQUANT Image Analysis, Nashville,
149 TN). Trabecular osteocytes were analyzed and normalized to trabecular bone area in the distal
150 femur excluding endocortical surfaces and primary spongiosa. Osteoblast numbers were
151 normalized to trabecular bone surface. Three mice per condition were used for these experiments.

152 **Micro-computed tomography (µCT):** Paraformaldehyde fixed femora from CKD and healthy
153 mice were scanned, reconstructed, and analyzed as previously described [16]. Femurs were
154 scanned at 10-µm resolution, 55-kV peak tube potential and 8W. Standard output parameters
155 related to cortical bone mass, geometry, and architecture were measured as reported [17].

156 **Single cell library preparation:** We applied a single cell master mix with lysis buffer and reverse
157 transcription reagents according to the Chromium Single Cell 3' Reagent Kits V3 User Guide,
158 CG000183 Rev A (10X Genomics, Inc.). This was followed with cDNA synthesis and library
159 preparation according to standard 10X Genomics methods; all libraries were sequenced on an
160 Illumina NovaSeq6000 platform in paired-end mode (28bp + 91bp).

161 **Data processing/bioinformatic analyses:** The 10x Genomics Cellranger (v. 6.0.0) pipeline was
162 used to demultiplex raw base call files to FASTQ files and reads were aligned to the mm10 murine
163 genome using STAR [18]. The Cellranger computational output was then analyzed in R (v 4.0.2)
164 using the Seurat package v. 4.1.0 4 [19]. Seurat objects were created, and the top principal
165 components were used to perform unsupervised clustering analysis and visualized using UMAP
166 dimensionality reduction. Using the Seurat package, annotation and grouping of clusters by cell
167 type was performed manually by inspection of differentially expressed genes using the MAST
168 method [20] for each cluster, based on canonical marker genes in the literature. The selected
169 markers were visualized on UMAP coordinates as gene expression density using the R package
170 Nebulosa [21]. To perform the pseudotime analysis on the integrated Seurat object, cells were
171 divided into individual gene expression data files organized by previously defined cell types. R
172 package Monocle v3 [22] was used for dataset analysis and outputs were obtained detailing the
173 pseudotime cell distributions for each cell type. Positional information for the Monocle plot was
174 used to subset and color cells for downstream analyses [22]. Ingenuity Pathway Analysis (IPA)
175 was used to predict the statistically significant canonical pathways regulated in osteolineage cells
176 (pre-osteoblasts, osteoblasts and osteocytes).

177 **scFEA:** The python package v1.2 of single cell flux estimation analysis (scFEA) was applied to
178 estimate cell-wise metabolic flux rate against the whole human metabolic map using the
179 generated mouse scRNAseq data [23]. Default parameters were utilized and statistical
180 significance of the differences in metabolic flux between cell groups was assessed by Mann-
181 Whitney test.

182 **Culture of the MPC cell line:** A conditionally immortalized mesenchymal stem cell line, Murine
183 progenitor cells clone 2 (MPC2) [24], was cultured in αMEM (Invitrogen, Thermo-Fisher Scientific)
184 supplemented with 10% fetal bovine serum (FBS; Hyclone), 25 mM L-glutamine, and 25 mM
185 penicillin-streptomycin (Sigma-Aldrich, St. Louis, MO, USA) at 33°C and 5% CO_2 to proliferate.

186 Cells were plated at a density of $1.0x10^5$ cells per well in 6-well plates and incubated overnight
187 before being transferred to a 37°C incubator for osteogenic differentiation by culturing in
188 maintenance media supplemented with 4 mM beta-glycerophosphate and 50 μg/mL ascorbic
189 acid. Cells were differentiated for 0-4 weeks with this 'osteogenic media', which was changed
190 every 2-3 days.

191 **Bulk mRNA sequencing:** MPC2 cells were differentiated for 3 weeks in osteogenic media or
192 plated in an undifferentiated state at 33°C (see cell culture methods above). Total RNA was
193 extracted and evaluated for its quantity and quality using an Agilent Bioanalyzer 2100; 100 ng of
194 total RNA was used for the cDNA libraries. Library preparation included mRNA
195 purification/enrichment, RNA fragmentation, cDNA synthesis, ligation of index adaptors, and
196 amplification, following the KAPA mRNA Hyper Prep Kit Technical Data Sheet, KR1352 – v4.17
197 (Roche Corporate). Each resulting indexed library was quantified, and its quality accessed by
198 Qubit and Agilent Bioanalyzers; multiple libraries were pooled in equal molarity. The pooled
199 libraries were denatured and neutralized before loading on a NovaSeq 6000 sequencer at 300
200 pM final concentration for 100b paired-end sequencing (Illumina, Inc.). Approximately 30-40M
201 reads per library were generated. A Phred quality score (Q score) was used to measure the quality
202 of sequencing. More than 90% of the sequencing reads reached Q30 (99.9% base call accuracy).
203 The sequencing data were first assessed using FastQC (Babraham Bioinformatics, Cambridge,
204 UK) for quality control. The sequencing reads were mapped to the mouse genome mm10 using
205 STAR (v2.7.2a) with the following parameter: '--outSAMmapqUnique 60' [18]. Uniquely mapped
206 sequencing reads were assigned to Gencode M22 gene using featureCounts (v1.6.2) [25] with
207 the following parameters: "–p –Q 10 -O". The genes were kept for further analysis if their read
208 counts > 10 in at least 3 of the samples, followed by the normalization using TMM (trimmed mean
209 of M values) method and subjected to differential expression analysis using edgeR (v3.24.3) [26].
210 Gene Ontology and KEGG pathway functional enrichment analysis was performed on selected

211 gene sets, e.g., genes undergoing both significant differential expressions and notable changes
212 of open chromatin accessibilities, with the cut-off of false discovery rate (FDR) < 0.05 using DAVID
213 [27]. Canonical pathways from RNAseq data were generated through the use of IPA (QIAGEN
214 Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis) [28]; n=3
215 samples per condition.

216 **Assay for Transposase-Accessible Chromatin sequencing (ATACseq):** Cells tested in
217 ATACseq were plated and differentiated to osteoblast/osteocyte-like cells at the same time as
218 those for RNAseq (3 weeks, see above). After differentiation, osteoblast/osteocyte and
219 undifferentiated MSC (control) cells were washed twice in 1X PBS, then dissociated with trypsin
220 (Hyclone) for 5 minutes. Cells were resuspended in ice cold 1X PBS, dead cells were removed,
221 and live cells were processed for nuclei isolation and ATAC sequencing according to published
222 protocols [29]. Briefly, cells were collected in cold PBS and cell membranes were disrupted in
223 cold lysis buffer (10 mM Tris–HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2 and 0.1% IGEPAL CA-630).
224 The nuclei were pelleted and resuspended in Tn5 enzyme and transposase buffer (Illumina
225 Nextera® DNA library preparation kit, FC-121-1030). The Nextera libraries were amplified using
226 the Nextera® PCR master mix and KAPA biosystems HiFi hotstart readymix successively.
227 AMPure XP beads (Beckman Coulter) were used to purify the transposed DNA and the amplified
228 PCR products. The resulting ATACseq libraries were sequenced on Illumina NovaSeq 6000 and
229 paired-end 50 bp reads were generated. Illumina adapter sequences and low-quality base calls
230 were trimmed off the paired-end reads with Trim Galore v0.4.3. Bowtie2 [30] was used for
231 ATACseq read alignments on the mouse genome (mm10). Duplicated reads were removed using
232 Picard developed by the Broad Institute [https://broadinstitute.github.io/picard/ (Accessed:
233 2018/02/21; version 2.17.8)]. Low mapping quality reads and mitochondrial reads were discarded
234 in further analysis. Peak calling of mapped ATACseq reads were performed by MACS2 [31] with
235 a Bonferroni adjusted cutoff of p-value less than 0.01. Peaks called from multiple samples were

236 merged, after removing peaks overlapping with ENCODE blacklist regions [32, 33]. Reads
237 locating within merged regions in different samples were counted by pyDNase [34]. The data was
238 filtered using at least 10 cut counts in more than one of the samples, then normalized using TMM
239 (trimmed mean of M values) method and subjected to differential analysis using edgeR (v3.24.3)
240 [26, 35]. Motif enrichment of differential accessibility peaks with a false discovery rate cut-off of
241 0.05 was performed using Homer [36].

242 **RNA isolation and qPCR:** To isolate total bone RNA, one femur and one tibia per mouse were
243 harvested, the bone marrow flushed using a 27G syringe, and the epiphyses removed, similar to
244 the approach undertaken to prepare the bones prior to collagenase digestion for the scRNAseq
245 studies. The remaining bones (femur and tibia) were harvested and homogenized in 1 ml of TRIzol
246 reagent (Invitrogen) according to the manufacturer's protocol using a Bullet Blender (Next
247 Advance, Inc.), then further purified using the RNeasy Kit (Qiagen). Mouse *β-actin* was used as
248 an internal control for RT-qPCR. The qPCR primers were purchased as pre-optimized reagents
249 (Applied Biosystems/Life Technologies, Inc.) and the TaqMan One-Step RT-PCR kit was used to
250 perform all reactions. PCR conditions were: 30 minutes 48°C, 10 minutes 95°C, followed by 40
251 cycles of 15 seconds 95°C and 1 minute 60°C. The data were collected and analyzed by a
252 StepOne Plus system (Applied Biosystems/Life Technologies, Inc.). The expression levels of
253 mRNAs were calculated relative to appropriate controls, and the $2^{-\Delta\Delta CT}$ method described by Livak
254 was used to analyze the data [37]. The primers used in the study were: Fgf23, Mm00445621_m1;
255 Mmp13, Mm00439491_m1; Tnc, Mm00495662_m1; Gdpd2, Mm00469948_m1; Spp1,
256 Mm00436767_m1; Col1a1, Mm00801666_g1; Bglap, Mm03413826_mH; Cthrc1,
257 Mm01163611_m1; Smpd3, Mm00491359_m1; Dmp1, Mm01208363_m1; Sost,
258 Mm04208528_m1; Pdpn, Mm01348912_g1; Phex, Mm00448119_m1; Ptprz1, Mm00478486_m1;
259 Pdpn(E11), Mm01348912_g1; Actin, Mm02619580_g1; Aldoa, Mm00833172_g1; Adpgk,
260 Mm00511302_m1; Pgam1, Mm02526975_g1; Acadm, Mm01323360_g1; Acp5,

261 Mm00475698_m1; Mmp9, Mm00442991_m1; and Tnfrsf11a, Mm00437132_m1 (Thermo Fisher,
262 Inc).

263 **Statistical analyses:** The most recent updated R software packages with robust affiliated
264 statistics were used to analyze the scRNAseq datasets. The cutoffs used for integration of RNA-
265 seq and ATAC-seq were: peaks withing 10kb upstream of gene, differential peak cutoff FDR <
266 0.05, differential expression cutoff FDR < 0.05, $\log_2$FC > 1 (up-regulation) or < -1 (down-
267 regulation). Statistical analyses of the *in vivo* data presented were performed by two-way ANOVA
268 to assess the differences between the same gender in response to adenine diet and to assess
269 the differences between genders and treatments. Significant changes were considered when at
270 $p<0.05$.

