A multiomic atlas of human early skeletal development

[1]Ken To*, [1]Lijiang Fei*, [1]Jan Patrick Pett*, [1]Kenny Roberts, [1]Krzysztof Polański, [1]Tong Li, [1,2]Nadav Yayon, [1,2]Peng He, [1]Chuan Xu, [1]James Cranley, [1]Ruoyan Li, [1]Kazumasa Kanemaru, [1]Ni Huang, [1]Stathis Megas, [1]Laura Richardson, [1]Rakesh Kapuge, [1]Shani Perera, [1]Elizabeth Tuck, [1]Anna Wilbrey-Clark, [1]Ilaria Mulas, [1]Fani Memi, [1]Batuhan Cakir, [1]Alexander V. Predeus, [1,3]David Horsfall, [1]Simon Murray, [1]Martin Prete, [1]Pavel Mazin, [4]Xiaoling He, [1]Kerstin B. Meyer, [1,3]Muzlifah Haniffa, [4,5]Roger A. Barker, [1]Omer Bayraktar, [6]Christopher D. Buckley, [1#]Sarah A. Teichmann

*These authors constibuted equally

[#]corresponding author

[1] Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, UK

[2] European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Cambridge, UK

[3] Newcastle University, Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK

[4] John van Geest Centre for Brain Repair, Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK

[5] Wellcome-MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK

[6] Kennedy Institute of Rheumatology, University of Oxford, Oxford, UK

Abstract

Bone and joint formation in the developing skeleton rely on co-ordinated differentiation of progenitors in the nascent developing limbs and joints. The cell states, epigenetic processes and key regulatory factors underlying their lineage commitment to osteogenic and other mesenchymal populations during ossification and joint formation remain poorly understood and are largely unexplored in human studies. Here, we apply paired single-nuclei transcriptional and epigenetic profiling of 336,000 droplets, in addition to spatial transcriptomics, to construct a comprehensive atlas of human bone, cartilage and joint development in the shoulder, hip, knee and cranium from 5 to 11 post-conception weeks. Spatial mapping of cell clusters to our highly multiplexed in situ sequencing (ISS) data using our newly developed tool ISS-Patcher revealed new cellular mechanisms of zonation during bone and joint formation. Combined modelling of chromatin accessibility and RNA expression allowed the identification of the transcriptional and epigenetic regulatory landscapes that drive differentiation of mesenchymal lineages including osteogenic and chondrogenic lineages, and novel chondrocyte cell states. In particular, we define regionally distinct limb and cranial osteoprogenitor populations and trajectories across the fetal skeleton and characterise differential regulatory networks that govern intramembranous and endochondral ossification. Through somatic mutation analysis, we predict two new potential cell origins for human chondrocyte development. We also introduce SNP2Cell, a tool to link cell-type specific regulatory networks to numerous polygenic traits such as osteoarthritis. We also conduct *in silico* perturbations of genes that cause monogenic craniosynostosis and implicate potential pathogenic cell states and disease mechanisms involved. This work forms a detailed and dynamic regulatory atlas of human fetal skeletal maturation and advances our fundamental understanding of cell fate determination in human skeletal development.

Main

Human bone formation begins between 6-8 post-conception weeks (PCW) in the period spanning the transition from embryonic to fetal stages. Within the calvaria of the cranial skeleton, progenitors differentiate into osteoblasts through intramembranous (IM) ossification and expand from ossification centres that eventually meet to establish the prenatal suture joints that house osteoprogenitors [1,2]. In the appendicular skeleton, the nascent synovial joint first appears as an interzone condensation in the limb bud at 5-6 PCW [3], which subsequently forms a cavitated region that articulates adjacent incipient cartilage templates. The latter acts

as a temporary scaffold to facilitate development of the body plane and is gradually replaced by bone tissue as development proceeds through endochondral (EC) ossification [4,5].

These two distinct and anatomically restricted modes of ossification govern osteogenesis and joint formation throughout the human skeleton and, to our knowledge, the cellular bases by which they form and mature remain partially described in human development at single-cell resolution. To address this, we applied single-nuclei paired RNA and ATAC sequencing, and multiple spatial methods to decipher the gene regulatory networks that mediate maturation of the distinct bone and joint-forming niches in the cranium and appendicular skeleton across space and time from 5-11 PCW. Through this, we discovered previously undescribed cellular diversity in the osteogenic and chondrogenic lineages. We develop ISS-Patcher, a tool to impute cell labels from the droplet data on our high-resolution 155-plex in situ sequencing (ISS) datasets, allowing us to gain detailed insights into spatially defined niches within the embryonic synovial joint. Applying OrganAxis, a new spatial transcriptomics annotation tool, we also define the spatial trajectory of the developing frontal bone of the fetal skull.

In addition, our comprehensive resource and new computational toolset including SNP2Cell enabled us to gain new insights into the molecular mechanisms of developmental bone disease, such as craniosynostosis[6–8] as well as to implicate region-specific contributions of bone and cartilage lineages in ageing diseases of the human skeleton, such as osteoarthritis[9,10].

Results

Cellular taxonomy across first trimester skeletal joint development

We applied paired droplet-based snRNA-seq and snATAC-seq (10x Genomics Multiome) to define the cellular taxonomy of the developing embryonic fetal skeleton between 5-11 post conception weeks (PCW) across 12 fetal donors. Focusing on three major synovial joints, we sampled whole-intact nascent and maturing shoulder, hip and knee joints across the developmental timepoints (Fig. 1a-b). For the developing skull and cranial suture joints, which had not been profiled in human across embryonic stages, we studied the anterior and posterior regions of the cranium separately, and sampled the calvaria and skull base individually to divide intramembranous and endochondral bone-forming niches. (Fig. 1a and Supplementary Table 1). We captured 336,162 high-quality transcriptome and chromatin-accessibility droplets and curated the dataset from all regions into eight shared major cellular compartments based on marker gene expression: mesenchymal, muscle, immune, endothelial, Schwann, neural, epithelial, and erythroid cells (Fig. 1c-e, Extended Data Fig. 1c). High concordance was

observed between the transcriptome and ATAC peak compartment structures (Extended Data Fig. 1f), and we observed that mesenchymal cells were the most abundant across all regions, while myogenic cells were absent in the calvaria (Fig. 1e). From these we defined 30 broad cell clusters and conducted sub-clustering of lineages which enabled the identification of 122 fine-grained clusters (Supplementary Table 2).

Previous single-cell (scRNA-seq) atlases of developing human and mouse limbs [11–13] profiled whole cells and captured low numbers of maturing osteoblast transcriptomes and *COL10A1+* hypertrophic chondrocyte transcriptomes. To facilitate the reconstruction of developmental trajectories in osteochondral lineages, we profiled nuclei-droplets from a large number of cells. This captured comparatively more diverse chondrogenic and osteogenic subcompartments, suggesting that matrix-rich stromal populations are relatively resistant to enzymatic digestion and less amenable to whole-cell profiling [11–13].

Our approach enabled the discovery of osteogenic cell state trajectories enriched in the appendicular joints and skull base (formed through endochondral ossification) and calvarium (formed through intramembranous ossification), reflecting different osteoblast origins from anatomically distinct regions (Extended Data Fig. 2). Chondrogenic clusters were relatively depleted in the calvarium (Fig. 2a), consistent with the mechanisms of intramembranous bone formation[14]. In addition, we uncovered novel cell states of the facial and pharyngeal regions, described *COL10A1*-expressing fetal hypertrophic chondrocyte development, discovered a novel *PAX7*$^{+}$ chondrocyte population, and described a potential pathway leading to its formation from common myogenic progenitors[15]. We delineate the previously unreported process of human Schwann cell development and the formation of fetal *PI16*$^{+}$ fibroblasts, the latter of which shares transcriptional similarities with postnatal pan-tissue perivascular fibroblasts.

Sequencing-based spatial transcriptomics previously deployed on the embryonic limb has highlighted challenges in capturing transcripts within the early ossifying bone and cartilaginous precursors (cartilage anlagen)[11]. In order to resolve bone lineage cell states in space and understand organisation of the nascent synovial joint that bridges the adjacent anlagen, we performed high resolution 155-plex ISS of the whole intact early embryonic upper (6.7 PCW) and lower (5.7 PCW) limbs, late embryonic (7.3 PCW) knee, and shoulder regions (Fig. 1f and Extended Data Fig. 2a-c). Additionally, we conduct sequencing-based spatial transcriptomics (10x Genomics Visium CytAssist) of the developing coronal suture (9 PCW) and frontal bone (Extended Data Fig. 2d) allowing us to capture the trajectory of osteo-lineage development across space. We leveraged these spatial data to systemically curate cell lineages within the

