ARTICLE

Check for updates

Type-I collagen produced by distinct fibroblast lineages reveals specific function during embryogenesis and Osteogenesis Imperfecta

Yang Chen [1], Sujuan Yang[1], Sara Lovisa [1], Catherine G. Ambrose[2], Kathleen M. McAndrews [1], Hikaru Sugimoto[1] & Raghu Kalluri [1,3,4 ✉]

Type I collagen (Col1) is the most abundant protein in mammals. Col1 contributes to 90% of the total organic component of bone matrix. However, the precise cellular origin and functional contribution of Col1 in embryogenesis and bone formation remain unknown. Single-cell RNA-sequencing analysis identifies Fap+ cells and Fsp1+ cells as the major contributors of Col1 in the bone. We generate transgenic mouse models to genetically delete Col1 in various cell lineages. Complete, whole-body Col1 deletion leads to failed gastrulation and early embryonic lethality. Specific Col1 deletion in Fap+ cells causes severe skeletal defects, with hemorrhage, edema, and prenatal lethality. Specific Col1 deletion in Fsp1+ cells results in Osteogenesis Imperfecta-like phenotypes in adult mice, with spontaneous fractures and compromised bone healing. This study demonstrates specific contributions of mesenchymal cell lineages to Col1 production in organogenesis, skeletal development, and bone formation/repair, with potential insights into cell-based therapy for patients with Osteogenesis Imperfecta.

---

[1] Department of Cancer Biology, University of Texas MD Anderson Cancer Center, Houston, TX 77054, USA. [2] Department of Orthopaedic Surgery, University of Texas Health Science Center at Houston, Houston, TX, USA. [3] Department of Bioengineering, Rice University, Houston, TX, USA. [4] Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX, USA. ✉email: rkalluri@mdanderson.org

Type I collagen (Col1), a fibrillar collagen, is the most abundant protein in mammals. The basic functional unit of Col1 is a heterotrimer consisting of two α1 chains and one α2 chain that assemble together to form a triple helical structure. Each α-chain polypeptide is synthesized in the cytosol and combines with two other α-chains to generate a triple-helical type I procollagen (in the format of α1α2α1) with N-terminal and C-terminal propeptides. Subsequently, the procollagen molecule is secreted into the extracellular space where the N-terminal and C-terminal propeptides are cleaved by propeptidases, generating the basic functional heterotrimeric unit of Col1. The Col1 triple helical rod-like molecules interact with each other to form fibrils and undergo further crosslinking to form large bundles of fibers.

During embryogenesis, many organs express Col1 to likely facilitate cellular migration, differentiation and structural compartmentalization[1]. Insertional mutation of *Col1a1* gene resulted in complete absence of Col1 and led to embryonic lethality[2]. Genetic mutations in *Col1a1* or *Col1a2* genes (or other genes affecting Col1 biosynthesis) in humans lead to multiple subtypes of Osteogenesis Imperfecta (OI) syndrome, a bone disease referred to as the 'brittle bone disease'[3–5]. Previous studies established and investigated various Osteogenesis Imperfecta mouse models using different genetic targeting/manipulating approaches with some success[6–19]. Particularly, recent studies demonstrated that *Col1a1* heterozygous mutation/deletion in transgenic mouse models lead to Osteogenesis Imperfecta phenotypes[14,19].

Col1 contributes around 90% of the total organic component of bone matrix. The development, formation, and homeostasis of bone tissue involve mesenchymal cell lineages in the synthesis, deposition, and remodeling of Col1[4,20]. These mesenchymal cell lineages include osteoblasts and fibroblasts, which can be differentiated from mesenchymal stem/stromal cells (MSCs)[3,4,21–25]. However, the precise contributions of Col1 produced by these mesenchymal cell lineages and fibroblasts in the development and function of bones remain unknown. Recent studies, using single-cell RNA sequencing analysis (sc-RNA-seq) and other techniques, highlighted the distinct features and potential functions of various osteogenic cell subpopulations in the bone and bone marrow, shedding some light on the potential mechanisms driving bone homeostasis and repair[26–28]. In this present study, we analyzed sc-RNA-seq dataset on bone and bone marrow stroma of healthy mice[27], and identified the expression of *Col1a1* gene dominantly in Fap+ (fibroblast activation protein; encoded by *Fap*) mesenchymal cells, as well as S100a4+ (also known as fibroblast-specific protein 1, Fsp1; encoded by *S100a4*) mesenchymal cells. Next, we generated novel genetically engineered mouse models (GEMMs) allowing the conditional deletion of Col1 specifically in Fap+ or Fsp1+ cells, as well as in other cell lineages. The *Col1a1^loxP/loxP* mouse strain was crossed with multiple Cre recombinase mouse strains including *CMV-Cre, Fap-Cre, Fsp1 (S100a4)-Cre, αSMA (Acta2)-Cre,* and *Cdh5-Cre* strains. These mice revealed distinct phenotypes due to Cre recombinase-driven Col1 deletion in targeted cell lineages. Taken together, this study provides comprehensive investigation of the functional roles of Col1 produced by various cell lineages in organogenesis, osteogenesis, and bone-related diseases.

Results

Analysis of single-cell RNA-sequencing of mouse bone and bone marrow identifies type I collagen (Col1)-producing fibroblast/mesenchymal lineages. Bone marrow is one of the sources of mesenchymal lineages, pluripotent cells that differentiate into a variety of cell types such as osteoblasts, chondrocytes, and fibroblasts[3,4,21–25]. To identify the cell populations involved in type I collagen (Col1) production in osteogenesis, we further analyzed the single-cell RNA-sequencing data[27] of combined cells from mouse bone (including the periosteum) and the bone marrow (Fig. 1a, Supplementary Fig. S1 and S2). Consistent with previous report[27], six distinct non-hematopoietic cell populations were identified in the bone and the bone marrow (Fig. 1a–c, Supplementary Fig. S1). These are MSCs (expressing *Lepr* and *Cxcl12*, also called Cxcl12-abundant-reticular cells)[28–30], MSC-descendent osteolineage cells (OLCs; expressing *Bglap* and *Runx2*)[31,32], chondrocytes (expressing *Acan* and *Col2a1*)[33,34], fibroblasts (expressing *S100a4* and *Dcn*)[35], bone marrow endothelial cells (BMECs; expressing *Cdh5*, encoding VE-cadherin)[36], and smooth-muscle cell/mesenchymal lineage (SMC/mesenchymal) (expressing *Acta2* and *Myh11*)[37].

Next, we determined that *Col1a1* is highly expressed by MSC-descendent OLCs and fibroblasts (Fig. 1d). We simultaneously examined the expression of several genes encoding common fibroblast/mesenchymal markers, such as *Fap* (encoding fibroblast activation protein, Fap), *S100a4* (encoding Fibroblast-specific protein 1, Fsp1), and *Acta2* (encoding α-smooth muscle actin, αSMA), among the cell populations of bone and bone marrow. *Fap* expression was observed in cell clusters defined as MSC-descendent OLCs (Fig. 1e), and also observed in chondrocytes, fibroblasts, and MSCs. *Fap* expression in OLCs was observed in the matrix-producing mature osteoblast subpopulation expressing *Bglap* (encoding osteocalcin) (Supplementary Fig. 1b), and also the osteoprogenitor cell subpopulation expressing *Sp7* (Osterix) and *Mmp13* (Supplementary Fig. 1c)[27,28]. *S100a4* (Fsp1) expression was dominantly detected in all fibroblast clusters[27,35], but also detected in SMC/mesenchymal cells, chondrocytes, and endothelial cells (Fig. 1f). *Acta2* (αSMA) was predominantly expressed in the SMC/mesenchymal cells, and was expressed in other cell types such as fibroblasts, chondrocytes, or osteolineage cells (Fig. 1g). It is important to note that *Col1a1*-expressing cell populations were largely constituted of Fap-expressing cells and Fsp1-expressing cells. *Col1a1* expression in OLCs was found predominantly in differentiated osteoblasts expressing *Bglap* (Fig. 1d, Supplementary Fig. 1b), but sparsely in osteoprogenitor cells expressing *Sp7* and *Mmp13* (Supplementary Fig. 1c), which is consistent with recent observations[26,27]. Taken together, these results suggest that Fap-expressing cells and Fsp1-expressing cells are likely associated with Col1 synthesis in the bones, prompting us to further investigate the functional role of Col1 produced by these cell lineages.