271

272 **Results**

273 **Mouse osteoblast/osteocyte transcriptomic profiling at single cell resolution**

274 The osteocyte transcriptome at the single cell level solely from long bones remains
275 uncharacterized. To identify cortical bone cells and enrich our cell sample preparation, we used
276 Sclerostin (Sost)-Cre/Ai9 'Tomato' reporter mice at 8 weeks of age to isolate fluorescently-labeled
277 osteoblasts/osteocytes. Consistent with previous characterization [13], SOST-Cre/Ai9 ('tdTomato'
278 reporter) mice had detectable basal Cre activation in osteocytes as evidenced by the presence of
279 red fluorescence in mouse tail vertebrae (**Figure 1A**), and were thus harvested in their native
280 state to avoid confounders due to delivery of tamoxifen. Following cortical bone digestion using
281 collagenase/EDTA, the cells were sorted, and the Ai9/tdTomato$^+$ fraction represented
282 approximately 1.5% of total cells (**Figure 1B**). As determined by qPCR, the Ai9/tdTomato$^+$ cells
283 were highly enriched for osteocyte marker mRNAs *Fgf23* (40-fold), *Phex* (102-fold), *Dmp1* (172-
284 fold), and *Pdpn* (12-fold), compared to the Ai9/tdTomato$^-$ population (**Figure 1B**).

285 To identify the transcriptome of the Ai9/tdTomato^{+} population, scRNAseq was performed on this
286 isolated cell population. Approximately 1,200 cells were successfully recovered after sequencing
287 with an estimation of 130,000 reads per cell and 97% valid barcodes. Further stringent
288 bioinformatic analyses were performed to exclude low-quality cells; for downstream analysis, cells
289 that had more than 3500 genes were excluded to reduce doublet nuclei, and less than 15% of
290 mitochondrial genes to exclude potentially dead cells (**Figure S1A**). To aggregate cells based on
291 their similarities, principal component analysis (PCA) was performed on highly variable genes and
292 the result was used as input for clustering using the Louvain algorithm with multilevel refinement
293 and Uniform Manifold Approximation and Projection (UMAP) for dimension reduction (**Figure 1C**).
294 To identify gene specific markers in each cluster, we used the function FindMarkersAll from the
295 MAST method algorithm. The quality of clustering was assessed by dot plot of the most significant
296 marker genes obtained from an individual cluster (**Figure 1D**).

297 The transcriptomic profiling of Ai9/tdTomato^{+} cortical bone cells identified a total of six cell
298 populations, including two osteoblast clusters and one osteocyte cluster (**Figures 1C and 1D**).
299 The sequenced cells were grouped into three partitions: hematopoietic cells that are Cd45 (Ptprc)
300 positive, endothelial cells that highly express Cdh5, and osteolineage that express the pan
301 mesenchymal lineage cell marker Pdgfra, which was found to be present in pre-osteoblasts,
302 osteoblasts and osteocytes (**Figure S1B-D**). Marker analysis revealed that within the osteoblast
303 population, one cell type was characterized by higher expression of *Smpd3*, *Bglap, Col1a1*, and
304 *Col11a1* mRNAs. The second, referred to as "Tnc/Mmp13 osteoblasts," was defined by high
305 expression of *Tnc*, *Mmp13, Serping1 and Spp1*, which was consistent with previous
306 subpopulation analyses [10]. Osteocytes were associated with a cluster showing the highest
307 expression of *Phex* and *Dmp1*. Other genes that defined osteocytes were *Cd109*, *Dkk1*, and
308 *Ptprz1* (**Figure 1D,** and osteolineage markers listed in **Supplementary Table 1**). Our analyses
309 also identified two distinct populations of hematopoietic cells. The first population annotated

310 “Hematopoietic I” showed high expression of *Cybb*, *Chil3*, and *Ltf* whereas the “Hematopoietic II”
311 cells were characterized by expression of *Clec4d*, *Il1r2*, and *Cxcl2*. Bone marrow derived
312 endothelial cells (‘BMEC’) were also identified and distinctly defined by *Cdh5*, *Plvap*, and *Ptprb*
313 gene transcripts (**Figure 1D**). As expected, *Col1a1* was identified only in osteolineage cells
314 (**Figure 1E**), underscoring the ability to identify transcripts that may be uniquely associated with
315 ossification.

316 Our UMAP analysis confirmed that the Tnc/Mmp13 osteoblast population with high Tnc, Mmp13
317 and Spp1 expression (**Figure 1E and Figure S1E-F**) clustered farther from osteocytes when
318 compared to the dimensional separation between the defined more mature osteoblasts (high
319 *Bglap* expression (**Figure 1E**)) and osteocytes (high *Phex* expression (**Figure 1E**)). These data
320 suggest that the Tnc-Mmp13 osteoblasts likely represented a precursor osteoblast population.
321 Importantly, we isolated 25-highly and preferentially expressed genes in osteocytes, with some
322 overlap of genes in a proposed osteocyte transcriptome [9]. The polar plot comparison of recently
323 reported osteocyte transcriptome highlights the genes in common such as *Dkk1*, *Dmp1*, *Irx5,*
324 *Col24a1, Ackr3* from our dataset (scRNAseq of long bone), Wang and colleagues (scRNAseq on
325 long bone and calvaria) [10] and Youlten and colleagues (bulk RNAseq on long bone) [9] (**Figure**
326 **1F**).

327 To further test whether the genes identified as highly expressed in osteolineage cells in our
328 scRNAseq analyses were not derived from bone marrow cells, we compared the mRNA
329 expression of these candidate genes in mouse cortical long bone (osteoblast/osteocyte-enriched
330 fraction) *versus* bone marrow using qPCR. In this regard, *Mmp13*, *Cthrc1*, *Smpd3*, *Ptprz1*, and
331 *Cd109* were tested, and these genes were highly expressed in cortical bone suggesting that they
332 were preferentially expressed in osteolineage cells (**Figure 1G**). The erythropoietin receptor
333 (*Epor)* mRNA, tested as a marrow positive control, was not increased in cortical bone when
334 compared to bone marrow expression levels, whereas defined osteocyte genes *Phex, Pdpn,*

335 *Dmp1, Fgf23,* and *Sost* were all increased in cortical bone samples (**Figure 1G**). To determine
336 the cellular differentiation patterns that govern the maturation of osteoblasts to osteocytes, we
337 performed pseudotime analysis. This algorithm was first performed on the partition of cells based
338 on the UMAP (**Figure 1C**) in three main compartments (p1, osteolineage; p2, hematopoietic cells;
339 and p3, marrow endothelial cells) as indicated in **Figure S1G**. After the initial analyses, Monocle3
340 simultaneously performed pseudotime analysis over the individual partitions. Trajectory analyses
341 from the partition p1, an osteogenic lineage (**Figure 1H**)**,** hematopoietic cells (**Figure S1H**), and
342 endothelial cells (**Figure S1I**) were next generated. Interestingly, a global cellular interaction
343 pattern was conserved with almost the same UMAP projection as for all cell populations (as in
344 **Figure 1C**). These analyses confirmed that the osteocyte cluster of cells identified from our
345 dataset derived from the osteoblast subset, with the Tnc/Mmp13 osteoblast as a likely precursor
346 population (**Figure 1H**).

347

348 **scRNAseq identifies pathways involved in the osteoblast to osteocyte transition**

349 To identify the molecular pathways involved in the transition from osteoblasts to osteocytes at
350 single cell resolution, we analyzed differentially expressed osteoblast/osteocyte genes using
351 Ingenuity Pathway Analysis (IPA). Further, a comparison analysis was performed to identify
352 signaling pathways in osteoblasts *versus* osteocytes. In this regard, we found an enrichment of
353 "*GP6 Signaling*", a major signaling receptor for collagen in osteoblasts, which was almost
354 completely shut down in osteocytes (**Figure 2A**). In contrast, the "*Axonal guidance*", the
355 "*Differentiation via BMP receptors*", "*TGFβ signaling*" and the "*Iron homeostasis signaling*"
356 pathways that were not highly enriched in osteoblasts were significantly increased in osteocytes
357 (**Figure 2A**). Consistent with osteocyte morphology, among the genes predicted to drive "*Axonal*
358 *guidance*" were *Adamts14*, *Bmp2*, *Bmp4*, and *Wasl*. Furthermore, IPA analysis predicted EGFR
359 among the top upstream regulators of osteoblastic pathways whereas FGF2 was predicted to be

360 a primary upstream regulator associated with the osteocyte maturation pathway (**Figure S2A**).
361 Additionally, our in-silico analysis identified an enriched cell activation in osteoblasts, associated
362 with genes such as *Col11a1*, *Col11a2*, and *Col1a1* which were significantly attenuated in
363 osteocytes (**Figure 2A**). Overall, these data support that osteoblasts and osteocytes can be
364 defined by unique pathways associated with their individual homeostatic functions.

365 To predict the transcriptional profiles of osteocyte genes across the osteolineage, we used
366 pseudotime analysis to track the computational-derived kinetics of gene expression during
367 differentiation. We found that "pan-osteolineage" markers *Col1a1*, *Col1a2* and *Runx2* showed
368 relatively stable expression during osteoblast to osteocyte differentiation (**Figure 2B**). In contrast,
369 *Mmp13, Tnc,* and *Spp1* pseudotemporal progression were predicted to decrease during
370 differentiation from Tnc/Mmp13 osteoblasts to osteocytes (**Figure 2C**). Further, *Bglap, Col11a2,*
371 and *Smpd3* expression increased from the Tnc/Mmp13 precursor osteoblasts to osteoblasts, but
372 decreased when the osteoblast population transitioned to osteocytes (**Figure 2D**). In osteocytes,
373 the kinetic profiling of *Dmp1, Irx5, Ackr3, Phex, Ptprz1,* and *Cd109* showed progressive positive
374 regulation during differentiation (**Figure 2E, Figure S2B,** and **Supplementary Table 1**). These
375 data confirmed the dynamic differentiation of canonical genes during the transcriptional
376 reprogramming required for osteoblast transition to osteocytes, as well as identified new genes
377 associated with the cortical bone osteolineage.

378 To distinguish the metabolic heterogeneity between osteoblasts and osteocytes, we applied the
379 recently developed computational method single-cell flux estimation analysis (scFEA) [23] that
380 uses a systematically reconstructed human metabolic map to infer the cell-wise cascade from the
381 transcriptome to metabolome. This approach applies multilayer neural networks to capitulate the
382 nonlinear dependency between enzymatic gene expression and reaction rate fluxome from
383 scRNAseq datasets. By reconducting the cell clustering analysis using the original UMAP labeling
384 (**Figure 1C**), scFEA identified high metabolic reactions including glycolytic, TCA, and fatty acid

385 metabolic reactions in Tnc/Mmp13 osteoblasts (**Figure S2C-S2E**) consistent with an overall
386 increased cell activation (**Figure 2A**). The most prominent change of metabolic route predicted
387 was the conversion of Fatty acid to Acetyl co-A during the transition from pre-osteoblasts to
388 osteoblasts/osteocytes with two distinct cell populations (**Figure 2F**). The gene *Acadm* (Acyl-CoA
389 Dehydrogenase Medium Chain) was the top predicted gene to drive this process. Whereas the
390 conversion of serine to cysteine decreased with osteoblast differentiation, the transformation of
391 cysteine to pyruvate increased when Tnc/Mmp13 osteoblasts transitioned to
392 osteoblasts/osteocytes, a potential alternative metabolic route to supply cells with pyruvate
393 (**Figure 2G**). This cysteine metabolism change predicted by scFEA correlated with differential
394 regulation of *Cdo1* (Cysteine Dioxygenase Type 1), a gene that was highly expressed in
395 osteoblasts (**Figure 2H**). Based on the scRNAseq dataset, the top predicted metabolites in
396 osteolineage cells were the inosine monophosphate (IMP), succinyl coenzyme A, and pyrimidine
397 in pre-osteoblasts; the farnesyl pyrophosphate (FPP), cytidine-5'-diphosphate (CDP), and malate
398 in osteoblasts; whereas the xanthosine monophosphate (XMP), acetyl CoA oxaloacetate, and
399 malate were enriched in osteocytes (**Figure 2I**). Collectively, the transcriptional profile across
400 cortical bone cell types predicted differential metabolic states between the osteoblast cell
401 population subsets and osteocytes.