133 mesenchymal compartment in a well-defined spatial context.
134
135 **Spatiotemporal zonation of the incipient synovial joints**
136 Synovial joint-site determination occurs between 5-6 PCW in the limbs, orchestrated by the
137 emergence of a mesenchymal condensation comprising *GDF5*-expressing populations
138 termed the interzone (IZ). Part of them undergo differentiation into joint-articulating
139 chondrocytes at the ends of the endochondral bone primordia, and chondrocytes that adopt a
140 hypertrophic phenotype in the cartilage anlagen[16]. To identify cell-states that constitute these
141 early joint-specific progenitors, we applied differential abundance testing focusing on the non-
142 myogenic mesenchyme across developmental time (5-11 PCW) and skeletal regions
143 (shoulder, hip, knee) (Fig. 2a). We identified four progenitor populations enriched in the early
144 appendicular joint: Interzone Chondrocytes (InterzoneChon), *HIC1*$^{+}$ mesenchyme (*HIC1*$^{+}$
145 Mes), *PI16*-expressing fibroblast-progenitors (FibroPro) and dermal fibroblasts (DermFibro).
146 These clusters support the notion that numerous early progenitors exist between 5-7 PCW
147 and give rise to the chondro-fibro lineages in the joint, instead of a single master *GDF5*$^{+}$
148 progenitor population[17].
149
150 We first focused on deciphering the composition of progenitors in the early synovial joint,
151 where fibrous ligaments[18], tendon[19] and cartilage components are thought to derive from a
152 progenitor *GDF5*$^{+}$ population in mice[17,20]. To this end, we performed sub-clustering of the
153 broad InterzoneChon (*GDF5*$^{+}$) population and leveraged RNA-velocity to infer their pseudo-
154 trajectory (Fig. 2b-c). We then applied SCENIC+ to identify predicted gene programs and TF
155 accessibility changes across the resulting subclusters (Fig. 2d). Our analysis revealed multiple
156 subclusters with specific tissue expression signatures, including an early *PRRX1*$^{+}$
157 mesenchyme population (EarlyIZ) enriched for mesenchyme-associated signatures (Fig. 2b).
158 Early IZ was predicted to express TFs associated with early limb development (e.g. *TBX18*,
159 *SHOX*, *LHX9*)[11] and demonstrated low *RUNX2* expression and accessibility in downstream
160 target regions but moderate *SOX5, SOX6* and *SOX9* expression and target region
161 accessibility, suggesting a poised trajectory favouring chondrogenesis over osteogenesis (Fig.
162 2d).
163
164 We defined four main trajectories that emerge from this latter population, characterised by
165 transcriptional signatures of fibroblasts (Fibro IZ), articular (Articular IZ) and hypertrophic
166 chondrocytes (Hypertrophic IZ-1, Hypertrophic IZ-2), and flanking dermal mesenchyme
167 (Flanking IZ) (Fig. 2b-c). Three clusters; Articular, Fibro and *GDF5*^{high} IZ, expressed *GDF5*.
168 Articular IZ was more prevalent in the knee joint, whereas Fibro IZ is enriched in the ball-and-
169 socket joints of the shoulder and hip (Fig. 2c). Articular IZ particularly expressed chondrogenic

170 markers (*COL2A1, ADGRG6*, *PIEZO2*, *LGR5*, *NOG*) comparable to mouse IZ progenitors for
171 articular chondrocytes [17,20] including *ENPP1*, which negatively regulates chondrocyte
172 hypertrophy and bone formation, in keeping with a role in articular cartilage formation. While
173 *GDF5*-high IZ was predicted to highly express *CREB5* and *DBX2* and its target genes, the
174 latter of which is associated with digital IZ formation regulated by *HOX* genes[17,21],
175 chondrogenesis TFs were not highly expressed, and predicted TF activity and accessibility for
176 *RUNX2* was lowest among the IZ populations. We therefore hypothesise that *GDF5high* IZ is
177 maintained as a progenitor pool, potentially to sustain continuous influx into the forming joint[18].
178 In the *GDF5 negative* hypertrophic clusters, progressive TF expression and accessibility of
179 *RUNX2*[22–24], and concurrent downregulation of articular chondrocyte TFs (*CREB5*, EGR1)
180 signifies a hypertrophic phenotype (Fig. 2d). Hypertrophic2 IZ transitions through
181 Hypertrophic1 in the inferred trajectory and the former expressed genes that are upregulated
182 during formation of anlagen or pre-hypertrophic chondrocytes in mice (*BMP5, HAPLN1, ACAN*
183 and *ZFHX3*)[19,25], suggesting they form the incipient cartilage template. Overall, our *in vivo* data
184 are consistent with *in vitro* observations of the propensity for human-iPSC and mouse-ESC
185 derived *GDF5*$^{+}$ stromal cells to form an articular, rather than a hypertrophic phenotype [26,27].
186
187 To understand spatial zonation of the nascent synovial joint we next visualised the pre-
188 cavitating knee joint ISS to identify regions of incipient hypertrophic anlagen (*RUNX2)*, early
189 progenitors (*TBX15, SHOX)* and nascent articular cartilage and interzone (*SOX9*) (Fig. 2e).
190 We then leveraged the snRNA-seq dataset and imputed genes missing in the 155-plex ISS
191 dataset through our newly developed ISS-Patcher function (see Methods) to infer cell labels
192 to the 155-plex clustered manifold (Fig. 2f,g). In the embryonic lower limb, Early IZ was
193 diffusely distributed across regions of the interzone and anlagen, and was surrounded by
194 Flanking IZ (Fig. 2c,d). Articular IZ (*GDF5*$^{+}$), were predominantly enriched in sites of incipient
195 knee articular cartilage formation which also showed SOX9 staining (Fig. 2e,g). In contrast,
196 Fibro IZ (*GDF5*$^{+}$) was enriched in the shoulder interzone region adjacent to the articular
197 surface of the humerus (Fig. 2g), and expressed *SCX* (Fig. 2b) which plays a role in ligament
198 formation in the IZ in mice[28]. The differential enrichment and signatures of the Articular and
199 Fibro IZ populations may be explained by the need for fibro-cartilage labrum formation in the
200 ball-and-socket joints of the shoulder and hip which emerge and mature later in development.
201 We coalesce the trajectories and spatial enrichment of the newly defined interzone progenitors
202 and describe a model for zonation of the embryonic joint. Within the model, the future joint
203 region (Articular phenotype) forms in the centre of the interzone defined by early pro-
204 chondrogenic and anti-osteogenic TF enrichment, the incipient fetal anlagen (Hypertrophic
205 phenotype) forms at the edges away from the joint, defined by *RUNX2* enrichment (Fig. 2h).
206

Emergence of fibroblast lineages in synovial joints

Fibroblast lineage cell states have previously been described in the developing mouse limb to arise from a master *HIC1*$^{+}$ precursor population[29], which contributes minimally to osteochondral components. In addition, a "universal" *PI16*$^{+}$ population is thought to persist in postnatal stages and govern adult fibroblast formation across tissues in mice and humans [29,30]. Here, we sought to uncover the taxonomy of the fibroblast lineage in first trimester human joints. We first identified fibroblast progenitors (FibroPro) and *HIC1*$^{+}$ mesenchyme (*HIC1*$^{+}$Mes) enriched in the appendicular joints during the embryonic period (<8 PCW), surrounding the nascent joint and with diffuse distribution in the limbs, respectively (Fig. 2a, Extended Data Fig.3a-c,f). *TWIST1*, a known activator of postnatal fibrosis and *TWIST2* a regulator of postnatal dermal fibroblast proliferation [31,32], were predicted to show high TF activity in DermFibro (Extended Data Fig. 3e), which was present across sites and time, and localised to the skin around the developing embryonic limbs (Extended Data Fig.3 f-g)**.** Interestingly, *EN1,* a TF required for the fibrotic response and associated scarring during postnatal wound healing[33], was highly expressed in DermFibro (Extended Data Fig. 3e) but showed low target gene expression, suggesting target repression, consistent with observations of scarless wound healing *in utero*. In the cranium, it formed the majority of early-stage enriched fibro-lineage progenitors (Fig. 2a).

To uncover developmental dynamics in the appendicular joints, we reconstructed pseudotime trajectory across fibroblasts and predicted *HIC1*$^{+}$ Mes as a progenitor to FibroPro during the embryonic phase at <8 PCW (Extended Data Fig. 3a). *HIC1*$^{+}$ *Mes* enriched for activity in numerous proliferation associated TFs including *WT1, SOX5,* and *FOXC1,* which is associated with an invasive and activated synovial fibroblast phenotype (Extended Data Fig. 3e) [34–37]. At ~8 PCW, FibroPro forms a "fibroblast hub" and expresses *PI16* and *DPT*, markers of pan-tissue adventitia-associated fibroblasts in postnatal health (Extended Data Fig. 3d). Here, it gives rise to tenocytes, synovial, dermal and myo- fibroblasts (Fig. 2a,b). *BNC2,* a myofibroblast-associated TF [38], had high activity in FibroPro, consistent with its postnatal pan-tissue presence in the vascular adventitia [30]. Additionally, *YBX1*, a TF shown to drive proliferation of mouse embryonic fibroblasts was also enriched. *HIC1*+ Mes additionally differentiated to the tenogenic lineage during the embryonic, and synovial fibroblasts in the fetal phase, highlighting potential parallel routes to synovial fibroblast formation in the prenatal skeleton (Extended Data Fig. 3a-d).

242 **Formation of the cranium and suture joints**

243 In the cranium, incipient suture mesenchyme matures beyond 7 PCW at meeting points of
244 bone fronts emanating from primary ossification centres in the flat bones, forming suture
245 joints[39]. We conducted differential abundance testing of the mesenchyme compartment and
246 revealed the embryonic cranium (<7 PCW) was dominated by dermal and myo- fibroblasts,
247 and two broad *PAX9*$^+$*RUNX2*$^+$ and *ALX1*$^+$*RUNX2*$^+$ mesenchymal clusters that had not been
248 described previously in human craniogenesis (Fig. 2a). The latter expressed markers such as
249 *TWIST1, CTSK, ZIC1,* consistent with mouse cranial progenitors[40], and *RUNX2*, suggesting
250 they form part of the osteogenic lineage (Extended Data Fig. 4a). *ALX1* is required for cranium
251 formation in the mouse[41], and neural crest cell (NCC) migration and differentiation in human
252 derived stem cells, with variants of the gene associated with frontonasal dysplasia [28,42]. Single-
253 cell studies of the human brain have also identified *ALX1* expression within mesenchyme-like
254 progenitors[43]. We therefore used metadata obtained during tissue sampling to examine
255 differences in anterior-posterior skull gene expression and found numerous neural crest cell
256 (NCC) TFs and targets linked in our GRN including *ALX1, PAX3, BMP5, COL4A2, COL4A1*
257 and *TSHZ2* which are differentially enriched in the anterior portion of the cranium (Extended
258 Data Fig. 4b). This suggests that *ALX1*$^+$ Mes may have a neural crest origin. Due to the
259 transient embryonic nature of NCCs[44] we did not capture early embryonic *bona fide SOX10*$^+$
260 NCCs. *RUNX2*-expressing *PAX9*$^+$ Mes was present in the cranium (skull base) and
261 appendicular skeleton (Fig. 2a), and therefore likely represents a lateral plate and axial
262 mesoderm-derived progenitor for osteoblasts in endochondral ossification.