Systemic deletion of *Col1a1* employing CMV-Cre results in failed gastrulation and early embryonic lethality. We generated *Col1a1^loxP/loxP* mice allowing for conditional deletion of floxed *Col1a1* gene (encoding type I collagen α1 chain). Complete Col1 deletion can be achieved with loss of *Col1a1* due to the fact that Col1α1 chain (encoded by *Col1a1* gene) is essential for Col1 triple helical trimer formation, while Col1α2 chain (encoded by *Col1a2* gene) alone cannot form Col1 trimers. Next, to investigate the functional role of Col1 in embryogenesis and organ development, we crossed *Col1a1^loxP/loxP* mice with *CMV-Cre* mice (Fig. 2a) to achieve the systemic (whole-body) deletion of Col1 (*CMV-Cre;-Col1a1^loxP/loxP*, referred to also as Col1a1^cmvKO). *CMV-Cre* transgene, under the control of human cytomegalovirus (CMV) minimal promoter, ubiquitously deletes *loxP*-flanked genes in all tissues including germline cells. Live progeny of Col1a1^cmvKO was not obtained, indicating the embryonic lethality (Fig. 2b). To further determine the stage of lethality, we dissected embryos at E9.5 and E12.5 (Fig. 2c–e). Specifically, Col1a1^cmvKO embryos at E9.5 revealed a retarded development with the morphology that was normally observed at the E6-E7 gastrulation stage (Fig. 2c).
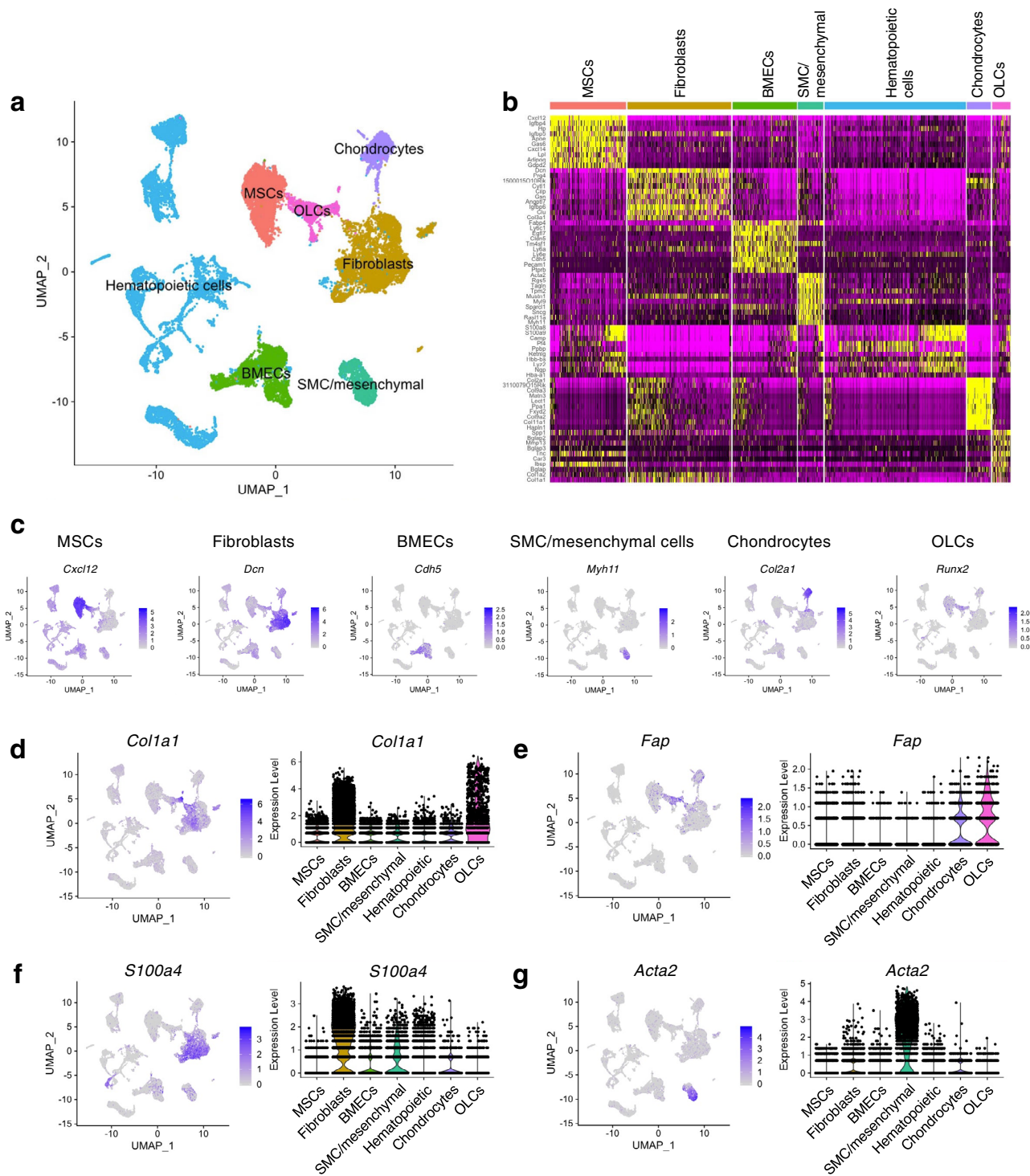
Fig. 1 Single-cell RNA-sequencing analysis identifies Col1-producing cells/lineages in bone and bone marrow fractions. a–c Single-cell RNA-sequencing (sc-RNA-seq) analysis of cell mixture of mouse bone and bone marrow fractions ($n = 6$ mice) from a recently published dataset (GSE128423). Functional clusters of cells were defined with group definition listed, as implemented in the Seurat R package (**a**). MSCs mesenchymal stem/stromal cells, OLCs osteolineage cells derived from MSCs, BMECs bone marrow endothelial cells, SMC/mesenchymal smooth-muscle cell/mesenchymal lineage. **b** Heat map showing the scaled expression values of discriminating signature genes for the functional clusters of cells defined above. (**c**) Expression profile of representative signature gene for each of the non-hematopoietic cell clusters. Continued in Supplementary Fig. 1 and 2. **d–g** Expression profile of type I collagen α1 chain (*Col1a1*; **d**), fibroblast activation protein (*Fap*; **e**), fibroblast-specific protein 1 (*S100a4*; **f**), and α-smooth muscle actin (*Acta2*; **g**) among defined cell clusters shown in UMAP plot and violin plot.
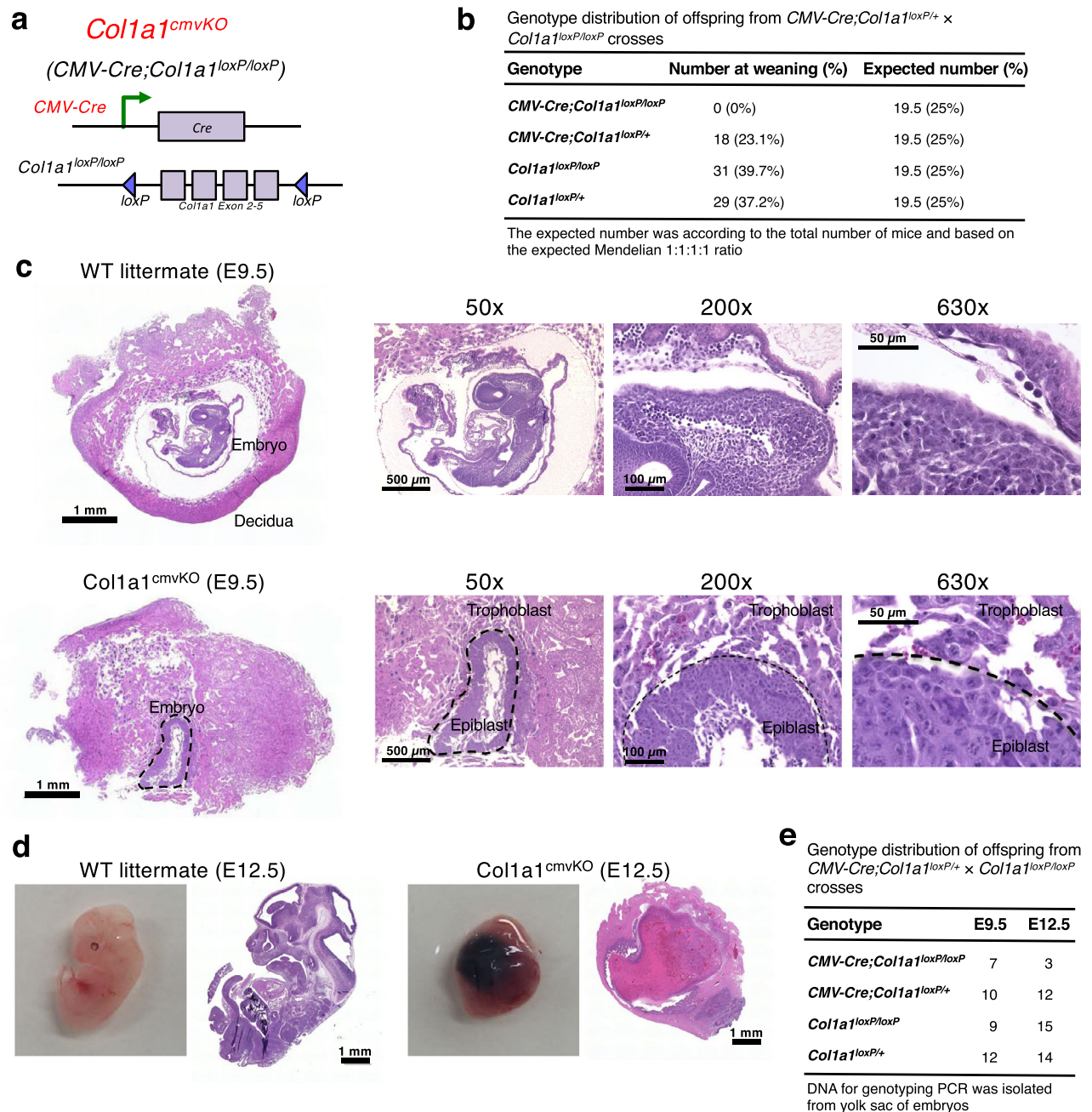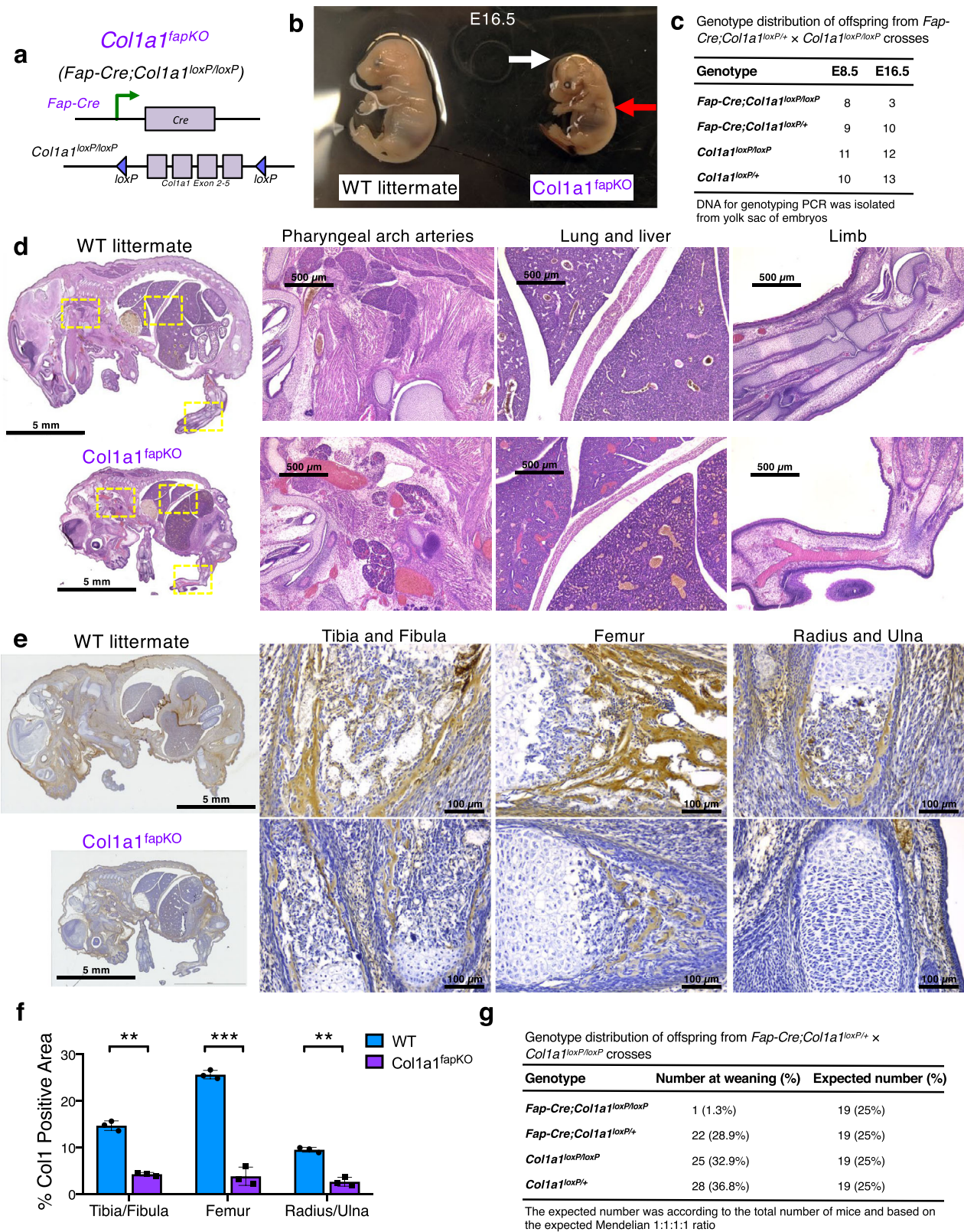
a

Col1a1^{cmvKO}

(*CMV-Cre;Col1a1^{loxP/loxP}*)

b Genotype distribution of offspring from *CMV-Cre;Col1a1^{loxP/+}* × *Col1a1^{loxP/loxP}* crosses

Genotype	Number at weaning (%)	Expected number (%)
CMV-Cre;Col1a1^{loxP/loxP}	0 (0%)	19.5 (25%)
CMV-Cre;Col1a1^{loxP/+}	18 (23.1%)	19.5 (25%)
Col1a1^{loxP/loxP}	31 (39.7%)	19.5 (25%)
Col1a1^{loxP/+}	29 (37.2%)	19.5 (25%)

The expected number was according to the total number of mice and based on the expected Mendelian 1:1:1:1 ratio

c

d

e Genotype distribution of offspring from *CMV-Cre;Col1a1^{loxP/+}* × *Col1a1^{loxP/loxP}* crosses

Genotype	E9.5	E12.5
CMV-Cre;Col1a1^{loxP/loxP}	7	3
CMV-Cre;Col1a1^{loxP/+}	10	12
Col1a1^{loxP/loxP}	9	15
Col1a1^{loxP/+}	12	14

DNA for genotyping PCR was isolated from yolk sac of embryos

Fig. 2 Systemic deletion of Col1 by CMV-Cre transgene results in early embryonic lethality. a Genetic strategy to achieve systemic (whole-body) knockout of type I collagen α1 chain (*Col1a1*) by crossing the *Col1a1^{loxP/loxP}* mice with generic *CMV-Cre* mice. **b** Genotype distribution in live offspring examined at the time of weaning from the crosses between *CMV-Cre;Col1a1^{loxP/+}* and *Col1a1^{loxP/loxP}* mice. **c, d** Hematoxylin and eosin (H&E) staining of the embryos of *CMV-Cre-negative;Col1a1^{loxP/loxP}* (WT; wild-type) and *CMV-Cre;Col1a1^{loxP/loxP}* (Col1a1^{cmvKO}) at embryonic day E9.5 (**c**) and E12.5 (**d**). Col1a1^{cmvKO} embryo of E9.5 was indicated by dashed line, with labeled epiblast and trophoblast areas. Representative images were shown for WT embryos (*n* = 6) and Col1a1^{cmvKO} embryos (*n* = 6) of E9.5 (**c**); and WT embryos (*n* = 6) and Col1a1^{cmvKO} embryos (*n* = 3, limited number of obtained embryos due to lethality) of E12.5 (**d**). **e** Genotype distribution in live offspring examined at E9.5 and E12.5 from the crosses between *CMV-Cre;Col1a1^{loxP/+}* and *Col1a1^{loxP/loxP}* mice. Scale bars of whole-mount sections, 1 mm; Scale bars at 50× magnification, 500 μm; Scale bars at 200× magnification, 100 μm; Scale bars at 630× magnification, 50 μm.