402

403 **Identification of genes during staged osteolineage differentiation**

404 To validate the osteoblast/osteocyte genes identified from our *in vivo* scRNAseq experiments, the
405 mRNA expression of these genes was tested in mesenchymal progenitor cell ('MPC2') cells *in*
406 *vitro*. MPC2 cells were recently characterized as harboring the ability to derive osteoblast- and
407 osteocyte-like cells when cultured in osteogenic media [24]. After 1-4 weeks of culture, the
408 differentiated cells can be stained by alizarin red reflecting the ability of the mature cells to

409 mineralize [24]. Our *in vitro* molecular analyses demonstrated temporal changes of canonical and
410 non-canonical osteolineage genes. *Col1a1* mRNA increased during cell differentiation with the
411 highest level observed at 4 weeks. *Runx2* gene expression reached an early and maximal
412 expression 1-week after differentiation and remained stable until 4 weeks. The expression of Tnc
413 significantly increased 1-week after differentiation, although *Cthrc1,* identified as an enriched
414 osteoblast gene based upon our scRNAseq dataset began increasing with MPC2 cell
415 differentiation at 3 weeks (maximal time point) before being significantly downregulated at 4
416 weeks (**Figure 3A**). Osteocyte genes, including *Pdpn*, *Phex*, *Sost* and *Ptprz1,* had high
417 expression at 3 and 4 weeks of osteocyte differentiation, confirming our finding of *Ptprz1* as a
418 likely late osteoblast/osteocyte gene. Further, genes associated with predictive glycolytic and β-
419 oxidation pathways identified by scFEA from the *in vivo* studies (**Figure 2F**) were assessed during
420 *in vitro* MPC2-osteoblast differentiation. We found that glycolytic genes increased during
421 osteoblast differentiation including *Adpgk* (ADP-Dependent Glucokinase)*, Aldoa* (Aldolase A)*,*
422 and *Pgam1* (Phosphoglycerate Mutase 1) reaching ~3-fold increases at 3 at 4 weeks. In contrast,
423 *Acadm* (Acyl-CoA Dehydrogenase Medium Chain), a gene associated with the conversion of fatty
424 acid to acetyl co-A remained stable during the first three weeks of differentiation before being
425 downregulated at 4 weeks (**Figure 3A**), consistent with the predicted osteocyte metabolic
426 transition *in vivo* (**Figure 2F**).

427 To identify the molecular events associated with osteolineage differentiation, a comprehensive
428 genomic and transcriptomic analyses using RNAseq and ATACseq was employed using MPC2
429 cells [38]. Based on the kinetic profile performed in **Figure 3A**, a time point of 3 weeks of MSC to
430 osteocyte-like cell differentiation was chosen for RNAseq and ATACseq to capture differentially
431 regulated genes in osteoblasts and osteocytes. Our unbiased approach identified several genes
432 with increased chromatin accessibility and/or gene expression during MPC2 differentiation
433 including osteocalcin (*Bglap*), osteopontin *(Spp1), Col11a1,* and *Col11a2* (**Figure 3B** and **Figure**

434 **S3A-F**). By performing integrative and correlative analyses of the ATACseq and RNAseq, we
435 confirmed our observed changes in mRNA expression and chromatin accessibility in the ATAC-
436 seq dataset. Using this approach, 150 genes were identified with significantly more open
437 chromatin status and higher gene expression levels after differentiation (**Figure 3C**), whereas 44
438 genes were downregulated with notably less DNA accessibility (**Figure 3C**). The gene sets with
439 both upregulated differential chromatin accessible regions (DAR) and differential expressed
440 genes (DGE), as well as downregulated DAR and DE showed a significantly high fold-enrichment
441 compared to the random selections. However, all other combinations did not demonstrate
442 significance *versus* the random selections (**Figure 3C**). Further, a significant Jaccard index score
443 was noted when gene expression and chromatin accessibility increased or decreased
444 simultaneously (FDR < 0.05, see **Supplementary Table 2**), suggesting that the chromatin
445 changes reflected by DARs, tended to be positively correlated with the DEGs. HOMER
446 transcription factor motif analysis of the ATACseq data detected an enrichment of several motifs
447 that are known to influence bone differentiation including an increased enrichment of Runx1,
448 Runx2, Foxo1, and DLX1/5/2 (**Figure 3D**).

449 The upregulated genes identified from differential expression analysis associated with changes
450 in chromatin accessibility (**Figure 3D**) were then subdivided by gene ontology (GO). Among the
451 GO terms that were increased during osteoblast/osteocyte differentiation were functions related
452 to bone mineralization, as well as cellular morphology changes including membrane branching
453 (**Figure 3E**). Genes that reflected pathways associated with “Cell Morphology” and “Ossification”
454 are shown in **Figure 3F**. Importantly, osteocyte genes identified from the scRNAseq analysis
455 including *Ccn4, Adamts14, Spns2,* and *Bmp2* were significantly upregulated in osteocyte-like
456 MPC2 cells when compared with parent undifferentiated MSC cells (**Figure 3F**). *Ptgis* identified
457 in our dataset and the Wang, et al dataset [10] was also significantly increased with differentiation.
458 *Col24a1* found in our dataset and the Youlten, et al dataset [9] was also increased (**Figure 3F**).

459 These mRNA expression levels detected during the cell transitions were robustly and positively
460 associated with their corresponding genomic changes (**Figure 3G**). For example, the DNA
461 promoter regions at *Bglap*, *Col8a1*, *Mgp*, *Igfbp5*, *Sema3a* and *Tnc* genes were more accessible
462 with cell maturation (**Figure 3G**).

463 Next, we tested the chromatin accessibility of a set of osteolineage genes identified in the
464 scRNAseq. For instance, among the Tnc/Mmp13 osteoblast signature genes, we detected
465 increased chromatin accessibility at the *Mmp13*, *Serpine2,* and *Lifr* promoter regions in
466 differentiated cells versus MSC control cells (**Figure S4A-C**). We also observed high chromatin
467 accessibility at the promoters of *Serpinf1* and *Cthrc1* (**Figure S4D-E**), that were defined as
468 expressed by fully differentiated osteoblasts based upon our *in vivo* scRNAseq data (see **Figure**
469 **1D**). For osteocyte genes, increased chromatin accessibility for the promoter regions of *Ackr3,*
470 *Cd109, Ptgis, Spns2,* and *Bmp2* was observed (**Figure S4F-J**). Taken together, these results
471 identified genomic regions of interest associated with osteoblast differentiation that may play
472 critical roles in cellular transcriptional reprogramming.

473

474 **Impairment of osteoblast/osteocyte genes prior to cortical bone deterioration in CKD**

475 We next assessed the dynamic transcriptional profile of the identified osteoblast and osteocyte
476 genes within cortical bone prior to the development of cortical porosity in CKD, a severe outcome
477 that leads to bone fracture. We used a mouse model in which CKD is induced by providing an
478 adenine-containing diet, which causes a progressive loss of kidney function [39]. This model
479 reflects the phenotype observed in humans, characterized by progressive tubular atrophy,
480 immune cell infiltration, and fibrosis [39], and has been extensively characterized as a reliable
481 model to investigate CKD-mineral and bone disorder [15]. In comparison to healthy controls,
482 previous studies showed that mice fed a 0.2% adenine diet exhibited biochemistries typical to

483 CKD such as hyperphosphatemia, hypocalcemia, secondary hyperparathyroidism, increased
484 blood urea nitrogen (BUN), and FGF23 induction, which are enhanced with the duration of
485 treatment, as well as associated with bone porosity during chronic adenine administration [15].

486 With a focus on testing osteoblast/osteocyte gene regulation in cortical bone at a timepoint prior
487 to major ultrastructural changes, cohorts of C57BL/6 mice were placed on control diet or a 0.2%
488 adenine-containing diet for 2 or 4 weeks. At 2 weeks, serum phosphate, alkaline phosphatase,
489 and calcium were not significantly changed compared to controls (**Table 1**). Blood urea nitrogen
490 (BUN) was monitored for declines in renal function, and as expected, the mice receiving the
491 adenine diet had significantly increased BUN (**Table 1**). The cortical porosity and bone volume
492 remained unchanged at 2 or 4 weeks in male or female CKD mice (**Figure 4C**). However, plasma
493 bioactive iFGF23 concentrations were markedly elevated in the mice with CKD (**Figure 4D**).

494 We next determined the effects of CKD on highly expressed osteoblast and osteocyte genes in
495 cortical bone samples. During early CKD, *Mmp13*, and *Spp1* identified in the osteoblast precursor
496 population, remained stable or increased with CKD (**Figure 4E**). In contrast, other precursor
497 osteoblast genes such as *Tnc,* and *Gdpd2,* were decreased during CKD progression. The
498 expression of the osteolineage gene *Col1a1*, as well as osteoblast-specific genes *Bglap, Cthrc1,*
499 *Smpd3* and osteoblast/osteocyte marker *Dmp1*, were strongly impaired in the mice with CKD
500 (**Figure 4G**). Increased circulating FGF23 was associated with significant upregulation of cortical
501 bone *Fgf23* mRNA at 4 weeks (**Figure 4G**). In a similar pattern to that of osteoblast genes,
502 osteocyte genes including *Sost*, *Phex,* and *Ptprz1* were dramatically downregulated early in CKD
503 (**Figure 4G**). To test whether the change of osteocyte genes in early CKD resulted from an overall
504 decrease of cell numbers, the osteocytes were counted in cortical and trabecular bone using
505 histological analysis. Osteocytes were found within lacunae, and their cell numbers were not
506 different between controls and mice with CKD (**Figure S5A-B**), suggesting that the observed
507 changes in gene expression in cortical bone were not due to the decrease of overall cell numbers

508 but rather a profound change at cellular level. Of note, empty lacunae were not observed. The
509 osteocyte numbers in trabecular bones were also unchanged, however the number of osteoblasts
510 in trabecular bone was increased in the mice with CKD (**Figure S5A-B**).