263

264 Next, we sought to delineate the contribution of the broad ALX1$^+$ and PAX9$^+$ Mes clusters to
265 the bone-forming suture joint. We first sub clustered these populations in combination with
266 osteoblasts (*SP7, ALP)*, osteocytes (*SOST, DMP1)* and perichondrium (*RUNX2, THBS2,*
267 *POSTN)* to form an osteo-lineage subcompartment. This revealed anatomical and age -
268 segregated clusters (Fig. 2a, Extended Data Fig. 4c-d). Among these were three early
269 progenitors, including facial (FacialMes) and pharyngeal mesenchyme (PArchMes) which
270 were located in the skull base and expressed markers of axial mesoderm (*PAX3 and LHX8*),
271 and a cranial mesenchyme (CranialMes) population which was abundant in the calvarium (Fig.
272 2a, Extended Data Fig. 4d-e).

273

274 Cranial sutures form in fetal stages following expansion of the primary ossification centres at
275 the end of the embryonic period[1]. Consistent with this, we define two human suture
276 mesenchyme clusters (SutureMes1/2) enriched in the cranium from 8 PCW (Fig. 2a). Classical
277 markers of fetal cranial sutures (*TWIST1, ZIC1, ZIC4)*, previously reported in NCC-derived
278 mesenchyme in mice, were enriched in both populations in addition to high expression levels

of *THBS2,* akin to the endochondral perichondrium cluster (Extended Data Fig. 4d)[45]. Notably, the SutureMes populations expressed *CTSK*, a marker of periosteal mesenchymal stem cells in the postnatal mouse cranial osteogenic niche shown to mediate intramembranous ossification[46]. Using Cell2location, we revealed the spatial distribution of SutureMes1 and SutureMes2 within the developing coronal suture joint, spanning outward toward the periosteum (Fig. 2i). Osteoprogenitors (*HHIP*$^{+}$PreOB) emerged at the opposing frontal and parietal bone boundaries of the suture populations (Fig. 2i). Additionally, the sutures were flanked by peri-suture (*CTSK*$^{-}$) in contrast to the *CTSK*$^{+}$ postnatal mouse periosteum previously reported[46] (Extended Data Fig. 4g).

Analogous to osteogenic repressor genes enriched in spatially-defined clusters of the articular IZ, SutureMes populations also showed predicted TF and target accessibility enrichment of negative regulators of osteogenesis in the mouse cranial sutures (*TWIST1, LMX1B*, *NFATC2)*[47] (Fig. 2i). Numerous TFs associated with osteogenesis (*EGR1, EGR2, SP7)*[48,49] and osteoblast formation (*FOXO1, ETS1, ETS2*)[50] were predicted suggesting a transcriptional network primed for impending osteogenesis. Akin to the molecular gradients of the joint-interface in the embryonic knee (Fig. 2h), we observed *LMX1B* and *TWIST1* expression within the suture region, dissipating toward the flanking bone edges alongside a rise in regulators of osteogenesis such as *RUNX2* (Fig. 2i), suggesting comparable mechanisms in sustaining the non-osteogenic interface in both the limb and skull bones.

To dissect this process and uncover the enhancer-driven gene regulatory network (eGRN) of the loci surrounding key osteogenic transcription factors *RUNX2* and calvaria-enriched *HHIP,* we visualised coverage of the network predicted by SCENIC+. Both *RUNX2* and *HHIP* were predicted to be inhibited by a shared set of anti-osteogenic TFs including *LMX1B*, *TWIST1* and *ALX4* via several intermediate repressors targeting various enhancers around their loci (Fig. 2k), illustrating the relationships maintaining the balance of osteogenesis initiation. *HHIP* was most accessible in *HHIP*$^{+}$PreOB and was indirectly repressed by *LMX1B* via *TWIST1* which positively regulated its immediate repressors *MAF* and *HMGA2* via two differentially accessible enhancers of *HHIP* enriched in SutureMes1 and *HHIP*+PreOB. *RUNX2* was most accessible in *HHIP*$^{+}$PreOB (IM) as well as Perichondrium (EC) and was predicted to be indirectly targeted by the same repressors via *TCF12* and *PRRX2*. Overall, the network illustrates the coherent regulation of bone-adjacent non-ossifying niches by key osteogenic regulators via multiple redundant paths.

315 **Distinct trajectories of osteogenesis across the skeleton**

316 Osteoblastogenesis, signifying the onset of ossification, commences from ~7-8 PCW onward
317 (Fig. 2a). To reconstruct this process, we utilised force-directed embedding on non-cycling
318 cells and uncovered two major trajectories from distinct progenitors converging toward
319 osteoblast and osteocyte formation (Fig. 3a)**.** Endochondral ossification (EC) consisted of two
320 routes, stemming from limb mesenchyme (LimbMes), FacialMes and PArchMes. LimbMes, a
321 cluster which encompasses upper and lower limb joints, demonstrated transcriptional
322 similarity (*ISL1, TBX5, WT1)* to lateral plate mesoderm previously reported in the fetal human
323 limb bud **(**Fig. 3a, Extended Data Fig. 4a-e**)** [11,30]. CranialMes and FacialMes were differentially
324 abundant in the anterior portion of the calvarium and skull base (Fig. 2a), respectively.
325 FacialMes expressed NCC-derived mesenchyme regulators *PAX3* and *PAX7 (*Extended Data
326 Fig. 4e)[51]. Likewise, *ALX1* was enriched in CranialMes, leading us to hypothesise that the
327 clusters represent, to our knowledge, previously undescribed human NCC-derived osteogenic
328 populations. *LEPR*+Mes was enriched in the calvarium, *LEPR* has been reported as a marker
329 of second trimester and post-natal osteogenic stromal population [12,52]. PArchMes emerged
330 between 6-8 PCW and expressed marker genes of the first pharyngeal arch (*PAX9, LHX,*
331 *DLK1)* [53], and is likely to comprise mixed origins from axial mesoderm and NCC. We
332 discovered a two-stage process in the intramembranous pathway (IM), defined by the
333 transition between embryonic and fetal stages. Suturogenesis occurs in stage 1, as
334 CranialMes differentiates into SutureMes1/2 (Fig. 3a). In stage 2, an osteogenic trajectory
335 emerges from SutureMes1 toward SutureMes2, followed by *HHIP*$^{+}$PreOB. *HHIP* was
336 previously reported in the mouse as a marker of osteogenic coronal suture mesenchyme [54].
337 We demonstrate here that they are a distinct population progeny to *TWIST1-enriched*
338 SutureMes and are distributed at the boundaries of bone formation in human fetal
339 development (Fig. 2i).
340

341 Within the two newly defined trajectories for EC and IM osteogenesis, we determined dynamic
342 expression and modules associated with pseudotime as they converge (Fig. 3a-b). Common
343 TFs (*RUNX2, DLX5, SP7, SATB2)* mediating osteogenesis were upregulated toward terminal
344 states across both pathways. The newly described SutureMes1/2 was enriched for osteogenic
345 genes (*SPARC, COL1A1, GAS1)* early in the trajectory and signatures associated with suture
346 formation and regulation in the fetal and postnatal mouse suture (*PRDM6, MSX1)* were most
347 regulated upon transition to SutureMes2 [55]. Inhibitory control of osteogenesis differed across
348 the pathways, *LMX1B* a common repressor, was enriched late in progenitors across both
349 pathways whereas *TWIST1* was only enriched late in IM progenitors, signifying additional
350 inhibitory regulation in the IM pathway. *NFATC2*, a chondro- and osteo- genesis repressor