Our results indicate the essential role of Col1 during gastrulation, consistent with previous observations identifying Col1 protein between embryonic endoderm and mesoderm[2]. Moreover, our observations that the whole-body KO of Col1 (truncated *Col1a1* exons 2-5 by CMV-Cre) resulted in failed gastrulation, arrested embryonic development at E6-E7, and the death of embryos (Fig. 2d, e), validated previous observations with a mouse model harboring the loss of Col1 by insertional mutation of *Col1a1* gene,

resulting in embryonic lethality at E12-E14 due to blood vessel rupture[2]. *CMV-Cre;Col1a1^{loxP/+}* mice harboring heterozygous *Col1a1* loss revealed minimal abnormalities in their general health, overall survival, or their breeding ability until 6-month age (Supplementary Fig. 3a). Mild spontaneous bone fractures could be observed by X-ray examination in 6-month-old *CMV-Cre;-Col1a1^{loxP/+}* mice (Supplementary Fig. 3b), consistent with previous studies using Osteogenesis Imperfecta type I mouse models

a Col1a1^{fapKO} (Fap-Cre;Col1a1$^{loxP/loxP}$)

b E16.5 — WT littermate, Col1a1^{fapKO}

c Genotype distribution of offspring from Fap-Cre;Col1a1$^{loxP/+}$ × Col1a1$^{loxP/loxP}$ crosses

Genotype	E8.5	E16.5
Fap-Cre;Col1a1$^{loxP/loxP}$	8	3
Fap-Cre;Col1a1$^{loxP/+}$	9	10
Col1a1$^{loxP/loxP}$	11	12
Col1a1$^{loxP/+}$	10	13

DNA for genotyping PCR was isolated from yolk sac of embryos

d WT littermate / Col1a1^{fapKO} — Pharyngeal arch arteries, Lung and liver, Limb

e WT littermate / Col1a1^{fapKO} — Tibia and Fibula, Femur, Radius and Ulna

f % Col1 Positive Area — WT, Col1a1^{fapKO} (Tibia/Fibula **, Femur ***, Radius/Ulna **)

g Genotype distribution of offspring from Fap-Cre;Col1a1$^{loxP/+}$ × Col1a1$^{loxP/loxP}$ crosses

Genotype	Number at weaning (%)	Expected number (%)
Fap-Cre;Col1a1$^{loxP/loxP}$	1 (1.3%)	19 (25%)
Fap-Cre;Col1a1$^{loxP/+}$	22 (28.9%)	19 (25%)
Col1a1$^{loxP/loxP}$	25 (32.9%)	19 (25%)
Col1a1$^{loxP/+}$	28 (36.8%)	19 (25%)

The expected number was according to the total number of mice and based on the expected Mendelian 1:1:1:1 ratio

harboring *Col1a1* heterozygous truncation/mutation[14,19]. Further investigations will be required to evaluate the mechanism by which such *Col1a1* haploinsufficiency in *CMV-Cre;Col1a1$^{loxP/+}$* mice may lead to the phenotype similar to human type I Osteogenesis Imperfecta (OI).

Specific deletion of Col1 in Fap$^+$ cells leads to impaired skeletal development and embryonic lethality. To test the functional role of Col1 produced by the Fap$^+$ cells in the bone, we generated the *Fap-Cre;Col1a1$^{loxP/loxP}$* (Col1a1^{fapKO}) mice (Fig. 3a). Most Col1a1^{fapKO} embryos were non-viable past E16.5

Fig. 3 Specific deletion of Col1 in Fap-lineage cells leads to impaired skeletal development and late embryonic lethality. a Genetic strategy to delete *Col1a1* specifically in Fap-lineage cells by crossing the *Col1a1^{loxP/loxP}* mice with *Fap-Cre* mice. **b** Wild-type (WT; with genotype of *Fap-Cre-negative;Col1a1^{loxP/loxP}*) and Col1a1^{fapKO} (*Fap-Cre;Col1a1^{loxP/loxP}*) embryos at E16.5. The Col1a1^{fapKO} embryo exhibits hydrops fetalis (white arrow) and hemorrhage (red arrow). **c** Genotype distribution in live offspring documented examined at E9.5 and E12.5 from the crosses between *Fap-Cre;Col1a1^{loxP/+}* and *Col1a1^{loxP/loxP}* mice. **d–f** H&E staining (**d**) and Col1 immunohistochemistry staining (**e**) of the embryos of WT and Col1a1^{fapKO} at E16.5. Quantification of % positive area for Col1 staining was based on three mice per group (**f**). The unpaired, two-tailed *t* test was used to compare the mean of two independent groups. **$P = 0.00116$ (tibia/fibula), ***$P = 0.00057$ (femur), **$P = 0.00149$ (radius/ulna). Data are represented as mean ± SEM. Additional Picrosirius Red staining was shown in Supplementary Fig. 4. **g** Genotype distribution in live offspring documented at the time of weaning from the crosses between *Fap-Cre;Col1a1^{loxP/+}* and *Col1a1^{loxP/loxP}* mice. Scale bars of whole-mount sections, 5 mm; Scale bars at 50× magnification, 500 μm; Scale bars at 200× magnification, 100 μm.

and revealed prominent hemorrhage and edema (Fig. 3b), while the number of Col1a1^{fapKO} embryos at E8.5 was comparable to the expected number based on Mendelian ratio (Fig. 3c). Histological analysis further confirmed the hemorrhagic phenotype in major organs and pharyngeal arch arteries in Col1a1^{fapKO} embryos (Fig. 3d). Significantly decreased collagen deposition in skeletal system was also observed in Col1a1^{fapKO} embryos, as examined by Col1 immunohistochemical staining (Fig. 3e, f) as well as Picrosirius Red staining (Supplementary Fig. 4 and 5a). Despite the decreased Col1 deposition, the skeletal morphology of Col1a1^{fapKO} embryos did not reveal prominent deformity as compared with WT embryos. Live progeny of Col1a1^{fapKO} genotype was almost never observed with only one exception, indicating that Col1a1^{fapKO} leads to embryonic lethality with almost 100% penetrance (Fig. 3g).

Despite the fact that most Col1a1^{fapKO} embryos died prenatally, we indeed observed one live progeny with Col1a1^{fapKO} genotype. This mouse was drastically smaller in size compared to its littermates with severely defective/fractured skeleton throughout the body (Fig. 4a). Histological analysis further revealed the dramatic fractures in Col1a1^{fapKO} mouse (Fig. 4b), associated with significantly decreased Col1 deposition in the bones (Fig. 4c, d). The fractures in Col1a1^{fapKO} mouse was accompanied by the significant accumulation of chondrocytes and cartilage deposition, as examined by H&E and Safranin-O/Fast Green staining (Fig. 4c, d). Prominent deformities were also observed in both the limb skeleton and the axial skeleton of the Col1a1^{fapKO} mouse (Supplementary Fig. 5b). A lineage-tracing *Fap-Cre;LSL-tdTomato* mouse model exhibited colocalization between *Fap-Cre*-tdTomato⁺ cells and osteoblast marker osteocalcin (Fig. 4e), consistent with previous sc-RNA-seq analysis showing *Fap*-expression in osteolineage cells (Fig. 1e). These results do not rule out the possibility that Fap⁺ fibroblasts may derive from osteolineage cells or osteoblasts. Taken together, these results demonstrate the fundamental role of Fap⁺ cells in the de novo production of Col1 in vascular organogenesis and bone development during embryogenesis.

Specific deletion of Col1 in Fsp1⁺ cells leads to Osteogenesis Imperfecta phenotype in mice. To test the function of Col1 produced by the Fsp1⁺ cells in the bone stroma, we generated the *Fsp1-Cre;Col1a1^{loxP/loxP}* (Col1a1^{fspKO}) mice (Fig. 5a). These mice were born normal but Col1 deletion in Fsp1-lineage cells led to the occurrence of spontaneous Osteogenesis Imperfecta (OI) symptoms in adult Col1a1^{fspKO} mice, which was analogous to the previous observations employing systemic transgenic mice harboring dominant mutations of *Col1a1* and *Col1a2*[6,12,14,15,38,39]. Col1a1^{fspKO} mice developed OI phenotype with full penetrance, resulting in death of mice at around 6-12 months of age (Fig. 5b). At the age of 6 month, all Col1a1^{fspKO} mice exhibited the inevitable occurrence of fractures and deformations, frequently observed in the joints, pelvis, ischium, and head of femur (Fig. 5c, d). Such OI-like phenotype of Col1a1^{fspKO} mice was very similar to the well-characterized *Col1a2^{OIM/OIM}* mice[6] and

phenotypically resembled the moderate to severe OI type III in humans. As shown in Supplementary Fig. 6a, b as a reference, the *Col1a2^{OIM/OIM}* mouse model harbors homozygous whole-body point-mutated *Col1a2* alleles and develops severe OI feature in the bones, similar to the Col1a1^{fspKO} mice.

Further radiographic analysis revealed that mild deformation of pelvis and ischium was observed as early as 2–3 months of age, and more severe fracture in hindlimbs associated with dysfunction by the 5–8 months or age (Fig. 5d, e). Noticeable deformation of elbow joints can also be seen as early as 3–4 months of age, resulting in forelimb dysfunction by 6-9 months of age (Fig. 5d, e). In contrast, the age-matched wild-type (*Fsp1-Cre-negative;Col1a1^{loxP/loxP}*) littermates did not reveal any of the above symptoms (Supplementary Fig. 7a). Of note, Col1a1^{fspKO} mice were born at the expected Mendelian ratio (Supplementary Fig. 7b). Additionally, Col1a1^{fspKO} mice exhibited no gross physical or behavioral abnormalities before one-month of age (Fig. 5e, f). Significant body weight loss and hampered movement were observed in all Col1a1^{fspKO} mice by 3–6 months of age (Fig. 5f and Supplementary Movie 1). Female and male Col1a1^{fspKO} mice revealed similar body weight loss and overall survival (Fig. 5f and Supplementary Fig. 7c). Micro-CT measurements confirmed the significantly decreased bone mineral density and bone material properties in six-month-old Col1a1^{fspKO} mice as compared with age-matched WT littermates (Fig. 5g and Supplementary Fig. 7d). Spontaneous fractures in the head of femur were ubiquitously observed in Col1a1^{fspKO} mice (Fig. 5g). Biomechanical assays validated the brittleness of femurs from six-month-old Col1a1^{fspKO} mice as compared with WT littermates (Fig. 5h).