511 To assess whether changes in bone forming genes at CKD onset occur concurrently with changes
512 in osteoclast markers, we measured gene expression of *Acp5* (Tartrate-resistant acid
513 phosphatase), *Mmp9* (Matrix metallopeptidase 9), and *Tnfrsf11a* (Rank) by qPCR. At 2 weeks, a
514 time point where serum phosphate remained unchanged and the cortical bone unaltered as tested
515 by µCT, the mRNA expression of these osteoclast markers was not different from controls (**Figure**
516 **4H**). However, by 4 weeks, these genes were upregulated, consistent with an onset of metabolic
517 bone disease. These findings suggest that at the early stage of the disease, the misregulation of
518 bone-forming osteolineage genes may be independent from loss of cortical bone or increased
519 osteoclast function.

520 In sum, this work demonstrated that scRNAseq can identify sub-populations of osteoblasts, and
521 together with metabolic profiling, predict differentiation to mature osteoblasts and osteocytes. In
522 addition, we identified genes not previously associated with cortical bone expression, and
523 confirmed their presence in osteoblast/osteocytes via corresponding changes in genomic
524 accessibility and expression using *in vitro* differentiation studies. These genes were then shown
525 to be misregulated in a mouse model of CKD. Our collective findings support that the molecular
526 changes observed in cortical bone associated with the severe musculoskeletal phenotypes in
527 CKD may occur more rapidly than recognized, and can now be pinpointed to specific cell types
528 within the osteolineage.

529 **Discussion**

530 There is an urgent need to understand bone-forming cell heterogeneity as well as isolate key
531 gene networks that impact skeletal homeostasis. This is especially apparent in diseases such as
532 CKD, where due to the progressive nature and specific calcium- and phosphate-related endocrine
533 disturbances, there are currently very little treatment options [40, 41]. Our cortical bone single cell
534 RNAseq analysis identified two osteoblast populations as well as a distinct osteocyte cluster. The
535 genes demarcating each cell type were confirmed as highly enriched in cortical bone as well as
536 being upregulated during osteoblast differentiation *in vitro*. Interestingly and unexpectedly, at
537 early-stage CKD onset in mice, there was a dramatic misregulation of gene expression in cortical
538 bone cell types prior to major bone ultrastructural changes.

539 Recent studies using bulk RNA sequencing developed an osteocyte-defining transcriptome which
540 was comprised of 28 genes [9], and we observed partial overlap of genes within the
541 transcriptomes reported by Youlten et al [9], as well as from analysis we performed using the
542 publicly available scRNAseq dataset of Wang et al [10]. Our scRNAseq analysis also identified
543 additional genes which were highly enriched in osteocytes (see **Figure 1F**). It is likely that some
544 of the observed differences in osteocyte gene profiling between our studies and those of others
545 are due to the fact that we exclusively used long bones for cell isolation whereas the previous
546 studies sequenced a mix of long bone and calvariae, or used bulk total cortical bone RNA as a
547 starting material [9, 10]. Future studies will be needed to further sub-set these gene profiles and
548 expand cell-specific targets as more cortical bone datasets become publicly available for multi-
549 study integration. This is especially the case for osteolineage genes in light of the idea that these
550 cells are derived from precursors that have distinct and overlapping gene expression.

551 Our bone scRNAseq analyses identified a population of Tnc/Mmp13 mRNA-containing
552 osteoblasts as potential precursors of a mature osteoblast population of cells. Mmp13, the most
553 highly expressed gene in this cell type has been described as a direct target of osteoblast-specific

554 transcription factor osterix (Sp7) in osteoblasts [42] and known to be activated by Runx2, the
555 master transcription factor for osteoblastogenesis [43]. A recent study defined the molecular
556 mechanisms driven by Sp7 to control osteocyte dendrite formation [10]. Whether directly targeting
557 Mmp13 could influence osteoblast differentiation and osteocyte maturation remains to be
558 investigated. Further analyses integrating our bone cell data sets with other published scRNAseq
559 data by including stem cells and osteoprogenitors may complement the work described herein to
560 predict a full osteogenic trajectory, and to precisely position the Tnc-Mmp13 osteoblast population
561 in the osteogenic commitment lineage. Although we identified osteolineage genes in common
562 with previous studies [9, 10] in addition to unique markers, we cannot exclude that using a system
563 of basal state SOST-Cre/Ai9 and cell sorting (with no tamoxifen induction) identified a specific set
564 of cortical bone cells. It will be important to continue to characterize additional bone cell
565 populations identified through multiple approaches. As new tools such as mouse reporter genes
566 for studying osteoblasts/osteocytes continue to emerge in concert with evolving approaches for
567 testing the single cell transcriptome and genomic accessibility with more sensitivity, it will be
568 important to continue to refine bone cell classification to fully understand osteolineage
569 heterogeneity.

570 Our work in modeled CKD identified a general misregulation of mature osteoblast and osteocyte
571 genes prior to major bone structural changes. This analysis detected a downregulation of multiple
572 osteoblast and osteocyte mRNAs except for bone *Fgf23* that increased in CKD, or remained
573 stable or were moderately induced, such as *Mmp13* and *Spp1*. The pre-osteoblast gene sets
574 tested in CKD suggested a step-wise decrease in progenitor alterations. Further, our histological
575 analyses found that the number of osteocytes in cortical and trabecular bone are unchanged at
576 the early stage of modeled CKD, suggesting that diminishing cell numbers with disease
577 progression was not responsible for the altered gene expression (**Figure S5A-B**). These fundings
578 are consistent with previous studies that quantified osteocytes in healthy controls, pre-dialysis

579 CKD patients, and pediatric dialysis patients. This report found that the osteocyte cell numbers in
580 bone biopsies did not differ between CKD patients when compared to the healthy group [44]. In
581 support of our findings, it is possible that the maturation mechanisms of osteocytes may instead
582 be affected [44]. Thus, the genes tested herein may serve as early onset markers for CKD bone
583 disease and set the stage for studies that could determine their function during osteolineage
584 differentiation in CKD. These studies also imply that targeting one bone cell type in CKD bone
585 disease will likely not be fully effective since damage may be done in early precursor cells, with
586 these cells potentially carrying these gene expression changes through differentiation where they
587 may impact mature cells. We also found a decrease in cortical bone Sclerostin (*Sost*) expression
588 at an early stage of CKD in mice (**Figure 4G**). In a paracrine signaling manner, rapid decreases
589 of osteocyte genes *Sost* and likely *Dkk1* may represent an adaptive mechanism to maintain
590 effective Wnt signaling as an attempt to delay the initial development of renal osteodystrophy.
591 Whether the decrease in cortical bone *Sost* stimulates trabecular osteoblast differentiation in a
592 paracrine manner to attempt to sustain bone formation at CKD onset remains to be determined.

593 It was demonstrated in previous studies that mice fed an adenine-containing diet over a chronic
594 10-week period developed deficits in cortical bone properties associated with an increase of
595 osteoclast surface and a higher bone formation rate [45]. Another adenine study conducted over
596 a 56-day protocol found that the bone resorption marker carboxy-terminal collagen crosslinks
597 (CTX) was decreased and the bone formation marker the N-terminal propeptide of type I
598 procollagen (P1NP) had a trend towards increasing (p=0.06) [46]. These chronic CKD models
599 suggested a high bone turnover phenotype when mice were fed with an adenine diet over a long-
600 term period. In our studies, mice exposed to adenine diet for only 2 weeks did not exhibit cortical
601 porosity, and osteoclast gene markers were unchanged at 2 weeks of dietary treatment although
602 osteoblast and osteocyte markers assessed in cortical bone were already downregulated (**Figure**

603 **4F-H**). These results suggest that misregulation of genes within the osteolineage may precede
604 the up-regulation osteoclast genes.

605 The skeleton is an important regulator of systemic glucose homeostasis, with osteocalcin and
606 insulin thought to represent prime mediators of the interplay between bone and energy
607 metabolism [47, 48]. Using the recently developed computational tool single-cell flux estimation
608 analysis (scFEA) [23] that calculates the cell metabolomic state in scRNAseq datasets, our data
609 predicted *in vivo* metabolic transitions during osteoblast differentiation. At the single cell level, our
610 data supported high glycolysis and glutamine metabolic states in osteoblast precursors versus
611 more mature osteoblasts and osteocytes (**Figure 2F** and **Figure S2G**). The metabolic flux
612 analysis suggested that metabolomic transition between pre-osteoblasts to osteoblasts was more
613 pronounced than between osteoblasts to osteocytes. This was correlated with the extent of
614 differentially expressed genes between osteolineage cells. Indeed, 198 genes were differentially
615 expressed when comparing pre-osteoblasts and osteoblasts *versus* 77 genes between
616 osteoblasts and osteocytes (logFC>0.5 and $p<0.05$; **Figure S2G**). The conversion of arginine to
617 ornithine was similar across all osteolineage cells (**Figure S2H**), and pathways including the
618 conversion of citrulline/aspartate to arginosuccinate or acetyl glucosamine to hyaluronic acid were
619 also predicted to be conserved regardless of cell differentiation status or cell type (**Figure S2H**).
620 In our analyses, we provided validation in a distinct cell system for the regulation of key genes
621 associated with these metabolic flux predictive data. However, it will be important to build upon
622 these findings to provide genetic evidence that specific osteoblast precursor genes assorted with
623 metabolism may be critical in supporting osteoblast differentiation, as well as influencing global
624 energy metabolism in normal states and in skeletal disease.

625 Although this study provided an extensive gene profile analysis of osteolineage cells, some
626 established osteocyte genes such as *Sost* and *Fgf23* were not detected under baseline
627 conditions. It is likely that current single cell sequencing technologies using the short coverage

628 length along each mRNA from the 3'UTR end (which is based on 28 bp of cell barcode and UMI
629 sequences and 91 bp RNA reads generated) have sensitivity limitations for detecting low, yet
630 specific gene expression in less prevalent cell populations. Overall, genes with lower expression
631 (mainly with a detection at the latest quarter threshold cycle using real time qPCR) were minimally
632 detected at single cell resolution using the current 10X Genomics expression profiling pipeline.
633 However, our approach of combining FACS sorting with cell isolation from the Sost-Cre/Ai9 mice
634 to develop an enriched population of cortical bone cells identified potentially new sets of osteocyte
635 genes which were validated in an independent mesenchymal cell line *in vitro* and in RNA from
636 cortical bone *ex vivo*.

637 In summary, by combining genomic, transcriptomic, and predictive metabolic profiling
638 approaches, we identified osteolineage genes associated with distinct cell populations of
639 osteoblast precursors, mature osteoblasts and osteocytes. Genes within these three cortical bone
640 cell populations were misregulated in a mouse model of CKD prior to the development of cortical
641 porosity. Thus, our findings support that the molecular events occurring during the bone disease
642 associated with CKD appear early and manifest more widely across the osteolineage cell
643 population.

644 **Acknowledgements**

645 The authors thank the members of the Indiana University Melvin and Bren Simon Cancer Center
646 Flow Cytometry Resource Facility for their outstanding technical support. This work was
647 supported in part by NIH. The Indiana University Melvin and Bren Simon Comprehensive Cancer
648 Center Flow Cytometry Resource Facility is funded in part by NIH, National Cancer Institute (NCI)
649 grant P30 CA082709 and National Institute of Diabetes and Digestive and Kidney Diseases
650 (NIDDK) grant U54 DK106846. The FCRF is supported in part by NIH instrumentation grant
651 1S10D012270. The authors would like to acknowledge NIH grants K99-DK129705 (RA), R21-
652 AR059278 (KEW), R01-DK112958 (KEW), P20GM125503 (NIGMS to IN), and the David Weaver
653 Professorship (KEW).