351 [56,57], was observed early in the IM pathway, but upregulated late in the EC pathway, potentially
352 reflecting context and timing dependent roles in maintenance of the suture niche (IM), and
353 cartilage primordia replacement (EC), respectively. In IM progenitors, accessibility across
354 activators (*RUNX2*) and repressors (*TWIST1, LMX1B*) was simultaneously high at early parts
355 of the trajectory, suggesting they are poised for osteogenesis but remain repressed in the
356 early stages (Fig. 3c). In contrast, at the transcriptome level, there was a reciprocal
357 relationship between repressor and *RUNX2* target expression across pseudotime, suggesting
358 additional, redundant layers of regulation limiting RUNX2 target expression. In the EC pseudo-
359 trajectory, repressor (*TWIST1, NFATC2*) accessibility peaked toward Perichondrium and then
360 decreased, whereas *RUNX2* target accessibility increased as Perichondrium formed,
361 suggesting the latter was critical in driving progression. This contrasts with the gradual
362 reduction in IM repressor accessibility marking the transition into *HHIP*+PreOB alongside a
363 persistently accessible activator (*RUNX2)* (Fig. 3c). In concert, the multi-omic dynamic
364 changes in the IM pathway, in contrast to EC osteogenesis, suggest additional layers of non-
365 transcriptional regulation.
366
367 Following suture formation, progressive waves of oriented differentiation emanate from the
368 cranial sutures toward the developing bone front [45]. In order to identify the newly defined cell
369 states in this progression (stage 2), we utilise OrganAxis (see Methods) to define a maturation
370 axis spanning the coronal suture into regions of the maturing frontal bone. Subsequently, we
371 create zonal bins within this continuous axis based on histological features, and evaluate cell-
372 state mapping along the anterior-posterior (AP) axis (Fig. 3d). Enrichment of *TWIST1*$^{+}$
373 SutureMes1/2 were substantial in the suture zones (1-3) which constitute the centre of the
374 coronal suture (Fig. 3e). Their enrichment peaked within the osteogenic front where
375 histological features of osteoprogenitors began to emerge along with *HHIP*$^{+}$ PreOB.
376 Establishment of the osteogenic zones coincided with downregulation of anti-osteogenic
377 (*LMX1B*, *TWIST1)*, and upregulation of pro-osteogenic (*RUNX2, DLX5, SP7)* TFs (Fig. 2i-j,
378 3d-e), signifying a spatially restricted molecular switch that defines territories of the suture.
379 Osteoblast-specific genes including *IRX5, SOST, SPP1, MMP9 and DMP1* gradually peaked
380 toward the most distant osteogenic zone[58]. Clustering of the spatial voxels aligned osteogenic
381 TFs with axis values away from the coronal suture (Extended Data Fig. 5c).
382
383 In EC ossification, bone formation initiates in the interfaces of the synovial joint, with
384 perichondrium formation surrounding the hypertrophic chondrocyte template, giving rise to
385 osteoblast in fetal stages and in postnatal injury repair [11,52,59]. At the border of the sphenoid in
386 the anterior skull base, we identified the anterior boundary between EC and IM osteogenesis

387 at the frontal-sphenoid interface, with co-localisation of SutureMes1 and SutureMes2, which
388 precede *HHIP*+PreOB and Osteoblast (Fig. 3e), forming a previously undescribed boundary-
389 defining niche. In the limb, we applied ISS-patcher to the ISS data, and inferred spatial
390 trajectories for differentiation of the LimbMes towards perichondrium and osteoblasts in a
391 centripetal pattern in the incipient bone at ~7 PCW (Fig. 3f). In the older 9 PCW skull base, a
392 comparable pattern was detected in the sphenoid where hypertrophic chondrocytes were
393 surrounded by perichondrium (Fig. 3g).
394
395 To further resolve the endothelial-bone relationship within the IM niche, which remains poorly
396 understood compared to the canonical vascularisation models described in the EC niches[60],
397 we leveraged both the spatial locations and inferred osteo-endothelial interactions to elucidate
398 a plausible cellular basis for endothelial sprouting. As expected within the IM niche, *MSR1*$^{+}$
399 Tip cells, Mural cells and Capillary Endothelial cells (ECs) progressively co-enriched along the
400 osteogenic zones with Osteoblast and Osteocytes (Fig. 3e, Extended Data Fig. 5a-b),
401 suggesting a spatial-temporal association between vascularisation and osteogenesis.
402 Concurrent to this trend, *ATF4*, a regulatory gene that promotes bone angiogenesis in
403 development [61], was also enriched along the axis.
404
405 Pro-angiogenic factors *VEGFA* and *VEGFB,* the concentration of which modulates vessel
406 sprouting, were highly enriched at the osteogenic front, which co-localised with SutureMes1/2
407 and *HHIP*$^{+}$PreOB, which highly expressed *VEGFA* and *VEGFB*, suggesting they are the likely
408 source. Endothelial populations highly express the *VEGF* receptor genes (*FLT1, KDR, NPR1*)
409 (Fig. 3h), which therefore suggests the former plays a pivotal role in promoting vascular
410 invasion into the bone in IM akin to the well-described function of hypertrophic chondrocytes
411 in EC niches[62,63].
412
413 To determine distinct cell-cell interactions across the EC and IM niches, we utilised NicheNet
414 to compare differences in signalling during sprouting angiogenesis, focusing on recipient tip
415 cells that lead the vascular sprouts. SutureMes1/2 distinctly enriched for *EPHB2* and *EPHB3*,
416 which promotes sprouting behaviour, motility, and vessel formation (Fig. 3i) [64]. *RACK1*, an
417 intracellular scaffold protein that promotes *VEGF-FLT1*-dependent cell migration[65], was
418 exclusively expressed in tip cells in the calvaria, signifying a highly motile state. *CXCR4*, which
419 promotes sprout anastomosis, was upregulated in tip cells[66]. In EC ossification, tip cells
420 distinctly express *CDON*, the receptor for *IHH* which promotes endothelial proliferation,
421 migration, and angiogenesis. Together, these data support the essential role of SutureMes1/2
422 and *HHIP*$^{+}$PreOB in coordinating blood vessel invasion and migration in the IM osteogenic
423 niche.

Next, we sought to unravel inferred signalling interactions from the invading endothelial cells toward osteogenic populations. We first defined a spatial gradient of angiogenesis along the defined bone-maturation axis by scoring for enrichment in sprouting angiogenesis pathways (Fig. 3e), which corroborated the coupling of angiogenesis and osteogenesis, orchestrated in the direction of bone growth. We then used CellphoneDB to predict significant signalling interactions from ligands in the endothelial populations toward the recipient osteogenic cell states (Fig. 3j)[67]. Endothelial cells, including capillary ECs, mural cells and tip cells were predicted to signal to osteoblastic cells via *DLL1/DLL4–NOTCH2* and *JAG1/JAG2-NOTCH2* interacting pairs, which have been reported to promote the differentiation of postnatal perivascular osteoprogenitors[68]. Of note, no *NOG* expression was observed in endothelial cells, unlike the previously described postnatal bone-associated type H vessels[68], which suggests a novel and unique pattern of crosstalk between endothelial and osteoblastic cells via *NOTCH* signalling in the embryo. Additionally, mural and capillary ECs promoted osteoblast differentiation through previously described interactions including *FGF2-FGFR2*[69] and *THBS1-CD36*[70]. Mural cells are also likely to be promoting bone mineralization via activation of canonical *WNT* signalling through the observed RSPO3-LGR5 pairs[71]. In addition to pro-osteogenic roles, the endothelial cells all support osteoclast recruitment and differentiation via *CCL14-CCR1*[72] and *CXCL12-DDP4* interactions[73]. Through these spatially defined interactions, we discern a model of the interplay between angiogenesis and osteogenesis in the intramembranous niche of the frontal bone whereby tip cells are first recruited by *VEGF* and *EPHB2* from suture mesenchyme and pre-osteoblastic cells, toward the pre-osteogenic condensation that form the bone-front. Vascularising endothelial cells then promote osteoblastic differentiation, osteocyte mineralization and osteoclast recruitment in the maturing bone (Extended Data Fig. 5e).

Other lineages, including neurons have also been reported to play a role in modulating osteogenesis[74,75]. Neuronal development and axon guidance modules were enriched in early stages of both osteogenic pathways (Supplementary Table 4), suggesting involvement of chemical cues from axons in niche formation. Specific IM modules were associated with epithelial mesenchymal transition (EMT), hypoxia response, and TGF signalling, whereas EC-enriched modules were related to neuroactive ligand-receptor interactions, axon guidance and development (GSEA, MSigDB Hallmark 2020 gene sets). To understand the potential role of axon guidance in skull formation, we compute gene-scores for axon guidance modules along the spatial trajectories of osteoblastogenesis in the cranium and revealed enrichment in the suture mesenchyme (Extended Data Fig. 5f), consistent with previous description of axonogenesis regulating suture formation in the mouse[76]. Future studies capturing

461 transcriptional profiles of neuronal cell bodies of the innervating axons may shed light on the
462 cell-cell interactions at play.
463

464 **Inferring *PAX7*$^+$ chondrocyte origins through somatic mutations**

465 Chondrogenesis is an essential process for the formation of various types of cartilage,
466 including hyaline, fibrous, and elastic cartilage. To reveal the transcriptional heterogeneity of
467 chondrocytes, the chondrocyte subcompartment was further sub-clustered across anatomical
468 regions and timepoints and annotated based on canonical marker genes (Fig. 4a, b, Extended
469 Data Fig. 6a and Supplementary Table 2). To dissect cellular heterogeneity within a spatial
470 context, we applied ISS-Patcher and Cell2location to transfer cell state labels from our single-
471 cell atlas to ISS and 10x Visium slides from the shoulder at 7.3 PCW and calvaria at 9 PCW
472 (Extended Data Figs. 6b, c).

473 Our chondrocyte clusters exhibited strong regional-specificity due to divergent development
474 origins (Fig. 4c)[77]. In terms of cell clusters, we identified known populations, such as
475 hypertrophic chondrocyte (HyperChon, *COL10A1*, *IHH*), cycling chondrocyte (CyclingChon,
476 *MKI67*, *TOP2A*), resting chondrocyte (RestingChon, *UCMA*), interzone chondrocyte
477 (InterzoneChon, *GDF5, PITX1*), as well as new populations, including two chondrocyte
478 progenitors (ChondroPro1, ChondroPro2), anatomical region-specific chondrocytes, *DLK1*$^+$
479 chondrocytes (*DLK1*$^+$ Chon, *DLK1*, *CD63*), and *PAX7*$^+$ chondrocytes (*PAX7*$^+$ Chon, *PAX7*,
480 *RBFOX1*). Two region-specific ChondroPros, which were enriched in appendicular joints and
481 skull respectively, transitionally expressed fibroblast differentiation markers (*POSTN*,
482 *COL1A1*, *PRRX1*, and *TWIST1*), which was consistent with findings in early chondrocyte
483 progenitors in mice[19]. In the skull base, facial chondrocytes highly express *PAX3*, indicating
484 an origin from neural crest[78]. Mandibular chondrocytes highly expressed *SEMA3D* in posterior
485 regions[79,80]. In appendicular regions, we observed two subtypes of articular chondrocytes, with
486 one population being enriched for *TRPV4* and *VEGFA*, while the other one was more mature,
487 with relatively low *SOX9* expression and high *EPYC* expression. In particular, *DLK1*$^+$ Chon
488 was highly enriched in ribosomal genes and distinctly expressed *CD63* and *DLK1*. *CD63* has
489 been identified in the pre-hypertrophic layer in the limb, while *DLK1* is a novel marker for
490 embryonic chondroprogenitor cells undergoing lineage progression from proliferation to pre-
491 hypertrophic stages. Spatially, CyclingChon, *DLK1*$^+$ Chon and HyperChon were also arranged
492 sequentially within the bone, spanning from the epiphysis toward to diaphysis, the primary
493 ossification centre (Extended Data Fig. 6b), which corroborates the transitional role of *DLK1*$^+$
494 Chon, from proliferative to pre-hypertrophic phenotypes during chondrogenesis.