Next, we performed histological analysis of joint lesions from Col1a1^{fspKO} mice (Fig. 6a and Supplementary Fig. 7e). In the joint lesions with fractures and deformations, we observed significant callus formation composed of chondrocytes and cartilage components adjacent to the fractured areas, as observed by Safranin-O staining of cartilage matrix (Fig. 6a, b). The joint lesions from Col1a1^{fspKO} mice also revealed significantly decreased overall levels of fibrillar Col1, as examined using circularly polarized light microscopy on Picrosirius Red-stained sections (Fig. 6c). Lineage-tracing *Fsp1-Cre;LSL-YFP* mouse model revealed *Fsp1-Cre*-induced YFP expression in the periosteum, connective tissue, and bone marrow (Fig. 6d). *Fsp1-Cre*-induced YFP expression was also observed in osteocalcin-expressing osteoblasts in a mosaic pattern (Fig. 6e). These observations indicate that the phenotypes of Col1a1^{fspKO} mice may result from the combined effect of Col1 deletion in Fsp1 cell lineage including both osteoblasts and other periosteal cells. Collectively, these results establish Col1a1^{fspKO} mice as a novel cell-specific transgenic mouse model for OI. Our observations that the Col1a1^{fspKO} mice reveal insignificant abnormalities in the bones before 1 month of age and develop OI symptoms after 2–3 months of age, emphasize the crucial role of Col1 produced specifically by Fsp1-lineage cells in regulating proper bone development and healing of fractures.
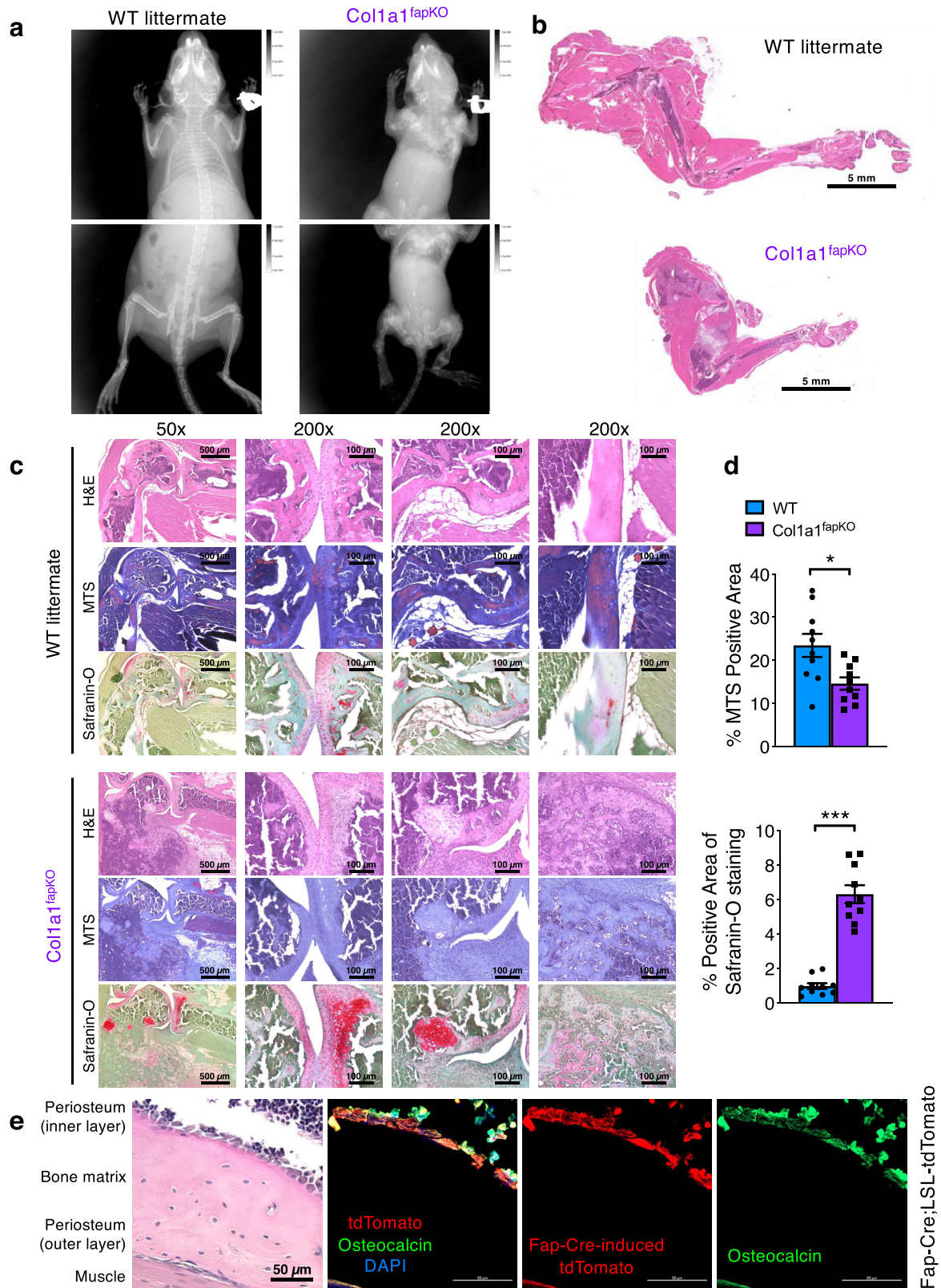
Col1 deletion in Fsp1$^+$ cells alters the composition of neutrophils and macrophages in the bone marrow. In order to further investigate the impact of Col1 deletion in the Fsp1$^+$ cells on the bone marrow microenvironment, we conducted single-cell RNA-sequencing analysis (sc-RNA-seq) on unfractionated live cell mixture from Col1a1^{fspKO} mouse bone marrow (Fig. 7a, Supplementary Fig. 8a–c). Col1a1^{fspKO} mouse bone marrow was

examined at 2-month age (when OI phenotype initiates in these mice), as compared with age-matched wild-type littermate bone marrow. Most of the major cell lineages (such as T cell, B cell, erythroid cell, eosinophil/basophil, and dendritic cell populations) remained largely unchanged in Col1a1^{fspKO} mouse bone marrow (Fig. 7a, b). Interestingly, Col1a1^{fspKO} mouse bone marrow revealed significantly increased percentage of a specific neutrophil

Fig. 4 The only live offspring of Col1a1fapKO genotype exhibits severe skeletal defects. a Radiograph of three-week-old wild-type (WT, *Fap-Cre-negative;Col1a1[loxP/loxP]*) and Col1a1[fapKO] mice. **b** H&E staining of the forelimb whole-mount sections of 3-week-old WT (*n* = 3) and Col1a1[fapKO] (*n* = 1, limited by the fact that only one live Col1a1[fapKO] mouse was ever observed) mice. **c** H&E staining, Masson's Trichrome staining (MTS), and Safranin-O/Fast Green staining on serial sections of the forelimbs from three-week-old WT (*n* = 3) and Col1a1[fapKO] (*n* = 1, limited by the fact that only one live Col1a1[fapKO] mouse was ever observed) mice. **d** Quantification of % positive area for MTS staining (blue indicating collagen deposition) and Safranin-O/Fast Green staining (red indicating cartilage) based on the staining shown in (**c**). Quantification of % positive area for staining was based on 10 bone samples from WT group (*n* = 3 mice) and Col1a1[fapKO] group (*n* = 1 mouse, limited by the fact that only one live Col1a1[fapKO] mouse was ever observed). The unpaired, two-tailed *t* test was used to compare the mean of two independent groups. *P = 0.0116, ***P = 1.11E-06. Data are represented as mean ± SEM. **e** Representative H&E and immunofluorescence images of humerus serial sections from *Fap-Cre;LSL-tdTomato* mice (*n* = 3 mice at 2-month-old) stained for osteoblast marker osteocalcin (green) and *Fap-Cre*-induced tdTomato (red). Scale bars, 50 μm. Scale bars of whole-mount sections, 5 mm; Scale bars at 50× magnification, 500 μm; Scale bars at 200× magnification, 100 μm.

cluster ("Neutrophil-1" cluster, Fig. 7a, b) when compared to wild-type littermate bone marrow. This uniquely increased neutrophil cluster revealed high expression of matrix metalloproteinase 9 (MMP9) (Fig. 7c, d), which has been previously shown to regulate bone fracture repair[26,40,41]. This MMP9-high Neutrophil-1 subpopulation highly expressed *Mmp8*, *Cxcr2*, and *Retnlg* (Fig. 7c, d, and Supplementary Fig. 8d, e), as well as the general neutrophil markers such as *S100a8*, *Ly6g*, and *Itgam/CD11b* (Supplementary Fig. 8d). The top upregulated genes in the Neutrophil-1 cluster of Col1a1[fspKO] bone marrow (in comparison with WT bone marrow) were largely anti-inflammatory neutrophil marker genes including *Ly6g*, *Ly6c1*, *Ly6c2*, and *Cxcr2* (Fig. 7e, f). The significant enrichment of Ly6G⁺ neutrophils in Col1a1[fspKO] bone marrow was also confirmed by immunofluorescence staining (Fig. 7g). Ingenuity Pathway Analysis (IPA) of signature genes in the Neutrophil-1 cluster (Supplementary Fig. 8e) revealed the correlation with neutrophil degranulation/activation. Additionally, the monocyte/macrophage cluster of Col1a1[fspKO] mouse bone marrow also revealed a noticeable shift towards Adgre1 (F4/80)-expressing subpopulation (Supplementary Fig. 8f), while the overall percentage of monocyte/macrophage cluster exhibiting a mild increase.

In conclusion, these results suggest an involvement of neutrophils and monocytes/macrophages in response to the bone fractures during early OI pathogenesis in Col1a1[fspKO] mice[42–46].

Normal phenotype of mice with specific deletion of Col1 in αSMA⁺ mesenchymal cells or Cdh5⁺ endothelial cells. Additionally, we also generated the *αSMA-Cre;Col1a1[loxP/loxP]* (Col1a1[smaKO]) and *Cdh5-Cre;Col1a1[loxP/loxP]* (Col1a1[cdh5KO]) mouse models, with the potential to delete Col1 in αSMA⁺ mesenchymal cells and Cdh5(VE-Cadherin)⁺ endothelial cells, respectively. Both Col1a1[smaKO] and Col1a1[cdh5KO] mice were born at the expected Mendelian ratio, when compared to their littermate controls (Supplementary Fig. 9a, b). These mice exhibit normal tissue throughout the body, normal litter size, and unaltered life span, without any overt phenotype. Fractures were not observed in either Col1a1[smaKO] or Col1a1[cdh5KO] mice at 12-month age (Supplementary Fig. 9c, d). The body weight of Col1a1[smaKO] or Col1a1[cdh5KO] mice at 12-month age was identical to that of their wild-type littermates (Supplementary Fig. 9e, f). These results were not surprising, given that αSMA⁺ mesenchymal cell populations and Cdh5⁺ endothelial cells in the bones showed minimal *Col1a1* expression (Fig. 1d). These results indicate that the contribution of αSMA⁺ cells and Cdh5⁺ cells to Col1 production in the bones is not rate limiting. Although Col1a1[smaKO] mice did not reveal spontaneous bone fractures, profound accumulation of αSMA-expressing cells in periosteal area was observed (Supplementary Fig. 9g) in the spontaneous bone lesions of Col1a1[fspKO] mice, consistent with previous studies showing the involvement in fracture repair. Future studies
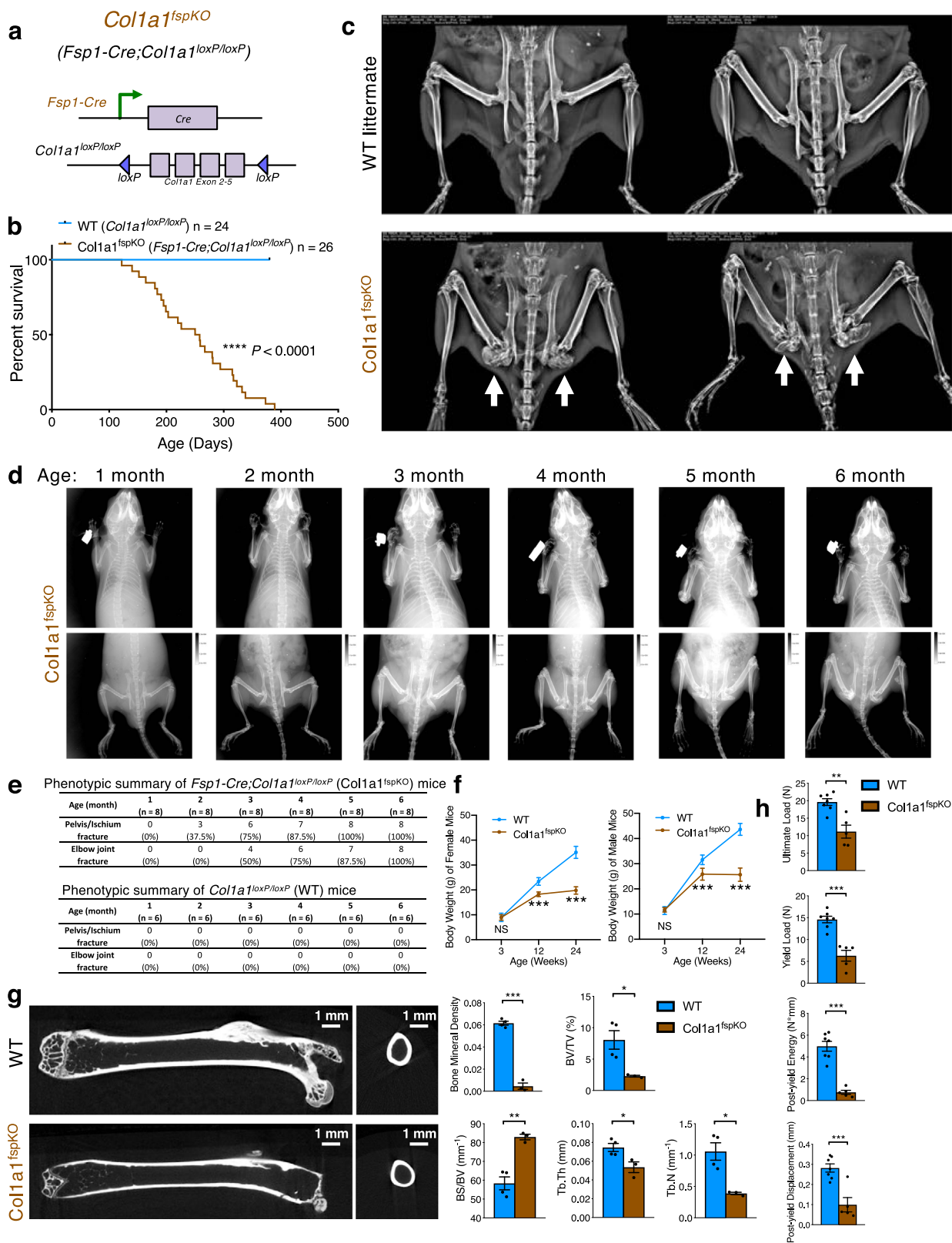
will hopefully unravel whether Col1a1[smaKO] mice may exhibit any bone healing deficiency[47–50].