654

655 **Disclosures**

656 KEW receives royalties for licensing FGF23 to Kyowa Hakko Kirin Co., Ltd; had previous funding
657 from Akebia, and current funding from Calico Labs; KEW also owns equity interest in FGF
658 Therapeutics. The other authors have nothing to declare.

659

660 **Data availability:**

661 The scRNAseq data have been deposited into the GEO under accession number GSE208152.
662 Raw data for MPC2 cell RNAseq and ATACseq were deposited into the GEO under accession
663 number GSE205792; the data analyzed herein were from undifferentiated (MSCs) and
664 differentiated osteoblasts/osteocytes (OB/OC).

References

1. Dallas, S.L. and L.F. Bonewald, *Dynamics of the transition from osteoblast to osteocyte.* Ann N Y Acad Sci, 2010. **1192**: p. 437-43.
2. Martin, B., *Aging and strength of bone as a structural material.* Calcif Tissue Int, 1993. **53 Suppl 1**: p. S34-9; discussion S39-40.
3. Seeman, E., *Age- and menopause-related bone loss compromise cortical and trabecular microstructure.* J Gerontol A Biol Sci Med Sci, 2013. **68**(10): p. 1218-25.
4. Nickolas, T.L., et al., *Rapid cortical bone loss in patients with chronic kidney disease.* J Bone Miner Res, 2013. **28**(8): p. 1811-20.
5. Luyckx, V.A., M. Tonelli, and J.W. Stanifer, *The global burden of kidney disease and the sustainable development goals.* Bull World Health Organ, 2018. **96**(6): p. 414-422D.
6. John, G.B., C.Y. Cheng, and M. Kuro-o, *Role of Klotho in aging, phosphate metabolism, and CKD.* Am J Kidney Dis, 2011. **58**(1): p. 127-34.
7. Nitsch, D., et al., *Chronic kidney disease and hip fracture-related mortality in older people in the UK.* Nephrol Dial Transplant, 2009. **24**(5): p. 1539-44.
8. Robertson, L., et al., *Hip fracture incidence and mortality in chronic kidney disease: the GLOMMS-II record linkage cohort study.* BMJ Open, 2018. **8**(4): p. e020312.
9. Youlten, S.E., et al., *Osteocyte transcriptome mapping identifies a molecular landscape controlling skeletal homeostasis and susceptibility to skeletal disease.* Nat Commun, 2021. **12**(1): p. 2444.
10. Wang, J.S., et al., *Control of osteocyte dendrite formation by Sp7 and its target gene osteocrin.* Nat Commun, 2021. **12**(1): p. 6271.
11. Lane, N.E., *Epidemiology, etiology, and diagnosis of osteoporosis.* Am J Obstet Gynecol, 2006. **194**(2 Suppl): p. S3-11.
12. in *Guide for the Care and Use of Laboratory Animals*, th, Editor. 2011: Washington (DC).
13. Maurel, D.B., et al., *Characterization of a novel murine Sost ER(T2) Cre model targeting osteocytes.* Bone Res, 2019. **7**: p. 6.
14. Stern, A.R., et al., *Isolation and culture of primary osteocytes from the long bones of skeletally mature and aged mice.* Biotechniques, 2012. **52**(6): p. 361-73.
15. Clinkenbeard, E.L., et al., *Increased FGF23 protects against detrimental cardio-renal consequences during elevated blood phosphate in CKD.* JCI Insight, 2019. **4**(4).
16. Kedlaya, R., et al., *Sclerostin inhibition reverses skeletal fragility in an Lrp5-deficient mouse model of OPPG syndrome.* Sci Transl Med, 2013. **5**(211): p. 211ra158.
17. Bouxsein, M.L., et al., *Guidelines for assessment of bone microstructure in rodents using micro-computed tomography.* J Bone Miner Res, 2010. **25**(7): p. 1468-86.
18. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.
19. Stuart, T., et al., *Comprehensive Integration of Single-Cell Data.* Cell, 2019. **177**(7): p. 1888-1902 e21.
20. Finak, G., et al., *MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data.* Genome Biol, 2015. **16**: p. 278.
21. Alquicira-Hernandez, J. and J.E. Powell, *Nebulosa recovers single cell gene expression signals by kernel density estimation.* Bioinformatics, 2021.
22. Trapnell, C., et al., *The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells.* Nat Biotechnol, 2014. **32**(4): p. 381-386.
23. Alghamdi, N., et al., *A graph neural network model to estimate cell-wise metabolic flux using single-cell RNA-seq data.* Genome Res, 2021. **31**(10): p. 1867-1884.

710 24. Prideaux, M., et al., *Generation of two multipotent mesenchymal progenitor cell lines capable of*
711 *osteogenic, mature osteocyte, adipogenic, and chondrogenic differentiation.* Sci Rep, 2021. **11**(1):
712 p. 22593.
713 25. Liao, Y., G.K. Smyth, and W. Shi, *featureCounts: an efficient general purpose program for assigning*
714 *sequence reads to genomic features.* Bioinformatics, 2014. **30**(7): p. 923-30.
715 26. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential*
716 *expression analysis of digital gene expression data.* Bioinformatics, 2010. **26**(1): p. 139-40.
717 27. Dennis, G., Jr., et al., *DAVID: Database for Annotation, Visualization, and Integrated Discovery.*
718 Genome Biol, 2003. **4**(5): p. P3.
719 28. Kramer, A., et al., *Causal analysis approaches in Ingenuity Pathway Analysis.* Bioinformatics, 2014.
720 **30**(4): p. 523-30.
721 29. Buenrostro, J.D., et al., *Transposition of native chromatin for fast and sensitive epigenomic*
722 *profiling of open chromatin, DNA-binding proteins and nucleosome position.* Nat Methods, 2013.
723 **10**(12): p. 1213-8.
724 30. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nat Methods, 2012.
725 **9**(4): p. 357-9.
726 31. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS).* Genome Biol, 2008. **9**(9): p. R137.
727 32. Consortium, E.P., *An integrated encyclopedia of DNA elements in the human genome.* Nature,
728 2012. **489**(7414): p. 57-74.
729 33. Amemiya, H.M., A. Kundaje, and A.P. Boyle, *The ENCODE Blacklist: Identification of Problematic*
730 *Regions of the Genome.* Sci Rep, 2019. **9**(1): p. 9354.
731 34. Piper, J., et al., *Wellington: a novel method for the accurate identification of digital genomic*
732 *footprints from DNase-seq data.* Nucleic Acids Res, 2013. **41**(21): p. e201.
733 35. McCarthy, D.J., Y. Chen, and G.K. Smyth, *Differential expression analysis of multifactor RNA-Seq*
734 *experiments with respect to biological variation.* Nucleic Acids Res, 2012. **40**(10): p. 4288-97.
735 36. Heinz, S., et al., *Simple combinations of lineage-determining transcription factors prime cis-*
736 *regulatory elements required for macrophage and B cell identities.* Mol Cell, 2010. **38**(4): p. 576-
737 89.
738 37. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time*
739 *quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.
740 38. Shashikant, T. and C.A. Ettensohn, *Genome-wide analysis of chromatin accessibility using ATAC-*
741 *seq.* Methods Cell Biol, 2019. **151**: p. 219-235.
742 39. Diwan, V., L. Brown, and G.C. Gobe, *Adenine-induced chronic kidney disease in rats.* Nephrology
743 (Carlton), 2018. **23**(1): p. 5-11.
744 40. Nickolas, T.L., et al., *Discriminants of prevalent fractures in chronic kidney disease.* J Am Soc
745 Nephrol, 2011. **22**(8): p. 1560-72.
746 41. Maravic, M., et al., *Incidence and risk factors for hip fractures in dialysis patients.* Osteoporos Int,
747 2014. **25**(1): p. 159-65.
748 42. Zhang, C., W. Tang, and Y. Li, *Matrix metalloproteinase 13 (MMP13) is a direct target of osteoblast-*
749 *specific transcription factor osterix (Osx) in osteoblasts.* PLoS One, 2012. **7**(11): p. e50525.
750 43. Hecht, J., et al., *Detection of novel skeletogenesis target genes by comprehensive analysis of a*
751 *Runx2(-/-) mouse model.* Gene Expr Patterns, 2007. **7**(1-2): p. 102-12.
752 44. Pereira, R.C., et al., *Impaired osteocyte maturation in the pathogenesis of renal osteodystrophy.*
753 Kidney Int, 2018. **94**(5): p. 1002-1012.
754 45. Metzger, C.E., et al., *Adenine-induced chronic kidney disease induces a similar skeletal phenotype*
755 *in male and female C57BL/6 mice with more severe deficits in cortical bone properties of male*
756 *mice.* PLoS One, 2021. **16**(4): p. e0250438.

757 46. Jia, T., et al., *A novel model of adenine-induced tubulointerstitial nephropathy in mice.* BMC
758 Nephrol, 2013. **14**: p. 116.
759 47. Karsenty, G. and M. Ferron, *The contribution of bone to whole-organism physiology.* Nature, 2012.
760 **481**(7381): p. 314-20.
761 48. Ducy, P., *The role of osteocalcin in the endocrine cross-talk between bone remodelling and energy*
762 *metabolism.* Diabetologia, 2011. **54**(6): p. 1291-7.

763

764

765 **Figure 1. Single-cell transcriptomic profiling of bone cells identified osteolineage**
766 **heterogeneity. A-B.** Long bones (femur, tibia, and humeri) were subjected to serial
767 collagenase/EDTA digestions (see Methods), and cells from fractions 4-9 were collected for FACS
768 sorting to isolate tdTomato-positive cells, followed by single-cell RNAseq library construction.
769 Analysis by qPCR confirmed high mRNA expression of osteocyte genes *Fgf23*, *Dmp1*, *Phex*, and
770 *Pdpn* in Ai9 positive cells *versus* Ai9 negative cells. **C.** Bioinformatic analysis was performed to
771 cluster the cells by UMAP. Each dot represents a single cell, and cells sharing the same color
772 code indicate discrete populations of transcriptionally similar cells. **D.** Defining cortical bone cell
773 types. Tnc/Mmp13 osteoblast cells were distinctly marked by *Tnc* and *Mmp13*. Osteoblast cells
774 showed high expression of *Smpd3* and *Bglap*. Osteocytes had the highest expression of *Dmp1*
775 and *Phex*. **E.** Expression density plots indicated cells with high transcription of *Col1a1*, *Tnc*, *Bglap*,
776 and *Phex*. **F.** The polar figure highlights the osteocyte markers isolated from our dataset (long
777 bone cell scRNAseq), Wang et al. (long bone and calvaria cell scRNAseq; [10]) and Youlten et al
778 (long bone bulk RNAseq; [9])**. G.** Canonical and non-canonical osteolineage genes were validated
779 as being highly expressed in cortical bone when compared to the expression detected in bone
780 marrow. **H.** Pseudotime analysis revealed Tnc/Mmp13 Osteoblasts as an osteoblast precursor
781 cell (P-OB), and shows their differentiation to osteoblasts (OB) and then osteocytes (OC).