495 With regard to gene regulatory networks, we identified common and region-specific co-
496 expressed gene modules using Hotspot[81] and performed gene ontology enrichment on each
497 module (Extended Data Figs. 6 d,e). Shared gene modules (M3, M5, etc) across regions
498 mainly related to chondrocyte cellular structure and functional development, including
499 biomineralization, collagen-containing extracellular matrix, and cartilage development.
500 Interestingly, synapse related pathways were enriched both in appendicular and cranial
501 samples, suggesting axons may contribute to development of cartilage[82] which in pathology
502 may be responsible for joint pain. In addition, chondrocytes from the skull base enriched for
503 axon guidance gene modules, while chondrocytes from appendicular joints had specific gene
504 modules, including glucose metabolism and cellular response to hypoxia (Extended Data Figs.
505 6f, g). Taken together, our data provides a comprehensive catalogue of chondrocyte molecular
506 profiles in multiple developing human joints.
507
508 Moreover, our work uncovered a previously undescribed developmental chondrocyte cluster
509 (*PAX7*$^{+}$ Chon), which first appeared at 7 PCW and peaked between 9-10 PCW across multiple
510 donors (Fig. 4c). *PAX7*$^{+}$ Chon co-expressed markers and gene modules of chondrocytes and
511 muscle cells on the transcriptomic and epigenetic level (Fig. 4d), and showed overlap of genes
512 that enrich in cartilage and muscle development pathways (Extended Data Fig. 6h, i). To
513 interrogate the possibility of doublet-capture artefact, we first applied stringent doublet
514 processing by performing extra statistical tests over the doublet score (see Methods, Extended
515 Data Fig. 7a, b). Then we verified the presence of *PAX7*$^{+}$ Chon through *in situ* hybridization
516 experiments (Fig. 4e), which was consistent with the multiplexed ISS data (Extended Data
517 Fig. 6b). Additionally, *PAX7*$^{+}$ Chon displayed transcriptional characteristics not explained by
518 doublet formation (Extended Data Fig. 7b,c). To reveal GRNs that govern the transcriptional
519 identity of *PAX7*$^{+}$ Chon, we applied SCENIC+ and constructed a GRN based on differentially
520 expressed transcription factors (TFs) with their predicted target genes. Aside from identifying
521 key myogenic and chondrogenic key regulators (*PAX7*, *MYF5*, *SOX5*, *SOX9*), we also
522 discovered core posterior axis HOX gene modules, suggesting that patterning of major
523 synovial joint chondrocytes may be conferred through their nested HOX expression [83].
524
525 Combined profiling of RNA and ATAC offers deep coverage for mutation detection at the single
526 cell level. Under the assumption that shared somatic mutations should be observed in cells
527 derived from common progenitors, we used SComatic[84] (see Methods) to detect somatic
528 mutations in the form of single-nucleotide variants (SNVs) at the single cell resolution to
529 ascertain whether chondrocytes may share origins with the muscle lineage, as suggested by
530 previous mouse studies[85,86,87]. We identified shared mutations between *PAX7*+ Chon and

531 myogenic cell states across multiple donors (Fig. 4g, h). The latter included three progenitor
532 cell states. Skeletal muscle mesenchymal cells (SkM.Mesen)[15] not only express classical
533 muscle markers at low levels, but also exhibit gene expression related to mesenchyme. Myo-
534 skeletal stem-like cells (MSSCs) expressed *CADM1*, *MECOM*, and *FOXP1*, previously
535 reported as makers of mouse skeletal/stem progenitors[88], along with genes related to muscle,
536 cartilage, and mesenchyme. In addition, MSSCs expressed low levels of lateral plate
537 mesoderm markers. Conversely, *PAX7*$^+$ MyoProgenitor cells highly expressed muscle-related
538 genes, suggesting they are a distinct branch separate from *PAX7*$^+$ Chon (Extended Data Fig.
539 7c). We further interrogated the relationship between these progenitors and *PAX7*$^+$ Chon,
540 applying scFate, to reveal a trajectory from MSSCs to *PAX7*$^+$ Chon, with upregulation of
541 chondrocyte genes and downregulation of stem cell genes, such as *SOX4*, *CDH2* and *CDH11*
542 on both the RNA and ATAC level (Extended Data Figs. 7d-f). Comparison of mouse and
543 human cell clusters revealed similarity between *PAX7*$^+$ Chon and a *MYF5*$^+$ progenitor-derived
544 mouse cluster from the extraocular muscle (Extended Data Fig. 7g), which was reported to
545 derive from progenitors that give rise to both myogenic and connective tissues in mouse[15,87].
546 Taken together these results indicate the possibility of a new source of chondrocytes from
547 muscle progenitors, which suggested the potential complexity of embryonic cell types prior to
548 reaching terminal differentiation.

549 Lineage tracing experiments in zebrafish and mice[89] have shown that Schwann cells possess
550 the evolutionarily conserved potential to differentiate into chondrocytes during embryonic
551 development in both species. Here, we identified SOX9$^+$ endoneurial Schwann cells (SOX9$^+$
552 enSC) within the Schwann cell compartment, characterised by the expression of typical
553 chondrocyte (*SOX9*, *COL9A1*, *ACAN*, *COL2A1*), and classical Schwann cell markers (*MPZ*,
554 *SOX10*) (Extended Data Figs. 8a-c). Velocity analysis suggested that multipotent hub
555 Schwann cells (Hub SCP) serve as the foundational cells in the differentiation hierarchy.
556 SOX9$^+$ enSC represented one of the endpoints on the mesenchymal trajectory, expressed a
557 mesenchymal gene signature (*PRRX1*, *PRRX2*, *PDGFRA*, *TWIST2*) and upregulated HOX
558 TFs (*HOXA9*, *HOXA10*, *HOXA11*, *HOXD10*) (Extended Data Figs. 8d, e). We applied
559 RNAscope (*SOX9*, *SOX10, MPZ*) to localise SOX9$^+$ EnSC in the acetabulum at 7.3PCW
560 (Extended Data Figs. 8g, h). Shared mutations between chondro-lineages and Schwann
561 lineages were enriched in SOX9$^+$ enSC, indicating that SOX9$^+$ enSC exhibit a close
562 differentiation relationship with chondrocytes (Extended Data Figs. 8i). Most of these
563 mutations were concentrated in the skull base, consistent with the process described in the
564 craniofacial and trunk skeleton of mice[89]. Notably, a limited number of mutations were also
565 detected in the hip, consistent with the findings from the RNAscope experiments. In summary,
566 the analyses suggest that Schwann lineage cells may carry the potential to serve as an

567 alternative source of chondrocytes in humans. However, further lineage tracing experiments
568 are essential to validate this. We therefore hypothesise that chondrocytes derive from multiple
569 potential origins, such as muscle and Schwann lineages during embryonic development (Fig.
570 4i).

571

572 **Developmental links to complex traits of the adult skeleton**

573 Numerous conditions of the postnatal and ageing skeleton have been linked to disrupted joint
574 and bone changes during the embryonic stages of life. Of note, enhancer-associated variants
575 linked to adult osteoarthritis (OA) appear to act on anatomical-region-specific regulatory
576 networks to influence synovial joint morphology during development[90]. However, the cell
577 states that mediate this process remain evasive.

578

579 To explore potential causal links in foetal skeletal cell states for adult complex traits, we first
580 conducted fGWAS (Fig. 1c, see Methods), which integrates full GWAS summary statistics with
581 cell-state specific epigenetic and transcriptomic signatures across clusters, computing cell-
582 cluster enrichments for phenotypes. Interestingly, we found distinct lineage enrichments for
583 knee and hip OA in chondrogenic and osteogenic cell populations, respectively (Fig. 5a). Knee
584 OA and its surrogate phenotype, total knee replacement (TKR) enriched in most chondrogenic
585 states, except for *PAX7*$^{+}$ chondrocytes and InterzoneChon (*GDF5*$^{+}$). In contrast, hip OA
586 enrichment was only significantly observed in two chondrocyte populations (ChondroPro and
587 Hypertrophic Chondro), but was additionally enriched in Perichondrium, Osteoblast as well as
588 Tenocyte. We further performed fGWAS on our fine-grained annotations (Fig. 5b) to explore
589 subcluster enrichment. Consistent with the broad annotations, various types of chondrocytes
590 show significant enrichments for knee OA (Fig. 5b) whereas the signal for InterzoneChon
591 implicated *Early* IZ in hip OA. These findings implicate a greater impact for bone development
592 affecting hip OA risk, potentially through modulating hip shape and a thereby altered risk of
593 mechanical stress at the hip joint over time[91]. In contrast, the knee-specific findings point
594 toward the central role of chondrogenesis in knee pathology, potentially implicating postnatal
595 regenerative mechanisms.