Discussion

Type I collagen (Col1) is the most abundant protein in the human body, especially in the bones. Over 25,000 publication have examined the role of Col1 in different aspects of biology in health and disease, but it remained unknown until this study what cellular lineages contribute to the synthesis and deposition of Col1 during embryogenesis and bone development[3,4]. Osteoblasts and fibroblasts, differentiated from mesenchymal stem/stromal cells (MSCs), are speculated to be the major source of Col1 in various organs. However, the precise contributions of Col1 by different cell lineages in the formation and development of bones remain unknown. In this regard, previous dataset[27] coupled with current analysis show that Col1 is predominantly produced by Fap⁺ and Fsp1⁺ mesenchymal cells in the bone. Based on this data, we generated novel genetically engineered mouse models (GEMMs) for the deletion of Col1 specifically in Fap⁺ or Fsp1⁺ cells. A *Col1a1[loxP/loxP]* mouse strain was generated and crossed with *CMV-Cre*, *Fap-Cre*, *Fsp1 (S100a4)-Cre*, *Cdh5-Cre*, and *αSMA (Acta2)-Cre* mice to probe the origin and function of Col1 during embryogenesis and bone development.

Our results demonstrate that systemic (whole-body) deletion of *Col1a1* (exons 2–5 truncation) employing *CMV-Cre* mice (Col1a1[cmvKO]) leads to arrested embryonic development and embryonic lethality around embryonic day 6–7 (E6-7), underscoring the essential role of Col1 during gastrulation stage. Deletion of Col1 in Fap⁺ cells leads to embryonic lethality at around E16 associated with several skeletal and vascular defects. This observation is in alliance with previous studies with Fap⁺ cells identified as mesenchymal cells surrounding the vasculature and in the skeletal structures[51,52]. These observations suggest the essential role of Fap⁺ cells in the synthesis of Col1 during the embryonic development of the bones. The observed phenotype of Col1a1[fapKO] mice is analogous to the human type II Osteogenesis Imperfecta, which is lethal in the perinatal period due to critical genetic defects in *COL1A1* or *COL1A2* resulting in multiple rib and/or long bone fractures[53–55].

Many different studies identified Fsp1⁺ cells in the developing skeletal tissue during embryogenesis[35,56,57]. Specific deletion of Col1 in Fsp1⁺ cells leads to an Osteogenesis Imperfecta-like phenotype and with most Col1a1[fspKO] mice dying at the age of 5–12 months. The phenotype of Col1a1[fspKO] mice is comparable to the human type III Osteogenesis Imperfecta, and relevant mouse models previously reported[6,13–15,19,38,58–60]. While previous models of OI were based on systemic Col1 mutation and/or deletion, our studies identity the precise Fsp1 cell lineage that is responsible for Col1 production in maturing bones starting about three weeks of age, and absence of such synthesis by the Fsp1⁺ cells leads to abnormal bones, resulting in easy fractures with minimal trauma. Our studies collectively suggest that early skeletal development during

a Col1a1^fspKO (Fsp1-Cre;Col1a1^loxP/loxP)

b WT (Col1a1^loxP/loxP) n = 24
Col1a1fspKO (Fsp1-Cre;Col1a1^loxP/loxP) n = 26
**** P < 0.0001

c WT littermate / Col1a1^fspKO

d Age: 1 month, 2 month, 3 month, 4 month, 5 month, 6 month — Col1a1^fspKO

e Phenotypic summary of Fsp1-Cre;Col1a1^loxP/loxP (Col1a1^fspKO) mice

Age (month)	1 (n = 8)	2 (n = 8)	3 (n = 8)	4 (n = 8)	5 (n = 8)	6 (n = 8)
Pelvis/Ischium fracture	0 (0%)	3 (37.5%)	6 (75%)	7 (87.5%)	8 (100%)	8 (100%)
Elbow joint fracture	0 (0%)	0 (0%)	4 (50%)	6 (75%)	7 (87.5%)	8 (100%)

Phenotypic summary of Col1a1^loxP/loxP (WT) mice

Age (month)	1 (n = 6)	2 (n = 6)	3 (n = 6)	4 (n = 6)	5 (n = 6)	6 (n = 6)
Pelvis/Ischium fracture	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Elbow joint fracture	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

f Body Weight (g) of Female Mice / Body Weight (g) of Male Mice — WT, Col1a1^fspKO; Age (Weeks) 3, 12, 24; NS, ***

g WT / Col1a1^fspKO micro-CT; Bone Mineral Density ***; BV/TV (%) *; BS/BV (mm⁻¹) **; Tb.Th (mm) *; Tb.N (mm⁻¹) *

h WT, Col1a1^fspKO; Ultimate Load (N) **; Yield Load (N) ***; Post-yield Energy (N*mm) ***; Post-yield Displacement (mm) ***

embryogenesis and the early postnatal period might rely on Fap⁺ cells for the production of Col1, and at a later stage the role of Fsp1⁺ cells becomes prominent. Deletion of Col1 from αSMA⁺ cells did not result in any adversity and mice were normal with expected lifespan. Same is true when Col1 was deleted in Cdh5⁺ endothelial cells (a common biomarker for most endothelial cells). This experimental arm served as an important control, because

some reports suggest that angiogenic endothelial cells during embryogenesis and organogenesis are shown to produce type I collagen. All together, these experiments suggest that αSMA⁺ mesenchymal cells do not compensate for the loss of Col1 from Fap or Fsp1 cell lineages. Moreover, the functions of Fap⁺ cells and Fsp1⁺ cells are mutually exclusive in the context of embryogenesis and bone formation.

Fig. 5 Specific deletion of Col1 in Fsp1-lineage cells causes Osteogenesis Imperfecta phenotype. a Genetic strategy to delete *Col1a1* specifically in Fsp1-lineage cells by crossing the *Col1a1^loxP/loxP^* mice with *Fsp1-Cre* mice. **b** Survival of wild-type (WT; with genotype of *Fsp1-Cre-negative;Col1a1^loxP/loxP^*) and Col1a1^fspKO^ (*Fsp1-Cre;Col1a1^loxP/loxP^*) mice (*n* = 24 and 26, respectively). Kaplan-Meier plots and the log rank Mantel-Cox test were used to evaluate statistical differences of survival. ****P < 0.0001. **c** Micro-CT scanning of 6-month-old WT and Col1a1^fspKO^ mice (with white arrows indicating the fractures in pelvis and ischium of Col1a1^fspKO^ mice). **d** Radiograph of Col1a1^fspKO^ mice at the age from 1 to 6 month. Radiograph of age-matched WT littermates was shown in Supplementary Fig. 7a. **e** Summary of the occurrence of fractures in pelvis, ischium, and/or joints, as examined by radiograph, of WT (*n* = 6) and Col1a1^fspKO^ (*n* = 8) mice at the age from 1 to 6 month. **f** Body weight of WT and Col1a1^fspKO^ mice (*n* = 6 mice per group) at the age of 3, 12, and 24 month. The unpaired, two-tailed *t* test was used to compare the mean of two independent groups. ***P = 9.57E-05 (female 12 weeks), 7.15E-07 (female 24 weeks), 0.00096 (male 12 weeks), 2.35E-07 (male 24 weeks). Data are represented as mean ± SEM. **g** The micro-CT images of trabecular and cortical bones of femurs from 6-month-old WT mice (*n* = 4) and Col1a1^fspKO^ mice (*n* = 3). The trabecular bone material property measurements are compared as bone mineral density (BMD; ***P = 1.04E-05), bone volume/total volume (BV/TV; * P = 0.0289), bone surface area/bone volume (BS/BV; **P = 0.00212), trabecular thickness (Tb.Th; *P = 0.0287), and trabecular number (Tb.N; *P = 0.0160). The unpaired, two-tailed *t* test was used to compare the mean of two independent groups. Data are represented as mean ± SEM. The cortical bone material property measurements are continued in Supplementary Fig. 7d. **h** Biomechanical (three-point bending) test results of femurs from 6-month-old WT mice (*n* = 7) and Col1a1^fspKO^ mice (*n* = 5) were presented as ultimate load (**P = 0.00135), yield load (***P = 0.00012), post-yield energy (***P = 2.01E-05), and post-yield displacement (***P = 0.00077). The unpaired, two-tailed *t* test was used to compare the mean of two independent groups. Data are represented as mean ± SEM.

These results also emphasize the crucial contribution of Col1 specifically produced by Fsp1+ osteoblasts and fibroblasts (in bone marrow and periosteum) to the structural maintenance and repair of the bones, which is consistent with previous notion that fibroblasts appear at fracture site and are responsible for collagen production/crosslinking upon encountering hematoma[61]. Our results also suggest that the Col1 deposited by fibroblasts of Fsp1-lineage is essential for the formation of hematoma, which further impacts the subsequent steps leading to the repair of fractures including osteoblast differentiation and bone formation in the fractured callus[62–64]. The identification of the Fsp1+ cells as critical and rate-limiting source of Col1 in developing bone and adult bone undergoing repair and regeneration provides new insights into potential cell-based therapy for Osteogenesis Imperfecta. Future studies may examine whether Fsp1+ cells may provide cell-based therapy opportunities for Osteogenesis Imperfecta via transplantation.

Taken together, our results demonstrate that Col1 produced by Fap+ cells is essential for the embryonic development of skeletal system, while Col1 produced by Fsp1+ cells is critical for the postnatal maintenance and repair of the bones. Utilizing multiple transgenic mouse models and single-cell RNA-sequencing analyses, this study for the first time provided comprehensive investigations on the functional roles of Col1 with respect to its specific cellular origin during embryogenesis, osteogenesis, and bone pathogenesis.

Methods
Mice. All mice were housed under standard housing conditions at MDACC animal facilities, and all animal procedures were reviewed and approved by the MDACC Institutional Animal Care and Use Committee.