782

783 **Figure 2. Differential signaling pathways in osteoblasts versus osteocytes. A.** Ingenuity
784 Pathway Analysis (IPA) software was used to predict the statistically significant canonical
785 pathways upregulated in Tnc/Mmp13 osteoblasts (P-OB), osteoblasts (OB) and osteocytes (OC)
786 from the scRNAseq data based upon a calculated probability score of ≥2 and a p-value of <0.05.
787 Further, a comparative analysis between osteolineage cells was performed to generate the
788 heatmap. The increase of blue intensity reflects the more significant pathways within cell types.
789 The gray dots indicate non-significant pathways. **B-E.** Monocle analysis shows pseudotime
790 trajectory mapping of osteolineage gene sets *Col1a1, Col1a2, Runx2;* pre-osteoblast gene sets
791 *Mmp13, Tnc, Spp1;* osteoblast gene sets *Bglap, Col11a2, Smpd3;* and osteocyte gene sets
792 *Dmp1, Irx5, Ackr3, Phex, Ptprz1,* and *Cd109*. **F.** The computational method scFEA was used to
793 infer cell-wise fluxome from the scRNAseq data to predict the metabolic profiling as well as the
794 differential metabolite conversion rate in cells that correspond to a metabolic flux value. Ridgeline
795 plots indicate the distribution values of metabolic flux in Tnc/Mmp13 osteoblasts (P-OB),
796 osteoblasts (OB), and osteocytes (OC). Each ridgeline represents the flux between two
797 metabolites, shown on the x-axis, for different osteolineage cells, shown on the y-axis. **G.** Violin
798 plot displaying *Cdo1* expression in scRNAseq dataset. **I.** The boxplots indicate the predictive
799 value of the conversion of methionine to cysteine and to pyruvate in P-OB, OB, and OC. **H.** The
800 bar plots show the top 10 different metabolites enriched in P-OB, OB, and OC.

801

802 **Figure 3. Molecular markers associated with osteoblast/osteocyte differentiation. A-top.**
803 Schematic representing cell differentiation from MSCs to osteoblasts/osteocytes. **A-lower.**
804 *Col1a1*, *Runx2, Tnc, Cthrc1, Pdpn, Phex, Sost, Ptprz1, Adpgk, Pgam1, Aldoa,* and *Acadm*
805 mRNAs were analyzed in MPC2 cells during osteogenic differentiation over the course of 1-4
806 weeks and normalized to β-actin. Data are represented as mean +/- standard deviation. **$p<0.01$,
807 and ***$p<0.001$ compared to Control (0 Wks). ###$p<0.001$ compared to 3 Wks. nd = not detected.
808 **B.** Volcano plots derived from the bulk RNAseq (left) and ATACseq datasets (right) show the
809 significant alterations of gene expression and chromatin accessibility detected in differentiated
810 *versus* undifferentiated cells. The x-axis shows the fold-change (in $\log_2$ scale). The y-axis
811 represents the significance. **C-top.** The 3 X 3 matrix heatmap highlights the number of genes with
812 differentially accessible regions (DARs), or differentially expressed genes (DEGs), or both DARs
813 and DEGs when compared differentiated *versus* undifferentiated cells. **C-lower.** Integration of
814 ATACseq and RNAseq data exhibited a higher correlation between genome-wide chromatin
815 accessibility changes (x-axis) and gene expression alterations (y-axis). **D.** Selected motifs of more
816 open regions of the genome in differentiated cells *versus* undifferentiated cells using the
817 ATACseq. **E.** Functional enrichment analysis using the Database for Annotation, Visualization
818 and Integrated Discovery (DAVID) on upregulated DEGs with more open DARs identified
819 biological pathways that were induced during osteogenic differentiation (numbers in parenthesis
820 are gene counts). **F.** Representative mRNA expression of upregulated genes in differentiated
821 versus undifferentiated cells enriched in cell morphogenesis and ossification. All the genes shown
822 were significantly upregulated (false discovery rate, FDR < 0.05). The osteocyte genes identified
823 from the scRNAseq dataset in **Figure 1J** were confirmed to be significantly upregulated in
824 osteoblasts/osteocytes cells *versus* control undifferentiated MPC2 cells. **G.** Representative
825 ATACseq peaks of undifferentiated cells (top track in gray) compared to differentiated cells
826 (osteoblasts/osteocytes bottom track in black) are shown.

827 **Figure 4. Precocious misregulation of osteoblast/osteocyte genes in CKD.** Eight-week old
828 wild type C57BL/6 male (shown in blue) and female (represented in red) mice were fed an adenine
829 diet (AD) to induce CKD for 2, or 4 weeks. Mice fed the casein diet for 2 weeks were used as
830 controls. **A.** Key serum biochemical analyses: phosphate, alkaline phosphatase, calcium, blood
831 urea nitrogen, creatinine, and iron are shown. **B.** Cortical porosity (left) and trabecular bone
832 volume (right) were measured using micro computed tomography (μCT). **C.** Representative
833 images of cortical bone μCT. **D.** Circulating FGF23 was assessed by ELISA. **E-H.** Real-time qPCR
834 was used to measure the mRNA expression of Tnc/Mmp13 osteoblast (pre-osteoblast),
835 osteoblast and osteocyte gene sets, as well as osteoclast genes. Data are shown as fold change
836 (2-ΔΔCt) relative to the housekeeping gene *β-Actin* and normalized to the experimental control
837 ('AD0'). The 'AD 0' mice were fed with a control diet (Casein) and sacrificed at 2 weeks; 'AD 2'
838 mice were fed with a CKD diet (adenine) and sacrificed at 2 weeks; and the 'AD 4' were mice fed
839 with the adenine diet and sacrificed at 4 weeks. Data are represented as mean +/- standard
840 deviation. *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.001$ compared to Control.

841

842 **Supplementary Figure 1. A.** Percentage of mitochondrial genes in scRNAseq dataset. **B-F.**
843 Expression density plots indicated cells with high transcription of *Ptprc (Cd45)*, *Cdh5*, *Pdgfra,*
844 *Mmp13,* and *Spp1*. **G.** The Monocle algorithm divided the cells based upon the original UMAP
845 into three partitions, as indicated by each color. The partition 'p1' identified the osteolineage cells,
846 'p2' corresponded to hematopoietic cells and 'p3' represented endothelial cells. **H-I.** Examples of
847 trajectory analysis performed on hematopoietic and marrow endothelial cells.

848

849 **Supplementary Figure 2. A.** Predictive upstream regulators in Tnc/Mmp13 osteoblast (pre-
850 osteoblasts), osteoblast and osteocyte. **B.** Dmp1 expression in different cell types **C.** scFEA
851 UMAP. **D.** Heatmap indicates the distribution of predicted cell-wise flux of glycolytic, TCA, serine
852 metabolism, fatty acid metabolism and glutamate metabolism relative to pre-osteoblast
853 (Tnc/Mmp13) values. The heatmap uses a column Z-score to show significant differences
854 between pre-osteoblasts (Tnc/Mmp13), osteoblasts, and osteocytes. **E-F.** Ridgeline plots indicate
855 the distribution values of metabolic flux in pre-osteoblasts (P-OB), osteoblasts (OB), and
856 osteocytes (OC). Each ridgeline represents the flux between two metabolites (x-axis) for different
857 cells that are plotted on the y-axis. **G.** Venn diagram shows the number of differentially expressed
858 genes for Tnc_mmp13_osteoblasts vs osteoblasts, and osteocytes vs osteoblasts. **H.** Ridgeline
859 plots indicate the distribution values of metabolic flux in pre-osteoblasts (Tnc/Mmp13), osteoblasts
860 (OB), osteocytes (OC), hematopoietic cells (Hem I and Hem II), and bone marrow endothelial
861 cells (BMEC).

862

863 **Supplementary Figure 3. A-F.** Chromatin accessibilities with corresponding gene expression of
864 Spp1 (p = 0.00933495), Col11a1 (p = 1.9363E-05), and Col11a2 (p = 7.2591E-06) in differentiated
865 (OB/OC) *versus* undifferentiated (MSC) cells.

866 **Supplementary Figure 4. A-C.** Assessment of chromatin accessibilities at *Mmp13* (p =
867 0.01025011), *Serpine2* (p = 3.4093E-09), and *Lifr* (p = 1.2547E-07) genomic regions. These
868 genes were detected as highly enriched in pre-osteoblasts. **D-E.** Assessment of chromatin
869 accessibilities at *Serpinf1* (p = 0.00099491), and *Cthrc1* (p = 0.013) genomic loci. These genes
870 were predicted to be highly enriched in osteoblasts. **D-E.** Assessment of chromatin accessibilities
871 at *Ackr3* (p = 2.503E-06), *Cd109* (p = 0.00179923), *Ptgis* (p = 6.8707E-06), *Spns2* (p =
872 0.00521604), and *Bmp2* (p = 0.00540629) genomic regions. These genes were detected primarily
873 in osteocytes. Top tracks (gray) corresponded to undifferentiated cells (MSC) and lower tracks
874 (black) corresponded to differentiated cells (OB/OC).