596

597 Complex diseases are often associated with many mutations across non-coding regions. We
598 therefore sought after a way to integrate the GWAS signals of hip and knee OA described
599 above across enhancers of our gene regulatory network to identify enriched TFs and
600 pathways, while accounting for cell cluster specific effects. To this end we developed
601 SNP2CELL, a tool that uses gene regulatory networks as a basis to aggregate scores of
602 individual SNPs and cluster marker genes across pathways of connected TFs (Fig. 5c).

603 Aggregation of scores was achieved through a network propagation approach[92]. In contrast to
604 simply linking SNPs with their closest genes, this also allows for the identification of upstream
605 TFs which might not be mutated, but whose effect on multiple binding sites and target genes
606 is often disrupted, by integrating a large number of disease associations. By using eGRNs
607 previously computed with SCENIC+, we can utilise SNP signals mapped to both gene regions
608 and distal enhancers in ATAC peaks. By overlapping integrated SNP scores with integrated
609 marker gene scores per cell cluster, and comparing the effects to random permutations, we
610 obtain cell cluster specific enrichments of sub-networks of our initial eGRN.
611
612 Deriving insights from cell states implicated in fGWAS, we next clustered enrichment scores
613 for knee and hip OA, focusing on osteogenesis for the latter. In knee OA, various chondrocyte
614 subtypes clustered together by high enrichment score and average gene enrichments (Fig.
615 5d), mirroring its fGWAS enrichment for chondrocytes. Likewise for hip OA, Perichondrium
616 expectedly had the largest average enrichment across the osteogenic trajectory (Fig. 5e)
617 consistent with fGWAS results. Utilising SNP2CELL, we identified sub-networks for knee OA
618 and hip OA, prioritising articular chondrocytes (Fig. 5f) and perichondrium (Fig. 5g),
619 respectively, and revealing similar regulatory pathways that balance chondro and osteogenic
620 functions.
621
622 In the knee articular chondrocyte knee OA network (Fig. 5f), a number of non-TF genes with
623 roles in cartilage makeup and chondrocyte differentiation (*COL27A1, PRKCA, SNORC,*
624 *CRISPLD2*) were identified and predicted to be regulated by *NFATC1* and *FOXA3*. *COL27A1*,
625 and the proteoglycan *SNORC*, are involved in chondrocyte ECM makeup and maturation[93].
626 However, *PRKCA* is a kinase that has been linked to mechano-sensing in articular
627 chondrocytes[94]. Notably, *NFATC1* itself has been described as a marker of articular cartilage
628 progenitors affecting differentiation[95] and also protects against OA[96]. For hip perichondrium in
629 hip OA (Fig. 5g), the osteogenic key regulator *RUNX2* showed significant enrichment, along
630 with multiple *NFAT* genes (*NFATC1, NFATC2, NFATC4, NFAT5*), which in conjunction with
631 additional TFs (*ZEB1, MAF, TEAD1*) implicated calcineurin and *WNT* signalling pathways.
632 Both canonical and non-canonical *WNT* signalling via Ca^{2+}/Calcineurin/*NFAT* have been
633 linked to bone formation and remodelling, affecting the balance and differentiation of both
634 osteoblasts and osteoclasts [97]. Various *WNT* signalling inhibitors including *DKK1* and *FRZB*
635 have also been linked to the shaping process of the hip and osteoarthritis, which may be
636 particularly important during foetal development[98]. To provide a summary of the potential
637 functional role of the enriched disease-associated cell-cluster specific genes, we conducted
638 gene set enrichment analysis (GSEA) on Gene Ontology (GO) terms. This expectedly featured
639 terms related to extracellular matrix organisation, cartilage development and chondrocyte

640 differentiation in articular chondrocytes in knee OA specifically, whereas hip perichondrium
641 scores showed enrichment for inositol-phosphate, calcineurin and *NFAT* signalling, and
642 cellular response to lipid (Fig. 5h). The latter points to the potential interplay of lipids with bone
643 formation and osteoarthritis[99]. Overall, through application of fGWAS and our newly developed
644 pipeline SNP2Cell, we were able to distil differential involvement of the chondrogenic and
645 osteogenic lineages respectively in hip and knee OA, and provide a cellular basis for the link
646 between developmental cell states and regionally distinct adult disease.
647

648 **Deciphering monogenic conditions of the embryonic skeleton**

649 Both syndromic and non-syndromic forms of craniosynostosis are congenital conditions that
650 involve disturbances in cranial ossification and suture formation during fetal and postnatal
651 development. Disease-associated clinical features include premature cranial suture fusion,
652 and depletion of osteoprogenitor pools that facilitate the postnatal expansion of the cranium,
653 leading to severe global developmental consequences. Disease mechanisms that underlie
654 these changes are reportedly linked to missense mutations and haploinsufficiency in genes
655 that govern persistence of IM osteo-progenitor pools within the suture joints throughout the
656 cranium [100–102].
657

658 To implicate cell states that underlie craniosynostosis disease mechanisms we first cross-
659 referenced pseudotime-associated DEGs enriched in the IM and EC pathways against a
660 candidate database of over 2700 genes known to cause congenital conditions in humans[11].
661 We then filtered for enriched DEGs that were associated with musculoskeletal (MSK)
662 conditions and created a global view of enriched expression and accessibility across the
663 intersecting genes (Fig. 6a). Within these DEGs, we then filtered for craniosynostosis
664 associated genes by combining a subset of the database, and reported genes obtained from
665 Genomics England Limited rare and common craniosynostosis panels (Supplementary Table
666 5), of which a high proportion of the top DEGs were associated with synostosis (Fig. 6a). Most
667 of these associations (n=13/22) were within the progenitor populations of the IM trajectory
668 (SutureMes1/2 and *HHIP*+PreOB). Broadly, most of the craniosynostosis enriched genes
669 were highly accessible and expressed in IM progenitors, apart from *FN1 and IHH*, suggesting
670 the embryonic period is critically affected in craniosynostosis.
671

672 To elucidate potential effects of the synostosis-associated TFs on gene changes during
673 osteogenesis, we applied CellOracle to predict velocity shifts within the newly defined
674 trajectory and eGRNs using *in silico* perturbation simulation of gene knockouts in the IM
675 osteogenesis sub-compartment (Fig. 6b-c). Of the top-genes predicted to cause greatest
676 velocity shifts in the trajectory, *TWIST1, MSX2, LMX1B* were all disease-associated genes

677 that lead to highest velocity shifts in SutureMes2 when perturbed (Fig. 6c). The former two
678 were also IM pseudotime-associated DEGs (Fig. 6a), suggesting the velocity shift is highly
679 likely to reflect the loss-of-function mechanism of disease pathogenesis in these TFs [103,104].
680 Visualisation of the inferred velocity shifts revealed expected patterns of knockout favouring
681 lineage progression and bone formation (Fig. 6d), which are likely to penetrate in regions of
682 the coronal suture that indeed correspond to spatial enrichment of the candidate TFs (Fig. 6e).
683 Numerous other disease-associated TFs demonstrated maximal enrichment in the Suture
684 Zones (*IL11RA, SIX1*), whereas others preferentially affected more mature parts of the
685 trajectory (*IHH, ALPL, VLDLR*) (Fig. 6f). Overall, these findings constitute novel explanations
686 for features of TF-mediated craniosynostosis including premature suture fusion and
687 ossification.
688
689 To explore potential cell-extrinsic influences on fetal bone development, we curated a list of
690 65 clinically approved drugs from the chEMBL database which carried warnings of
691 teratogenicity (Supplementary Table. 6). We then utilised to score target gene-group
692 expression per cell state within our osteogenic sub compartment. Interestingly, this revealed
693 overall greater target enrichment of teratogenic drugs within IM progenitors and downstream
694 osteoblast/osteocyte cell states compared to EC progenitors (Extended Data Fig. 9a). Retinoid
695 drugs expectedly enriched for targets in Osteocytes, consistent with observations of retinoic
696 acid enhancing *in vitro* osteogenesis in mouse suture cells[105]. Likewise, the well-known
697 skeletal development teratogen thalidomide enriched for targets in Osteocytes. Aside from
698 antihypertensives targeting the renin-angiotensin pathway (ACEi, ARB, Renin inhibitor) which
699 enriched for targets in SutureMes1 and *HHIP*$^{+}$PreOB, we also identified high enrichment for
700 endothelin receptor antagonists (ERAs) specifically in SutureMes1. Mice lacking the
701 endothelin A receptor recapitulate wide-spread cranial neural crest-related defects *in utero*[106].
702 Informed by teratogenic effects of ERA use in animals[107], they are not routinely prescribed for
703 pulmonary arterial hypertension (PAH) during pregnancy. However, there remains a lack of
704 evidence for teratogenicity in humans, and women of child-bearing age continue to utilise them
705 for treatment of PAH. Our analyses support the notion of potential targets for ERAs within the
706 human suture cells and lend support to a contraindication for their use during pregnancy.
707
708 To further resolve the regulatory relationship between the prioritised TFs (*TWIST1, MSX2,*
709 *LMX1B)* within the osteo-lineage, we utilised our paired snRNA-seq and scATAC-seq dataset
710 to reconstruct interaction networks across the TFs (Fig. 6g). Our predictions revealed a rich
711 web of inter-regulation of shared coding and non-coding targets across the three candidate
712 TFs. Of these connected nodes, numerous TFs are associated with loss-of-function mutations
713 leading to craniosynostosis (*SIX1, TCF12, NFIX, ALX4*), suggesting a tightly regulated

network conferring suture patency, inhibiting osteogenesis in this region. Notably, *TCF12* was reported to govern normal coronal suture development through heterodimer formation with *TWIST1*, and severe phenotypes are observed in mice with doubly deleterious mutations[103]. *TCF12* served as a node connecting *TWIST1* to *RUNX2*, suggesting that disruption of this critical co-regulatory network leads to human disease. While direct transcriptional regulations of *RUNX2* by the three inhibitory TFs were predicted to be weaker than their stronger inter-regulation and not visualised in Fig. 6g, predicted regulations of *RUNX2* (Extended Data Fig. 9b) therefore likely contribute to known complex regulatory mechanisms, including protein-level interactions[108]. In addition, cell-extrinsic signals and control mechanisms are likely to act on the connecting nodal points such as *NFATC2* [97], governing the balance of osteogenesis.