Fsp1-Cre[65,66] and *αSMA-Cre*[67] mouse strains were previously documented. *Cdh5-Cre* (017968; B6.129-Tg-Cdh5-cre-1Spe/J), *CMV-Cre* (006054; B6.C-Tg-CMV-cre-1Cgn/J), and *LSL-YFP* (006148; B6.129 × 1-Gt-ROSA-26Sortm1-EYFP-Cos/J) mouse strains were purchased from Jackson Laboratory. Osteogenesis imperfecta murine (OIM) strain harboring *Col1a2* mutation (001815; B6C3Fe a/a-Col1a2oim/J) was purchased from Jackson Laboratory. This *Col1a2* spontaneous mutation is a single nucleotide deletion (G) that alters the terminal approximately 50 amino acids of the pro-alpha 2 C-propeptide and prevents association with the pro-alpha1 chains. *Col1a1^loxP/loxP^* mouse strain (with *loxP*-flanked exons 2-5) was established in the Genetically Engineered Mouse Facility at MD Anderson Cancer Center (MDACC) using the Col1a1^tm1a(EUCOMM)Wtsi^ embryonic stem cells that were obtained from the European Mouse Mutant Cell Repository (EuMMCR)[68]. Both female and male mice with desired genotype(s) were used for experimental mice.

Genotyping. Tail DNA was used for genotyping PCR analysis of animals and embryos older than E14.5. Yolk sac DNA was used for genotyping PCR of embryos at earlier stages. Recombination PCR of *Col1a1^loxP/loxP^* allele on DNA from multiple major organs of all experimental mice was routinely performed to exclude the possibility of germline recombination.

Histology and immunohistochemistry. For paraffin-fixed samples, mouse tissues were fixed in 10% neutral buffered formalin, embedded in paraffin, and sectioned at 5 μm thickness. Sections were processed for hematoxylin and eosin (H&E) staining. Masson's trichrome stain (MTS) was conducted using Gomori's Tri-chrome Stain Kit (38016SS2, Leica Biosystems). Picrosirius red staining for collagen was conducted using 0.1% Picrosirius Red (Direct Red80; Sigma-Aldrich) and counterstained with Weigert's haematoxylin. Images were captured with a Leica DM 1000 LED microscope and an MC120 HD Microscope Camera with LAS V4.4 Software (Leica). Safranin-O/Fast Green staining was conducted using 0.1% Safranin-O solution and 0.1% Fast Green solution (Electron Microscopy Sciences).

Formalin-fixed, paraffin-embedded sections were processed for immunohistochemical staining as previously documented[69,70]. Sections were incubated with primary antibody against type I collagen (Col1; SouthernBiotech, 1310-01, lot#B2918-XB89, 1:200), followed by biotinylated donkey anti-goat secondary antibody (Jackson ImmunoResearch, 705-066-147, lot#127411, 1:400), and streptavidin HRP (Biocare Medical). For all immunolabeled experiments, sections were developed by DAB and counterstained with hematoxylin. Immunofluorescence staining on formalin-fixed, paraffin-embedded sections was conducted using the following primary antibodies: rabbit anti-GFP/YFP (Abcam, ab290, lot#GR3321614-1, 1:500), chicken anti-GFP/YFP (Abcam, ab13970, lot# GR3190550-13, 1:400), rabbit anti-osteocalcin/OCN (Millipore/Sigma-Aldrich, ab10911, lot#3670782, 1:200), rat anti-Ly6G (Abcam, ab25377, clone RB6-8C5, lot#GR3391587-2, 1:100), and mouse anti-RFP/tdTomato (Thermo Fisher, MA5-15257, clone RF5R, lot#SD247965, 1:100). The staining was then followed by AF488/594-labeled secondary antibodies: Thermo Fisher A11012 (lot#1933366), A11039 (lot#2079383), A11008 (lot#1937184), and A21203 (lot#1722995), all in 1:400 dilution. Slides were then mounted with DAPI-containing Vectashield Mounting Medium (Vector Laboratories), visualized under the LSM800 confocal laser scanning microscope, and analyzed with ZEN software version 2.1 (Zeiss).

Radiography (X-ray) and micro computed tomography (micro-CT) Scanning. Imaging operations were conducted at MDACC Small Animal Imaging Facility. Radiography of mice was conducted using the Bruker Xtreme system. Micro-CT scanning was performed using the Bruker SkyScan 1276 micro-CT system (Bruker BioSpin Corporation). The scanning parameters are: resolution: 13 μm; pixel size: 13.083 μm; voltage: 55 kV; current: 200 μA; rotation step (deg): 0.400; duration: 11 min; exposure time: 500 ms; filter: 0.25 mm Al; matrix: 2016 ×1344. The femurs of 6-month-old Col1a1^fspKO^ mice and wild-type (WT) littermates were scanned and measured for the bone material properties. The trabecular bone parameters were measured in the distal metaphysis of the femurs, while the cortical bone parameters were measured at the femur mid-diaphysis. The detected bone parameters included trabecular or cortical bone mineral density (BMD), bone volume/total volume (BV/TV), trabecular or cortical bone surface area/bone volume (BS/BV), trabecular thickness (Tb.Th), trabecular number (Tb.N), cortical bone total area (Tt.Ar), cortical bone area (Ct.Ar), cortical bone marrow area (Ma.Ar),), cortical bone area/total area (Ct.Ar/Tt.Ar), and cortical thickness (Ct.Th). The analysis was conducted using Bruker SkyScan DataViewer software version 1.5.0.

Biomechanical (three-point bending) examination. The femurs from Col1a1^fspKO^ mice and WT littermates of 6-month age were freshly collected and tested at room temperature using an Instron 5848 microtester (Instron Inc.), as previously documented[17]. The femurs were examined to failure in three-point bending test at a rate of 0.1 mm/s. A 100 N load cell was used to collect data. Load and displacement data were captured at a rate of 40 Hz by Bluehill Software version 2.9 (Instron). The span length used in the examination was 7.3 mm. The maximum load was determined by the highest load value recorded before the specimen fractured.
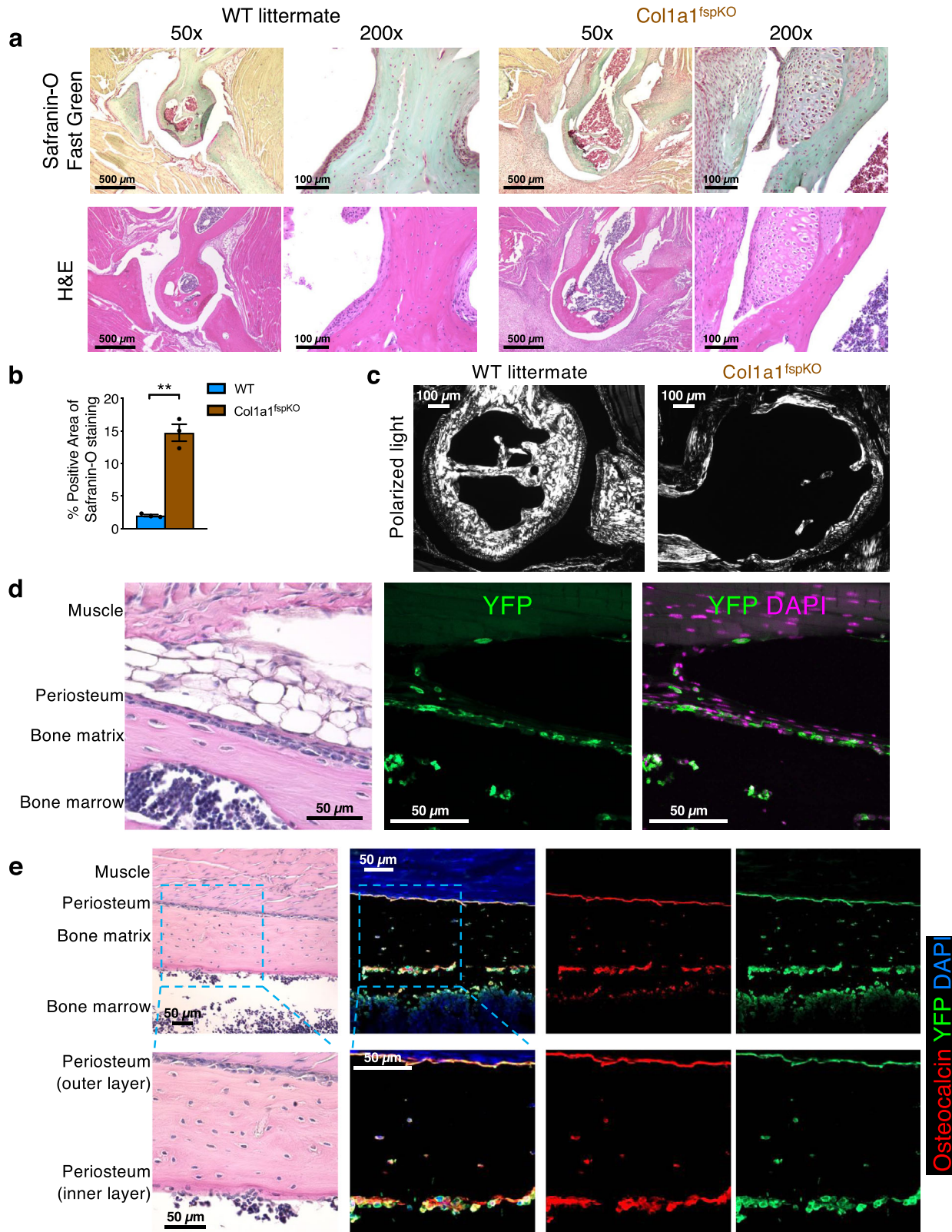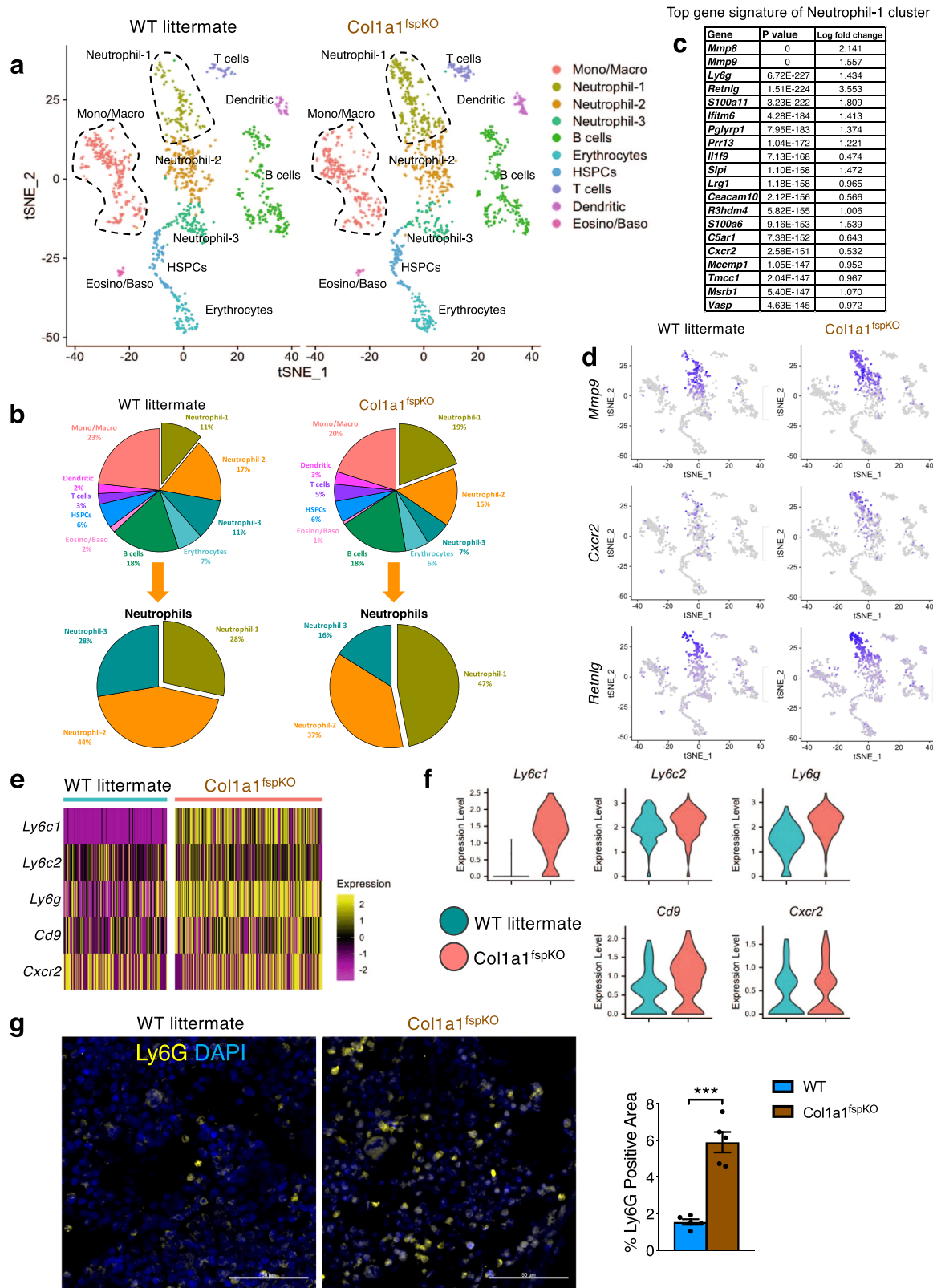
Fig. 6 Histological analysis of bone tissues from Col1a1fspKO mice and Fsp1-Cre;LSL-YFP mice. a, b H&E staining and Safranin-O/Fast Green staining of the forelimb elbow joints of 6-month-old WT and Col1a1[fspKO] mice (**a**). Quantification of % positive area for Safranin-O/Fast Green staining was based on three mice per group (**b**). The unpaired, two-tailed t test was used to compare the mean of two independent groups. **$P = 0.00946$. Data are represented as mean ± SEM. Scale bars at 50× magnification, 500 μm; Scale bars at 200× magnification, 100 μm. **c** Picrosirius Red staining of collagen content in the forelimb elbow joints from six-month-old WT and Col1a1[fspKO] mice ($n = 3$ mice per group) examined by circularly polarized light microscopy. Scale bars, 100 μm. **d** Representative immunofluorescence images of humerus serial sections from lineage-tracing *Fsp1-Cre;LSL-YFP* mice ($n = 3$ mice at two-month-old) stained for *Fsp1-Cre*-induced YFP. Scale bars, 50 μm. **e** Representative H&E and immunofluorescence images of humerus serial sections from *Fsp1-Cre;LSL-YFP* mice ($n = 3$ mice at two-month-old) stained for osteoblast marker osteocalcin (red) and *Fsp1-Cre*-induced YFP (green). Scale bars, 50 μm.