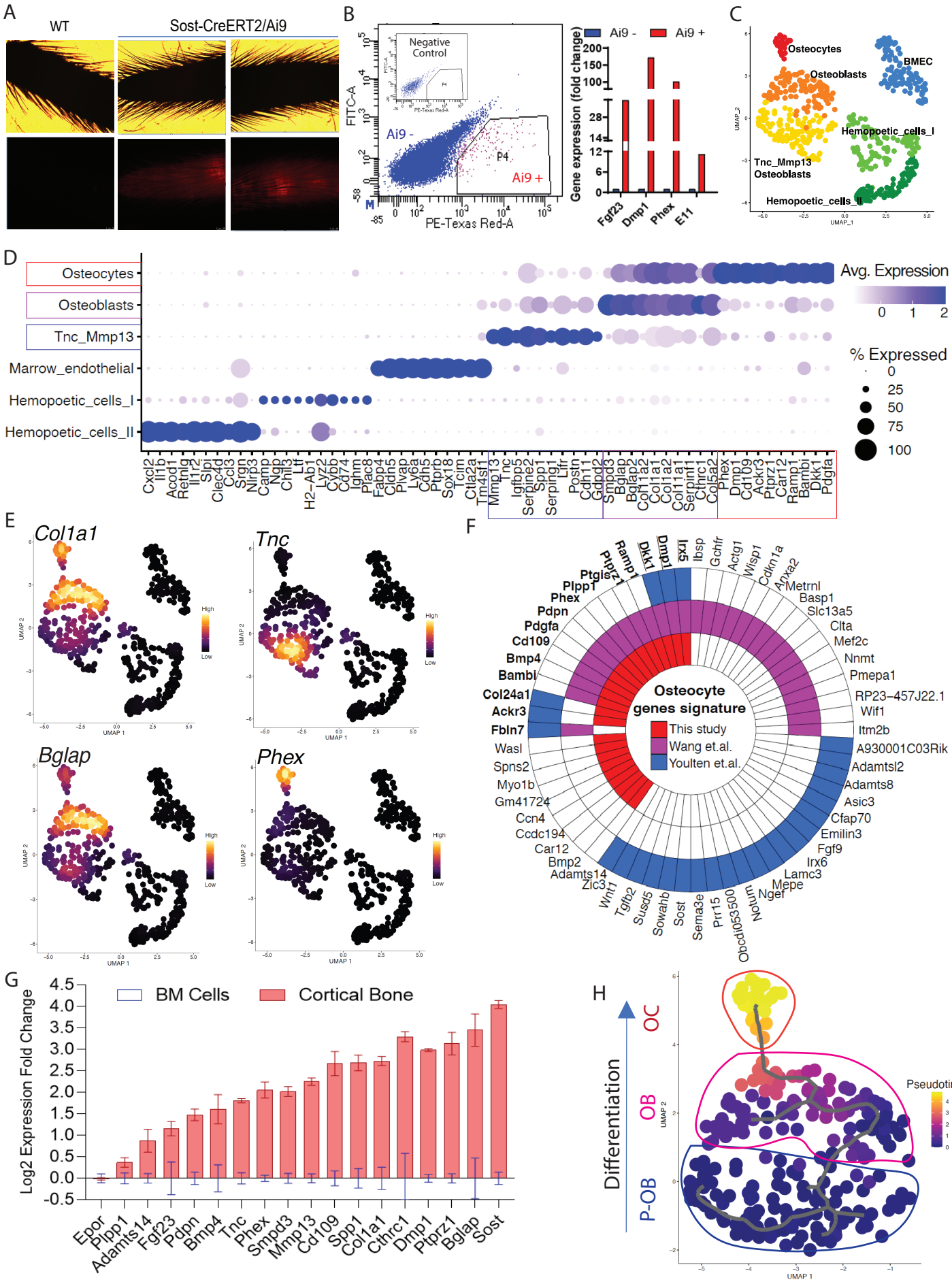
875

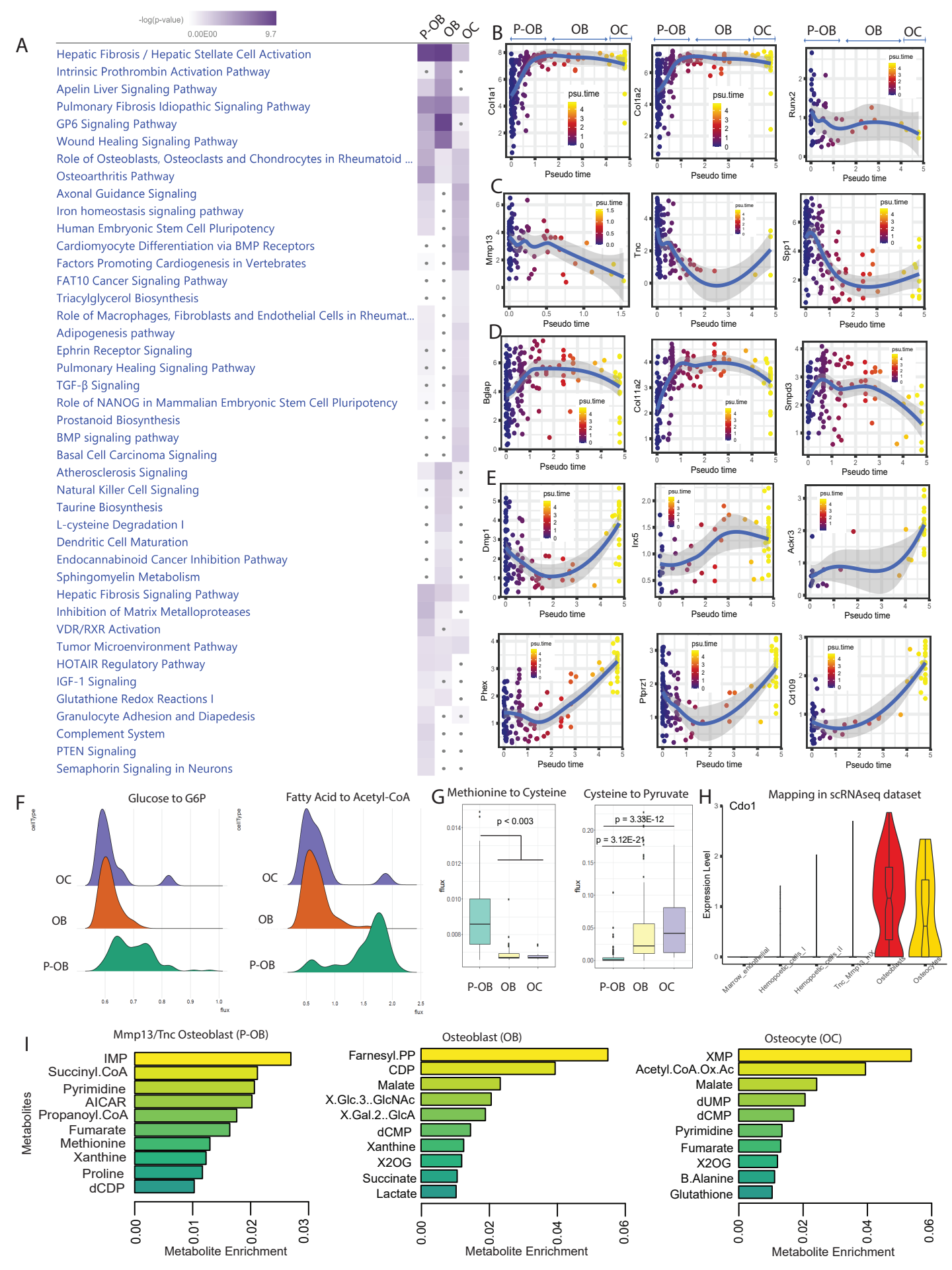
876 **Supplementary Figure 5. Osteocyte and osteoblast cell numbers in cortical and trabecular**
877 **bone. A.** The image shows osteocytes in the lacunae from femurs that were stained with
878 hematoxylin and eosin. **B.** Osteocyte numbers from cortical bone were counted in the midshaft of
879 bone and normalized to bone area. For counting of osteocytes in trabecular bone, cells were
880 counted in distal femur excluding endocortical surfaces and primary spongiosa, then normalized
881 to trabecular bone area. Osteoblasts were counted in distal femur in the same trabecular region
882 as where osteocytes were counted and normalized to trabecular bone surface.