Multi-omic information additionally allows exploration of enhancer-mediated regulation of transcription. This is of particular importance for monogenic bone-forming conditions such as Van Buchem disease (VBD) whereby mutation in the non-coding *ECR5* enhancer of *SOST* leads to sclerosing dysplasia of bone[109]. We therefore constructed a regulatory network centred around *SOST* and identified the *ECR5* enhancer to directly regulate *SOST* (Fig. 6h). Numerous osteogenesis-associated TFs were predicted to regulate the region containing the ECR5 enhancer (*DLX5, KLF2, KLF4, KLF13, MEF2C)*. Notably, *MEF2C* has previously been shown to regulate *ECR5* in mice[110], confirming its direct role in controlling *SOST* transcription. Overall, these predictions form a basis for constructing cellular models of TF and enhancer-driven monogenic conditions of the bone lineage.

Discussion

We present the first multiomic cell atlas to capture the spatially-resolved cellular composition of bone, synovial and suture joint development across the first trimester in humans. First, we dissected the cellular taxonomy of interzone cells, including *GDF5*-expressing progenitors in the synovial joints, and developed ISS-Patcher, a tool to aid inference of IZ subclusters in our high-resolution 155-plex ISS data. We highlighted Articular, Fibro (*GDF5*$^{+}$) and (*GDF5*$^{-}$*)* Hypertrophic IZ subclusters which zonate the articular chondrocyte and cartilage template-forming boundaries of the embryonic synovial joints, respectively.

Within lineages described in mice, *GDF5*$^{+}$ cells are posited to give rise to the numerous cell components within the joint including ligaments and tendons. Here we describe a novel two-stage process whereby the fibroblasts and tenocytes of the synovial joints arise from embryonic *HIC1*$^{+}$ fibroblasts, a population previously reported in the mouse embryonic limb [29], and a fetal (*Pl16*$^{+}$) fibroblast progenitor which is transcriptionally similar to human postnatal

750 universal fibroblasts [30]. Together, these shed new light on the cellular architecture of the
751 developing human embryonic synovial joint in the first trimester.
752
753 We have identified numerous previously undescribed human cranial embryonic osteo-lineage
754 cell states including CranialMes (*HAND2*)[111,112], FacialMes (*PAX3*)[51] and PArchMes (*LHX8*)[113],
755 by their expression of comparable marker genes in mice counterparts. We leveraged our
756 spatial analyses of the fetal cranium to define two novel fetal *TWIST1*$^+$ SutureMes populations,
757 and *HHIP*-expressing pre-osteoblast which contributes to osteogenesis, and mirrors
758 populations of the fetal mouse suture [30,54]. We also discovered distinct IM and EC osteogenic
759 cellular trajectories in their respective regional niches. Using *OrganAxis*, we illustrate the
760 spatial component of these trajectories as they develop and show a regulatory shift in the
761 poised osteogenic suture population as it progresses in space toward maturing pre-
762 osteoblastic cell states where we uncover a region that critically recruits endothelial cell states
763 to drive ossification. We show that the EC niche in the appendicular anlagen derives from an
764 undifferentiated progenitor expressing *ISL1* and *TBX5*, which mark an equivalent population
765 in the human limb [11], and proceeds to form perichondrium and osteoblasts through
766 developmental time and space in the first trimester. These novel cellular trajectories form a
767 valuable reference for future studies of bone development. Further studies of the second
768 trimester, when adipose cell states and mature ligamentous tissues develop, will facilitate a
769 complete census of the joint.
770
771 Through our application of a modified workflow to conduct somatic mutation tracing within our
772 diverse chondrocyte subcompartment, we lend support to a potential myoprogenitor-related
773 origin for the newly reported *PAX7*$^+$ chondrocyte, and likewise suggest that a novel *SOX9*$^+$
774 Schwann cell state may confer chondrogenic potential. These proposed cross-lineage cell
775 fates uncover potential similarities to mouse development described above, and crucially, the
776 value of computational tool development in unravelling comparative human biology. In the
777 future, more focused investigations that involve isolation of these cell lineages will further
778 inform their biology.
779
780 Leveraging our dataset, we were able to identify developmental links to both monogenic and
781 complex musculoskeletal diseases. We observed marked differences in the enrichment of hip
782 and knee OA GWAS signals across cell clusters, implicating bone formation in the former, and
783 chondrogenesis in the latter, which we traced back to the osteogenic perichondrium, and
784 numerous chondrogenic cell states, respectively. In future, this framework may be applied to
785 other conditions to gain insights into cell-state specific pathogenesis. Lastly, by systematically
786 simulating *in silico* KO of TFs known to be associated with craniosynostosis, we identified a

network of regulators inhibiting osteogenesis, which provides mechanistic detail of how previously reported monogenic loss-of-function mutations in, for example, *TWIST1, MSX2 and LMX1B,* may lead to early suture fusion. This approach, applied to the newly mapped cross-region osteogenic trajectory, holds immense value in predicting potential effects of gene changes across other diseases involving the cellular lineage.

Our comprehensive multimodal human fetal skeletal atlas is a unique fundamental resource for the understanding of human cartilage and bone development in the first trimester. It also has the potential to serve as a cellular roadmap for *in vitro* endeavours to differentiate osteoblast and chondrocyte and mesenchymal cell states in the developing musculoskeletal system.

Acknowledgements

We thank the donors and their relatives who contributed to this study. We thank the staff of the Teichmann group wet lab team, Cellular Genetics Informatics team, Cytometry Core Facility team, Core DNA Pipelines team, and Spatial Genomics Platform team at the Wellcome Sanger Institute for their support. We also thank M. Storer, S. Kimber, J.E.G. Lawrence, A. Maartens for assistance in manuscript writing, A. Oszlanczi and A. Wilk for sample management and administrative assistance, C. Battrum for administrative assistance and S. Davidson for helpful discussions.

This work was made possible in part by the Wellcome Trust (220540/Z/20/A, 221052/Z/20/Z, 221052/A/20/Z, 221052/B/20/Z, 221052/C/20/Z, and 221052/E/20/Z); the NIHR Cambridge Biomedical Research Centre (NIHR203312) – (The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care); the MRC Clinical Research Training Fellowship to K.T; the Wellcome Trust Sanger Institute Postdoctoral Fellowship to LF; the Marie Curie Actions Individual Fellowship to J.P.P.

Author contributions

K.T., L.F. AND J.P.P. conceived the study, performed data generation and experiments, conceived methodology, performed analysis, collected biological samples, and wrote the paper. K.R. performed data generation, conceived methodology, collected biological samples, performed analysis, conducted imaging, and wrote the manuscript. T.L., N.Y. performed computational analysis and conceived methodology. P.H. performed computational analysis and provided key biological interpretation. C.X. performed computational analysis and provided key biological interpretation. J.C. performed experiments, conceived methodology, and collected samples. R.L.. performed computational analysis and provided key biological interpretation. K.K. performed computational analysis and provided key biological interpretation. N.H. performed analysis and provided key biological interpretation. S.M. performed analysis and provided key biological interpretation. L.R., R.K., S.P., A.W.C. performed experiments and sample processing and collection, generating single-cell data. E.T, I.M. and F.M. performed imaging experiments, provided key biological insights, and collected tissue samples, and generated spatial data. B.C., A.V.P., D.H., S.M., M.P. and P.M. performed computational processing of data and performed data analysis. X.H. conducted tissue processing and sample collection. K.B.M. conceived computational methodology and provided key biological interpretations. M.H. conceived the study and provided key biological interpretations. R.A.B. conceived the study and provided key biological interpretations. O.B.

833 conceived the imaging experiments, provided key biological interpretations and analysed the
834 data. C.D.B. conceived the study, conceived methodology, provided key biological
835 interpretation, and supervised the project. S.A.T. conceived the study, conceived
836 methodology, provided key biological interpretation, and supervised the project.
837

838 Conflict of interest

839 C.D.B is a founder of Mestag Therapeutics. S.A.T. has consulted for or been a member of
840 scientific advisory boards at Qiagen, Sanofi, GlaxoSmithKline, and ForeSite Labs. She is a
841 consultant and equity holder for TransitionBio and EnsoCell. The remaining authors declare
842 no competing interests.

843

844 Data availability

845 High-throughput raw sequencing data in this study is available from ArrayExpress
846 (www.ebi.ac.uk/arrayexpress) with accession numbers. Processed snRNA-seq/scATAC-seq
847 data, visium and ISS data are available for visualisation and downloading from
848 **https://developmental.cellatlas.io/skeleton-development**

849

850 Code availability

851 ISS-Patcher is available at **https://github.com/Teichlab/iss_patcher .**
852 SNP2CELL is available at **https://github.com/Teichlab/snp2cell.**
853

854 Materials and Methods

855 **Sample acquisition and ethics**
856 Developing human limb and cranium tissue samples were obtained from elective terminations
857 under REC 96/085 (East of England - Cambridge Central Research Ethics Committee). Briefly,
858 samples were kept suspended in PBS and at -4°C on ice during dissection. Shoulder,
859 hip and knee joints were dissected en-bloc from the limbs. For the shoulder joint a
860 proximal incision was made at the distal third of the clavicle, and a distal incision was
861 created at the neck of the humerus. For embryonic shoulder samples where distinctive
862 bone features had not formed, approximations were made to capture the entirety of
863 the glenohumeral and acromioclavicular joints. For the cranium samples (<8 PCW),

two regions were dissected for each of the calvaria and skull base, separated at the posterior border of the frontal bone in both cases. For older cranial samples (>8 PCW), tissues were dissected to separate the frontal, parietal, sphenoid, ethmoid, occipital and temporal bones where feasible. Samples were initially embedded in optimal cutting temperature medium (OCT) and frozen at -80°C on an isopentane-dry ice slurry. Cryosections were then cut at a thickness of 10 µm using a Leica CM1950 cryostat and placed onto SuperFrost Plus slides (VWR) for ISS or Visium CytAssist, or used directly for single-nuclei processing.