c Top gene signature of Neutrophil-1 cluster

Gene	P value	Log fold change
Mmp8	0	2.141
Mmp9	0	1.557
Ly6g	6.72E-227	1.434
Retnlg	1.51E-224	3.553
S100a11	3.23E-222	1.809
Ifitm6	4.28E-184	1.413
Pglyrp1	7.95E-183	1.374
Prr13	1.04E-172	1.221
Il1f9	7.13E-168	0.474
Slpi	1.10E-158	1.472
Lrg1	1.18E-158	0.965
Ceacam10	2.12E-156	0.566
R3hdm4	5.82E-155	1.006
S100a6	9.16E-153	1.539
C5ar1	7.38E-152	0.643
Cxcr2	2.58E-151	0.532
Mcemp1	1.05E-147	0.952
Tmcc1	2.04E-147	0.967
Msrb1	5.40E-147	1.070
Vasp	4.63E-145	0.972

Fig. 7 Single-cell RNA-sequencing analysis of bone marrow fractions from Col1a1fspKO mice. a–c Single-cell RNA-sequencing analysis of cell mixture of bone marrow fractions from 2-month-old WT and Col1a1[fspKO] mice (*n* = 2 per group). Functional clusters of cells were defined with group definition listed, as implemented in the Seurat R package (**a**). Mono/Macro: monocyte/macrophage cluster. HSPCs: hematopoietic stem and progenitor cells. Dendritic: dendritic cell cluster. Eosino/Baso: eosinophil/basophil cluster. **b** Pie charts showing the percentage of each of the functional cell clusters in bone marrow fractions from WT and Col1a1[fspKO] mice. Total neutrophil clusters containing 3 subpopulations (1–3) were further plotted as pie charts. (**c**) List of top signature genes of "Neutrophil-1" cluster shown in (**a**) and (**b**), presented with *P* value (based on the non-parameteric Wilcoxon Rank Sum test) and Log2 fold change. Continued in Supplementary Fig. 8a–c. **d** Expression profile of *Mmp9*, *Cxcl12*, and *Retnlg* among defined cell clusters in the bone marrow fractions from WT and Col1a1[fspKO] mice shown in UMAP plot. Continued in Supplementary Fig. 8d. **e**, **f** Top upregulated genes of Neutrophil-1 cluster in Col1a1[fspKO] mice than WT mice shown in heat map plot (**e**) or violin plot (**f**). **g** Representative immunofluorescence images and quantification results of humerus bone marrow from two-month-old WT and Col1a1[fspKO] mice (*n* = 5 per group) stained for Neutrophil-1 marker Ly6G (yellow). Scale bars, 50 μm. The unpaired, two-tailed *t* test was used to compare the mean of two independent groups. ***P = 0.00096. Data are represented as mean ± SEM.

manufacturer's guidelines. Cells from unfractionated tumor were encapsulated using 10X Genomics' Chromium controller and Single Cell 3′ Reagent Kits v2. Following capture and lysis, cDNA was synthesized and amplified to construct Illumina sequencing libraries. The libraries from about 1,000 cells per sample were sequenced with Illumina Nextseq 500. The run format was 26 cycles for read 1, 8 cycles index 1, and 124 cycles for read 2. sc-RNA-seq data was processed by the Sequencing and Microarray Facility in MD Anderson Cancer Center. Library Seurat (version 3.6.1), dplyr (version 1.0.7) and cowplot (version 1.1.1) were loaded into R (version 3.6.1) to explore QC metrics, filter cells, normalize data, cluster cells, and identify cluster bio-markers. To filter out low-quality cells, a threshold with a minimum of 200 and a maximum of 7000 genes per cell was used. Cells with more than 10% of the mito-chondrial genome were also removed for further analysis. "RunTSNE" function was used for clustering the cells. Based on the "JackStrawPlot" and "ElbowPlot" functions, the first 27 principal components were used for TSNE projection and clustering ana-lysis. "FindAllMarkers" function was used to identify the specific markers for each cell cluster. "DoHeatmap" function was used to show the top 10 genes in each cluster. "VlnPlot" function was used to show expression probability distributions across cell clusters of the genes we selected to assign the cell type identity, and the genes that we were interested in.

The previous published single-cell RNA-sequencing (sc-RNA-seq) dataset on bone and bone marrow cell populations of healthy mice[27] were retrieved from the Gene Expression Omnibus (GEO) repository (GSE128423). As previously described[27], femur and tibia bones of healthy C57BL/6 mice were collected and separated into bone and bone marrow fractions. Stromal cells from the bone marrow were isolated after STEMxyme1/DispaseII digestion. Stromal cells from the bone fraction were isolated after STEMxyme1/DispaseII digestion of the cut/crushed bone fragments. Cell mixtures were then filtered through 70 μm cell strainers, lysed with ACK-lysis buffer, and prepared for FACS cell sorting. Bone marrow stromal cells were enriched by sorting of live cells that are negative for erythroid markers (CD71/Ter119) and immune lineage markers (CD45/CD3/B220/CD19/Gr-1/CD11b). This dataset was analyzed using a similar methodology as the above sc-RNA-seq data generated in this study. A total of 12530 cells from bone fraction and 14875 cells from bone marrow fraction were analyzed. Functional clusters of cells were defined with indicated group definition, as implemented in the Seurat R package. "RunUMAP" function was used for clustering the cells. "FindAllMarkers" function was used to identify the specific markers for each cell cluster. "FindMarkers" function was used to compare the differentially expressed marker genes for indicated cell cluster between control and knockout groups. "DoHeatmap" function was used to show the top 10 genes in each cluster. "VlnPlot" function was used to show expression probability distributions across cell clusters of the genes we selected to assign the cell type identity, and the genes that we interested in.

Statistics. Statistical analyses were performed with unpaired, two-tailed *t* test, or one-way ANOVA with Tukey's multiple comparison test using GraphPad Prism version 8.0.0 (GraphPad Software, San Diego, CA, USA). Kaplan–Meier plots were drawn for survival analysis and the log rank Mantel-Cox test was used to evaluate statistical differences. Data met the assumptions of each statistical test, where variance was not equal (determined by an *F*-test) Welch's correction for unequal variances was applied. A *P* value <0.05 was considered statistically significant. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, NS not significant. Error bars represented standard error of the mean (S.E.M.) when multiple visual fields were averaged to produce a single value for each animal, which was then averaged again to represent the mean bar for the group in each graph.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The authors declare that all data supporting the findings of this study are available within the paper and the supplementary information. The single-cell RNA-sequencing source data generated by this study have been deposited in Gene Expression Omnibus (GEO)

database with the accession number GSE184360. The single-cell RNA-sequencing source data of a previously published dataset[27] have been deposited in Gene Expression Omnibus (GEO) database with the accession number GEO GSE128423. Source data are provided with this paper.

References
1. Hay, E. D. Extracellular matrix. *J. cell Biol.* **91**, 205s–223s (1981).
2. Lohler, J., Timpl, R. & Jaenisch, R. Embryonic lethal mutation in mouse collagen I gene causes rupture of blood vessels and is associated with erythropoietic and mesenchymal cell death. *Cell* **38**, 597–607 (1984).
3. Zaidi, M. Skeletal remodeling in health and disease. *Nat. Med.* **13**, 791–801 (2007).
4. Marini, J. C. et al. Osteogenesis imperfecta. *Nat. Rev. Dis. Prim.* **3**, 17052 (2017).
5. Rauch, F. & Glorieux, F. H. Osteogenesis imperfecta. *Lancet* **363**, 1377–1385 (2004).
6. Chipman, S. D. et al. Defective pro alpha 2(I) collagen synthesis in a recessive mutation in mice: a model of human osteogenesis imperfecta. *Proc. Natl Acad. Sci. USA* **90**, 1701–1705 (1993).
7. Kuznetsova, N., McBride, D. J. Jr. & Leikin, S. Osteogenesis imperfecta murine: interaction between type I collagen homotrimers. *J. Mol. Biol.* **309**, 807–815 (2001).
8. Miles, C. A., Sims, T. J., Camacho, N. P. & Bailey, A. J. The role of the alpha2 chain in the stabilization of the collagen type I heterotrimer: a study of the type I homotrimer in oim mouse tissues. *J. Mol. Biol.* **321**, 797–805 (2002).
9. Kuznetsova, N. V., McBride, D. J. & Leikin, S. Changes in thermal stability and microunfolding pattern of collagen helix resulting from the loss of alpha1(I) chain in osteogenesis imperfecta murine. *J. Mol. Biol.* **331**, 191–200 (2003).
10. Kuznetsova, N. V., Forlino, A., Cabral, W. A., Marini, J. C. & Leikin, S. Structure, stability and interactions of type I collagen with GLY349-CYS substitution in alpha 1(I) chain in a murine Osteogenesis Imperfecta model. *Matrix Biol.: J. Int. Soc. Matrix Biol.* **23**, 101–112 (2004).
11. Schnieke, A., Harbers, K. & Jaenisch, R. Embryonic lethal mutation in mice induced by retrovirus insertion into the alpha 1(I) collagen gene. *Nature* **304**, 315–320 (1983).
12. Bonadio, J. et al. A murine skeletal adaptation that significantly increases cortical bone mechanical properties. Implications for human skeletal fragility. *J. Clin. Investig.* **92**, 1697–1705 (1993).
13. Lisse, T. S. et al. ER stress-mediated apoptosis in a new mouse model of osteogenesis imperfecta. *PLoS Genet* **4**, e7 (2008).
14. Chen, F. et al. First mouse model for combined osteogenesis imperfecta and Ehlers-Danlos syndrome. *J. Bone Min. Res.* **29**, 1412–1423 (2014).
15. Besio, R. et al. Early fracture healing is delayed in the Col1a2(+/G610C) osteogenesis imperfecta murine model. *Calcif. Tissue Int* **103**, 653–662 (2018).
16. Daley, E. et al. Variable bone fragility associated with an Amish COL1A2 variant and a knock-in mouse model. *J. Bone Min. Res.* **25**, 247–261 (2010).
17. Zieba J., et al. Fracture healing in collagen-related preclinical models of osteogenesis imperfecta. *J Bone Miner Res.* **35**, 1132–1148 (2020).
18. Morello, R. et al. CRTAP is required for prolyl 3- hydroxylation and mutations cause recessive osteogenesis imperfecta. *Cell* **127**, 291–304 (2006).
19. Liu, Y. et al. A novel transgenic murine model with persistently brittle bones simulating osteogenesis imperfecta type I. *Bone* **127**, 646–655 (2019).
20. Florencio-Silva, R., Sasso, G. R., Sasso-Cerri, E., Simoes, M. J. & Cerri, P. S. Biology of bone tissue: structure, function, and factors that influence bone cells. *Biomed. Res Int.* **2015**, 421746 (2015).
21. Friedenstein, A. J., Chailakhyan, R. K., Latsinik, N. V., Panasyuk, A. F. & Keiliss-Borok, I. V. Stromal cells responsible for transferring the

microenvironment of the hemopoietic tissues. Cloning in vitro and retransplantation in vivo. *Transplantation* **17**, 331–340 (1974).