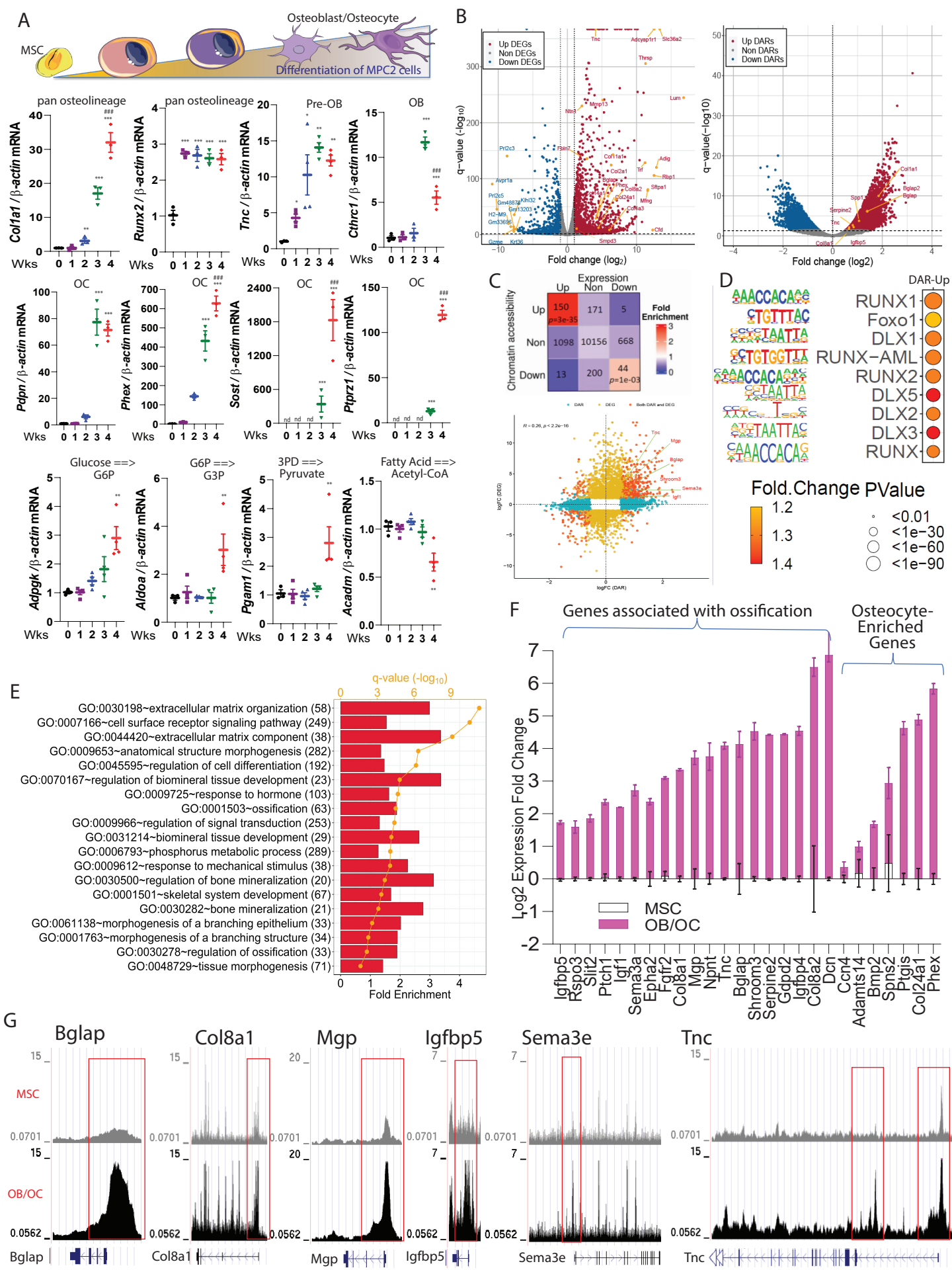
883

884

885

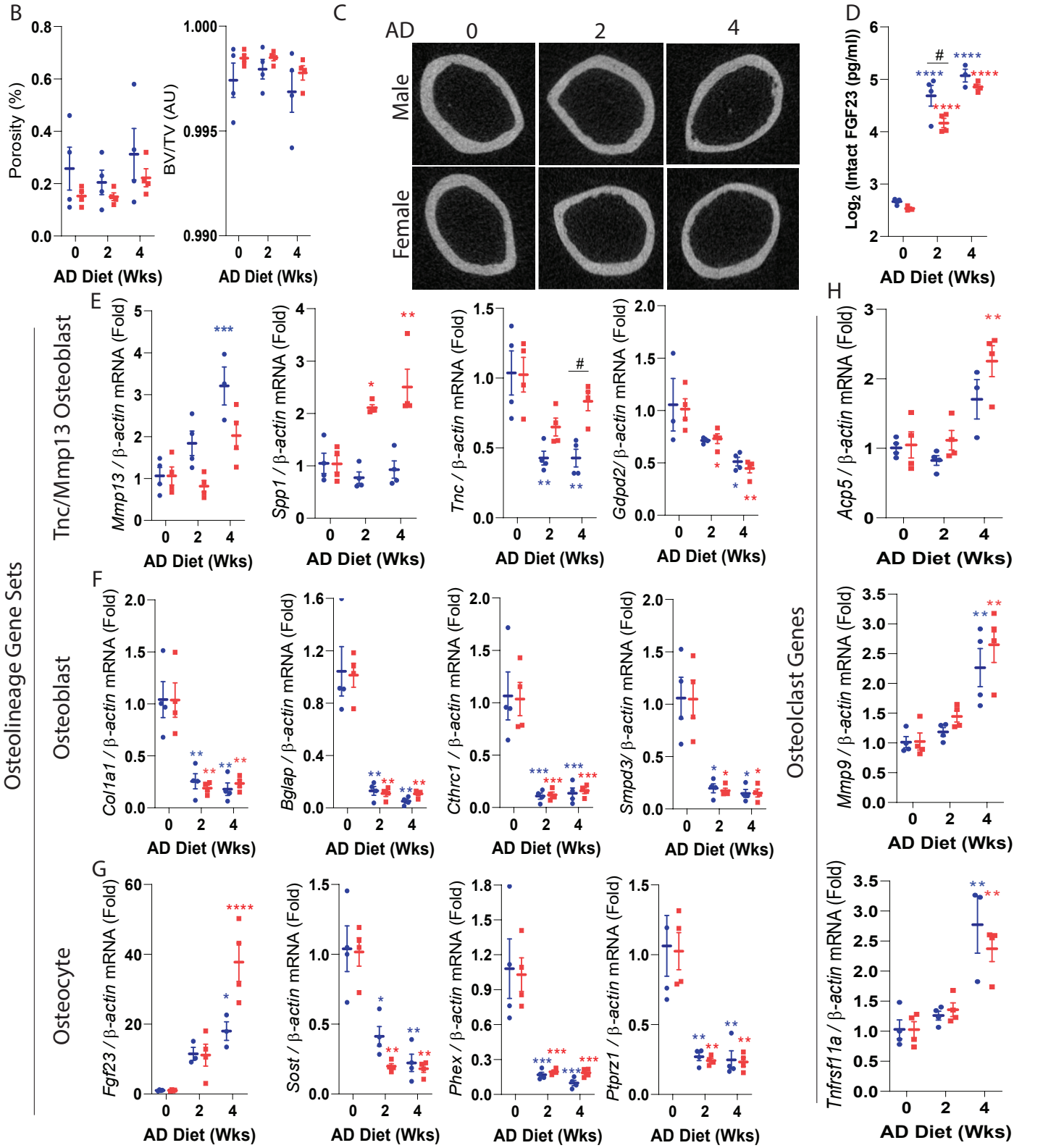

A
WT
Sost-CreERT2/Ai9
B
Negative Control
Ai9 -
Ai9 +
P4
FITC-A
PE-Texas Red-A
Gene expression (fold change)
Fgf23
Dmp1
Phex
E11
C
Osteocytes
Osteoblasts
BMEC
Hemopoetic_cells_I
Tnc_Mmp13 Osteoblasts
Hemopoetic_cells_II
UMAP_1
UMAP_2
D
Osteocytes
Osteoblasts
Tnc_Mmp13
Marrow_endothelial
Hemopoetic_cells_I
Hemopoetic_cells_II
Avg. Expression
% Expressed
E
Col1a1
Tnc
Bglap
Phex
High
Low
UMAP 1
UMAP 2
F
Osteocyte genes signature
This study
Wang et.al.
Youlten et.al.
G
BM Cells
Cortical Bone
Log2 Expression Fold Change
Epor
Plpp1
Adamts14
Fgf23
Pdpn
Bmp4
Tnc
Phex
Smpd3
Mmp13
Cd109
Spp1
Col1a1
Cthrc1
Dmp1
Ptprz1
Bglap
Sost
H
Differentiation
OC
OB
P-OB
Pseudotime
UMAP 1
UMAP 2


A
-log(p-value)
0.00E00
9.7
P-OB
OB
OC
Hepatic Fibrosis / Hepatic Stellate Cell Activation
Intrinsic Prothrombin Activation Pathway
Apelin Liver Signaling Pathway
Pulmonary Fibrosis Idiopathic Signaling Pathway
GP6 Signaling Pathway
Wound Healing Signaling Pathway
Role of Osteoblasts, Osteoclasts and Chondrocytes in Rheumatoid ...
Osteoarthritis Pathway
Axonal Guidance Signaling
Iron homeostasis signaling pathway
Human Embryonic Stem Cell Pluripotency
Cardiomyocyte Differentiation via BMP Receptors
Factors Promoting Cardiogenesis in Vertebrates
FAT10 Cancer Signaling Pathway
Triacylglycerol Biosynthesis
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumat...
Adipogenesis pathway
Ephrin Receptor Signaling
Pulmonary Healing Signaling Pathway
TGF-β Signaling
Role of NANOG in Mammalian Embryonic Stem Cell Pluripotency
Prostanoid Biosynthesis
BMP signaling pathway
Basal Cell Carcinoma Signaling
Atherosclerosis Signaling
Natural Killer Cell Signaling
Taurine Biosynthesis
L-cysteine Degradation I
Dendritic Cell Maturation
Endocannabinoid Cancer Inhibition Pathway
Sphingomyelin Metabolism
Hepatic Fibrosis Signaling Pathway
Inhibition of Matrix Metalloproteases
VDR/RXR Activation
Tumor Microenvironment Pathway
HOTAIR Regulatory Pathway
IGF-1 Signaling
Glutathione Redox Reactions I
Granulocyte Adhesion and Diapedesis
Complement System
PTEN Signaling
Semaphorin Signaling in Neurons
B
C
D
E
Pseudo time
psu.time
Col1a1
Col1a2
Runx2
Mmp13
Tnc
Spp1
Bglap
Col11a2
Smpd3
Dmp1
Irx5
Ackr3
Phex
Ptprz1
Cd109
F
Glucose to G6P
Fatty Acid to Acetyl-CoA
cellType
OC
OB
P-OB
flux
G
Methionine to Cysteine
Cysteine to Pyruvate
p < 0.003
p = 3.33E-12
p = 3.12E-21
H
Mapping in scRNAseq dataset
Cdo1
Expression Level
Marrow_endothelial
Hemopoietic_cells_I
Hemopoietic_cells_II
Tnc_Mmp13_HX
Osteoblasts
Osteocytes
I
Mmp13/Tnc Osteoblast (P-OB)
Metabolites
IMP
Succinyl.CoA
Pyrimidine
AICAR
Propanoyl.CoA
Fumarate
Methionine
Xanthine
Proline
dCDP
Metabolite Enrichment
Osteoblast (OB)
Farnesyl.PP
CDP
Malate
X.Glc.3..GlcNAc
X.Gal.2..GlcA
dCMP
Xanthine
X2OG
Succinate
Lactate
Osteocyte (OC)
XMP
Acetyl.CoA.Ox.Ac
Malate
dUMP
dCMP
Pyrimidine
Fumarate
X2OG
B.Alanine
Glutathione


A
MSC
Osteoblast/Osteocyte
Differentiation of MPC2 cells
pan osteolineage
Pre-OB
OB
OC
Col1a1 / β-actin mRNA
Runx2 / β-actin mRNA
Tnc / β-actin mRNA
Cthrc1 / β-actin mRNA
Pdpn / β-actin mRNA
Phex / β-actin mRNA
Sost / β-actin mRNA
Ptprz1 / β-actin mRNA
Wks
Glucose ==> G6P
G6P ==> G3P
3PD ==> Pyruvate
Fatty Acid ==> Acetyl-CoA
Adpgk / β-actin mRNA
Aldoa / β-actin mRNA
Pgam1 / β-actin mRNA
Acadm / β-actin mRNA
B
Up DEGs
Non DEGs
Down DEGs
q-value (-log10)
Fold change (log2)
Up DARs
Non DARs
Down DARs
q-value(-log10)
Fold change (log2)
C
Expression
Up Non Down
Chromatin accessibility
Up 150 171 5
p=3e-35
Non 1098 10156 668
Down 13 200 44
p=1e-03
Fold Enrichment
DAR
DEG
Both DAR and DEG
logFC (DEG)
logFC (DAR)
D
DAR-Up
RUNX1
Foxo1
DLX1
RUNX-AML
RUNX2
DLX5
DLX2
DLX3
RUNX
Fold.Change
PValue
<0.01
<1e-30
<1e-60
<1e-90
E
q-value (-log10)
GO:0030198~extracellular matrix organization (58)
GO:0007166~cell surface receptor signaling pathway (249)
GO:0044420~extracellular matrix component (38)
GO:0009653~anatomical structure morphogenesis (282)
GO:0045595~regulation of cell differentiation (192)
GO:0070167~regulation of biomineral tissue development (23)
GO:0009725~response to hormone (103)
GO:0001503~ossification (63)
GO:0009966~regulation of signal transduction (253)
GO:0031214~biomineral tissue development (29)
GO:0006793~phosphorus metabolic process (289)
GO:0009612~response to mechanical stimulus (38)
GO:0030500~regulation of bone mineralization (20)
GO:0001501~skeletal system development (67)
GO:0030282~bone mineralization (21)
GO:0061138~morphogenesis of a branching epithelium (33)
GO:0001763~morphogenesis of a branching structure (34)
GO:0030278~regulation of ossification (33)
GO:0048729~tissue morphogenesis (71)
Fold Enrichment
F
Genes associated with ossification
Osteocyte-Enriched Genes
Log2 Expression Fold Change
MSC
OB/OC
Igfbp5
Rspo3
Slit2
Ptch1
Igf1
Sema3a
Epha2
Fgfr2
Col8a1
Mgp
Npnt
Tnc
Bglap
Shroom3
Serpine2
Gdpd2
Igfbp4
Col8a2
Dcn
Ccn4
Adamts14
Bmp2
Spns2
Ptgis
Col24a1
Phex
G
Bglap
Col8a1
Mgp
Igfbp5
Sema3e
Tnc
MSC
OB/OC

A

Table 1. Serum biochemical analysis

	Male			Female		
	Ctl	Adenine 2 weeks	Adenine 4 weeks	Ctl	Adenine 2 weeks	Adenine 4 weeks
Pi (mg/dL)	12.22 ± 1.49	15.01 ± 0.95	17.23 ± 1.39[A]	12.53 ± 1.55	12.85 ± 1.45	14.88 ± 1.74
ALP (U/L)	103.75 ± 21.02	95 ± 10.66	113.75 ± 17.09	131.25 ± 34.16	88.75 ± 8.92	105 ± 11.18
Ca (mg/dL)	11.2 ± 0.29	11.86 ± 0.43	10.25 ± 1.28	10.5 ± 0.32	11.08 ± 0.62	11.26 ± 0.69
BUN (mg/dL)	64.51 ± 15.63	140.00 ± 23.11[B]	176.26 ± 66.17[C]	56.15 ± 6.62	115.88 ± 13.61	175.54 ± 29.14[C]
Creatinine mg/dL)	0.40 ± 0.19	0.69 ± 0.13	0.71 ± 0.20	0.50 ± 0.08	0.65 ± 0.11	0.71 ± 0.06
Iron (μM)	37.13 ± 7.81	15.38 ± 6.05[B]	23.75 ± 4.48[A]	28.38 ± 7.81	8.88 ± 4.44[B]	12.38 ± 2.88[A]

Data presented as mean ± SD. N = 4 mice per group. [A]p < 0.05; [B]p < 0.01; [C]p < 0.001 vs. Control (Ctl)