22. Pittenger, M. F. et al. Multilineage potential of adult human mesenchymal stem cells. *Science* **284**, 143–147 (1999).
23. Nakahara, H., Goldberg, V. M. & Caplan, A. I. Culture-expanded human periosteal-derived cells exhibit osteochondral potential in vivo. *J. Orthop. Res.* **9**, 465–476 (1991).
24. Friedenstein, A. J., Chailakhyan, R. K. & Gerasimov, U. V. Bone marrow osteogenic stem cells: in vitro cultivation and transplantation in diffusion chambers. *Cell Tissue Kinet.* **20**, 263–272 (1987).
25. Friedenstein, A. J., Chailakhjan, R. K. & Lalykina, K. S. The development of fibroblast colonies in monolayer cultures of guinea-pig bone marrow and spleen cells. *Cell Tissue Kinet.* **3**, 393–403 (1970).
26. Bohm, A. M. et al. Activation of skeletal stem and progenitor cells for bone regeneration is driven by PDGFRbeta signaling. *Developmental cell* **51**, 236–254 e212 (2019).
27. Baryawno, N. et al. A cellular taxonomy of the bone marrow stroma in homeostasis and leukemia. *Cell* **177**, 1915–1932 e1916 (2019).
28. Baccin, C. et al. Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. *Nat. cell Biol.* **22**, 38–48 (2020).
29. Ding, L. & Morrison, S. J. Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature* **495**, 231–235 (2013).
30. Greenbaum, A. et al. CXCL12 in early mesenchymal progenitors is required for haematopoietic stem-cell maintenance. *Nature* **495**, 227–230 (2013).
31. Delmas, P. D., Demiaux, B., Malaval, L., Chapuy, M. C. & Meunier, P. J. [Osteocalcin (or bone gla-protein), a new biological marker for studying bone pathology]. *Presse Med* **15**, 643–646 (1986).
32. Huang, W., Yang, S., Shao, J. & Li, Y. P. Signaling and transcriptional regulation in osteoblast commitment and differentiation. *Front. Biosci.: a J. virtual Libr.* **12**, 3068–3092 (2007).
33. Doege, K. J., Sasaki, M., Kimura, T. & Yamada, Y. Complete coding sequence and deduced primary structure of the human cartilage large aggregating proteoglycan, aggrecan. Human-specific repeats, and additional alternatively spliced forms. *J. Biol. Chem.* **266**, 894–902 (1991).
34. Chan, D., Cole, W. G., Chow, C. W., Mundlos, S. & Bateman, J. F. A COL2A1 mutation in achondrogenesis type II results in the replacement of type II collagen by type I and III collagens in cartilage. *J. Biol. Chem.* **270**, 1747–1753 (1995).
35. Strutz, F. et al. Identification and characterization of a fibroblast marker: FSP1. *J. cell Biol.* **130**, 393–405 (1995).
36. Corada, M. et al. Monoclonal antibodies directed to different regions of vascular endothelial cadherin extracellular domain affect adhesion and clustering of the protein and modulate endothelial permeability. *Blood* **97**, 1679–1684 (2001).
37. Armulik, A., Genove, G. & Betsholtz, C. Pericytes: developmental, physiological, and pathological perspectives, problems, and promises. *Developmental cell* **21**, 193–215 (2011).
38. Lim, J., Grafe, I., Alexander, S. & Lee, B. Genetic causes and mechanisms of Osteogenesis Imperfecta. *Bone* **102**, 40–49 (2017).
39. Kamoun-Goldrat, A. S., Le & Merrer, M. F. Animal models of osteogenesis imperfecta and related syndromes. *J. Bone Min. Metab.* **25**, 211–218 (2007).
40. Colnot, C., Thompson, Z., Miclau, T., Werb, Z. & Helms, J. A. Altered fracture repair in the absence of MMP9. *Development* **130**, 4123–4133 (2003).
41. Wang, X. et al. MMP9 regulates the cellular response to inflammation after skeletal injury. *Bone* **52**, 111–119 (2013).
42. Baht, G. S., Vi, L. & Alman, B. A. The role of the immune cells in fracture healing. *Curr. Osteoporos. Rep.* **16**, 138–145 (2018).
43. Einhorn, T. A. & Gerstenfeld, L. C. Fracture healing: mechanisms and interventions. *Nat. Rev. Rheumatol.* **11**, 45–54 (2015).
44. Kovtun, A., Messerer, D. A. C., Scharffetter-Kochanek, K., Huber-Lang, M. & Ignatius, A. Neutrophils in tissue trauma of the skin, bone, and lung: two sides of the same coin. *J. Immunol. Res.* **2018**, 8173983 (2018).
45. Tsukasaki, M. & Takayanagi, H. Osteoimmunology: evolving concepts in bone-immune interactions in health and disease. *Nat. Rev. Immunol.* **19**, 626–642 (2019).
46. Kovtun, A. et al. The crucial role of neutrophil granulocytes in bone fracture healing. *Eur. Cell Mater.* **32**, 152–162 (2016).
47. Kalajzic, Z. et al. Use of an alpha-smooth muscle actin GFP reporter to identify an osteoprogenitor population. *Bone* **43**, 501–510 (2008).
48. Grcevic, D. et al. In vivo fate mapping identifies mesenchymal progenitor cells. *Stem Cells* **30**, 187–196 (2012).
49. Matthews, B. G. et al. Analysis of alphaSMA-labeled progenitor cell commitment identifies notch signaling as an important pathway in fracture healing. *J. Bone Min. Res.* **29**, 1283–1294 (2014).
50. Matthews, B. G. et al. Osteogenic potential of alpha smooth muscle actin expressing muscle resident progenitor cells. *Bone* **84**, 69–77 (2016).

51. Niedermeyer, J. et al. Targeted disruption of mouse fibroblast activation protein. *Mol. Cell Biol.* **20**, 1089–1094 (2000).
52. Niedermeyer, J. et al. Expression of the fibroblast activation protein during mouse embryo development. *Int J. Dev. Biol.* **45**, 445–447 (2001).
53. Sillence, D. O., Barlow, K. K., Garber, A. P., Hall, J. G. & Rimoin, D. L. Osteogenesis imperfecta type II delineation of the phenotype with reference to genetic heterogeneity. *Am. J. Med Genet* **17**, 407–423 (1984).
54. Sillence, D. O., Senn, A. & Danks, D. M. Genetic heterogeneity in osteogenesis imperfecta. *J. Med Genet* **16**, 101–116 (1979).
55. Van Dijk, F. S., Pals, G., Van Rijn, R. R., Nikkels, P. G. & Cobben, J. M. Classification of Osteogenesis Imperfecta revisited. *Eur. J. Med Genet* **53**, 1–5 (2010).
56. Inubushi, T., Nozawa, S., Matsumoto K., Irie F., Yamaguchi Y. Aberrant perichondrial BMP signaling mediates multiple osteochondroma genesis in mice. *JCI insight* **2**, e90049 (2017).
57. Bowen, M. E., Ayturk, U. M., Kurek, K. C., Yang, W. & Warman, M. L. SHP2 regulates chondrocyte terminal differentiation, growth plate architecture and skeletal cell fates. *PLoS Genet* **10**, e1004364 (2014).
58. Camacho, N. P. et al. The material basis for reduced mechanical properties in oim mice bones. *J. Bone Min. Res.* **14**, 264–272 (1999).
59. Phillips, C. L. et al. Oim mice exhibit altered femur and incisor mineral composition and decreased bone mineral density. *Bone* **27**, 219–226 (2000).
60. Cardinal, M. et al. Sclerostin antibody reduces long bone fractures in the oim/oim model of osteogenesis imperfecta. *Bone* **124**, 137–147 (2019).
61. Claes, L., Recknagel, S. & Ignatius, A. Fracture healing under healthy and inflammatory conditions. *Nat. Rev. Rheumatol.* **8**, 133–143 (2012).
62. Jikko, A., Harris, S. E., Chen, D., Mendrick, D. L. & Damsky, C. H. Collagen integrin receptors regulate early osteoblast differentiation induced by BMP-2. *J. Bone Min. Res.* **14**, 1075–1083 (1999).
63. Li, H. et al. Immature osteoblast lineage cells increase osteoclastogenesis in osteogenesis imperfecta murine. *Am. J. Pathol.* **176**, 2405–2413 (2010).
64. Lynch, M. P., Stein, J. L., Stein, G. S. & Lian, J. B. The influence of type I collagen on the development and maintenance of the osteoblast phenotype in primary and passaged rat calvarial osteoblasts: modification of expression of genes supporting cell growth, adhesion, and extracellular matrix mineralization. *Exp. cell Res.* **216**, 35–45 (1995).
65. Xue, C., Plieth, D., Venkov, C., Xu, C. & Neilson, E. G. The gatekeeper effect of epithelial-mesenchymal transition regulates the frequency of breast cancer metastasis. *Cancer Res.* **63**, 3386–3394 (2003).
66. Bhowmick, N. A. et al. TGF-beta signaling in fibroblasts modulates the oncogenic potential of adjacent epithelia. *Science* **303**, 848–851 (2004).
67. LeBleu, V. S. et al. Origin and function of myofibroblasts in kidney fibrosis. *Nat. Med.* **19**, 1047–1053 (2013).
68. Chen, Y. et al. Type I collagen deletion in alphaSMA(+) myofibroblasts augments immune suppression and accelerates progression of pancreatic cancer. *Cancer cell* **39**, 548–565 e546 (2021).
69. Chen Y., et al. Dual reporter genetic mouse models of pancreatic cancer identify an epithelial-to-mesenchymal transition-independent metastasis program. *EMBO molecular medicine* **10**, e9085 (2018).
70. Zheng, X. et al. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* **527**, 525–530 (2015).

Acknowledgements

Author contributions
Conceptualization: Y.C., R.K. Data curation: Y.C. Investigation: Y.C., S.Y., S.L., H.S. Methodology: Y.C., S.Y., C.G.A. Single-cell RNA-Seq Analyses: S.Y., Y.C. Resources: R.K. Supervision: Y.C., R.K. Validation: Y.C. Writing – original draft: Y.C. and R.K. Writing – review & editing: Y.C., S.Y., S.L., K.M.McA., H.S., R.K.

Competing interests
The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-27563-3.

Correspondence and requests for materials should be addressed to Raghu Kalluri.

Peer review information *Nature Communications* thanks Guillaume Mabilleau and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.