In situ sequencing (ISS) and high-resolution imaging

ISS was performed using the 10X Genomics CARTANA HS Library Preparation Kit (1110-02, following protocol D025) and In Situ Sequencing Kit (3110-02, following protocol D100), which comprise a commercialised version of HybRISS[117]. Probe panel design was based on fold-change thresholds in cell states of the limbs (Supplementary Table 3). Briefly: cryosections of developing limbs were fixed in 3.7% formaldehyde (Merck 252549) in PBS for 30 minutes and washed twice in PBS for 1 minute each prior to permeabilization: sections were briefly digested with 0.5 mg/ml pepsin (Merck P7012) in 0.1 M HCl (Fisher 10325710) at 37°C for 15 seconds (5 PCW) or 30 seconds (6 PCW and older), then washed twice again in PBS, all at room temperature. Following dehydration in 70% and 100% ethanol for 2 minutes each, a 9 mm diameter (50 µl volume) SecureSeal hybridisation chamber (Grace Bio-Labs GBL621505-20EA) was adhered to each slide and used to hold subsequent reaction mixtures. Following rehydration in buffer WB3, probe hybridisation in buffer RM1 was conducted for 16 hours at 37°C. The 158-plex probe panel included 5 padlock probes per gene, the sequences of which are proprietary (10X Genomics CARTANA). The section was washed with PBS-T (PBS with 0.05% Tween-20) twice, then with buffer WB4 for 30 minutes at 37°C, and thrice again with PBS-T. Probe ligation in RM2 was conducted for 2 hours at 37°C and the section washed thrice with PBS-T, then rolling circle amplification in RM3 was conducted for 18 hours at 30°C. Following PBS-T washes, all rolling circle products (RCPs) were hybridised with LM (Cy5 labelling mix with DAPI) for 30 minutes at room temperature, the section was washed with PBS-T and dehydrated with 70% and 100% ethanol. The hybridisation chamber was removed and the slide mounted with SlowFade Gold Antifade Mountant (Thermo S36937). Imaging of Cy5-labelled RCPs at this stage acted as a QC step to confirm RCP (‘anchor’)

898 generation and served to identify spots during decoding. Imaging was conducted using
899 a Perkin Elmer Opera Phenix Plus High-Content Screening System in confocal mode
900 with 1 µm z-step size, using a 63X (NA 1.15, 0.097 µm/pixel) water-immersion
901 objective. Channels: DAPI (excitation 375 nm, emission 435-480 nm), Atto 425 (ex.
902 425 nm, em. 463-501 nm), Alexa Fluor 488 (ex. 488 nm, em. 500-550 nm), Cy3 (ex.
903 561 nm, em. 570-630 nm), Cy5 (ex. 640 nm, em. 650-760 nm).
904
905 Following imaging, each slide was de-coverslipped vertically in PBS (gently, with
906 minimal agitation such that the coverslip 'fell' off to prevent damage to the tissue). The
907 section was dehydrated with 70% and 100% ethanol, and a new hybridisation chamber
908 secured to the slide. The previous cycle was stripped using 100% formamide (Thermo
909 AM9342), which was applied fresh each minute for 5 minutes, then washed with PBS-
910 T. Barcode labelling was conducted using two rounds of hybridisation, first an adapter
911 probe pool (AP mixes AP1-AP6, in subsequent cycles), then a sequencing pool (SP
912 mix, customised with Atto 425), each for 1 hour at 37°C with PBS-T washes in between
913 and after. The section was dehydrated, the chamber removed, and the slide mounted
914 and imaged as previously. This was repeated another five times to generate the full
915 dataset of 7 cycles (anchor and 6 barcode bits).
916
917 **Multiplexed smFISH**
918 Cryosections were processed using a Leica BOND RX to automate staining with the
919 RNAscope Multiplex Fluorescent Reagent Kit v2 Assay (Advanced Cell Diagnostics,
920 Bio-Techne), according to the manufacturers' instructions. Probes may be found in
921 Supplementary Table 7. Prior to staining, fresh frozen sections were post-fixed in 4%
922 paraformaldehyde in PBS for 45 minutes at 4°C, then dehydrated through a series of
923 50%, 70%, 100%, and 100% ethanol, for 5 minutes each. Following manual pre-
924 treatment, automated processing included digestion with Protease III for 15 minutes
925 prior to probe hybridisation. Tyramide signal amplification with Opal 520, Opal 570,
926 and Opal 650 (Akoya Biosciences) and TSA-biotin (TSA Plus Biotin Kit, Perkin Elmer)
927 and streptavidin-conjugated Atto 425 (Sigma Aldrich) was used to develop RNAscope
928 probe channels. Stained sections were imaged as for ISS above.
929

930 **Image-based in situ sequencing (ISS) decoding**

931 We employed the ISS decoding pipeline outlined in Li et al [118]. This pipeline consists
932 of five distinct steps. Firstly, we performed image stitching using Acapella scripts
933 provided by Perkin Elmer, which generated two-dimensional maximum intensity
934 projections of all channels for each cycle. Next, we employed Microaligner[119] to
935 register all cycles based on DAPI signals using the default parameters. For cell
936 segmentation, we utilised a scalable algorithm that leverages CellPose[120] as the
937 segmentation method. The expected cell size is set to 70 pixels in diameter and further
938 expanded 10 pixels to mimic the cytoplasm. To decode the RNA molecules, we
939 employed the PoSTcode algorithm[121] with the following parameters: rna_spot_size=5,
940 prob_threshold=0.6, trackpy_percentile=90, trackpy_separation=2. Furthermore, we
941 assigned the decoded RNA molecules to segmented cells using STRtree and
942 subsequently generated AnnData objects following the approach described by Virshup
943 et al. [122]. Finally, only cells with more than 4 RNA molecules were kept for downstream
944 analysis.

945

946 **Visium processing and library preparation**

947 Visium CytAssist Spatial Gene Expression for Fresh Frozen (10x Genomics) was
948 performed following the manufacturer's protocol. Regions of interest (ROIs) were
949 selected based on the presence of microenvironments of bone-formation relevant to
950 the droplet data (e.g. coronal suture) and aligned to the CytAssist machine gasket
951 accordingly. Images were captured using a Hamamatsu S60 slide scanner at 40x
952 magnification prior to conducting the Visium CytAssist protocol for subsequent
953 alignment. Libraries were mapped with SpaceRanger (10x Genomics).

954

955 **Single-nuclei isolation and library preparation**

956 Single-nuclei were isolated from fresh frozen samples through cryosectioning followed
957 by mechanical dissociation as described in previous work [123]. Briefly, 10 µm sections
958 were homogenised in homogenization buffer (250 mM sucrose, 25 mM KCl, 5 mM
959 $MgCl_2$, 10 mM Tris-HCl, 1 mM dithiothreitol (DTT), 1× protease inhibitor, 0.4 U μl^{-1}
960 RNaseIn, 0.2 U μl^{-1} SUPERaseIn and 0.1% Triton X-100 in nuclease-free water) using
961 a glass Dounce tissue grinder set (Merck). Samples were dissociated with 10-20
962 strokes of a loose pestle 'A' followed by 10 strokes of a tight pestle 'B' when tissue

963 fragments remained . The resulting mixture was passed through a 50 μm cell strainer,
964 followed by centrifugation (500 g, 5 mins) the pellet was then resuspended in 300 μl
965 of storage buffer (1× PBS, 4% BSA and 0.2 U μl^{-1} Protector RNaseIn) and passed
966 through a 20 μm cell strainer. Nuclei were visualised and assessed for viability under
967 microscopy following staining with trypan blue solution and were further processed for
968 10x Genomics Chromium Single Cell Multiome ATAC + Gene Expression according
969 to the manufacturer's protocol. Nuclei suspensions were loaded with a targeted nuclei
970 recovery of 10,000 droplets per reaction. For some of the nuclei samples, mixtures of
971 different donors were pooled within one reaction and later demultiplexed by genotype.
972 Quality control of cDNA and final libraries was done using Bioanalyzer High Sensitivity
973 DNA Analysis (Agilent). Libraries were sequenced using a NovaSeq 6000 (Illumina)
974 with a minimum sequencing depth of 20,000 read pairs per droplet.
975
976 **Data preprocessing**
977 Sequencing data were aligned to the human reference genome (GRCh38-2020-A-
978 2.0.0) using CellRanger-ARC software (v.2.0.0). The called barcodes from 10x
979 Multiome lanes with pooled genotypes from multiple donors were demultiplexed per
980 Genotype using BAM outputs through Souporcell (v2.0) [124]. Subsequently, the
981 Souporcell outputs were clustered by genotype for metadata assignment to each
982 barcode. Visium data were mapped to SpaceRanger (v. 1.1.0) using default input
983 settings, and low-res Cytassist images were aligned to hi-res microscopy images of
984 the processed slides using 10x Genomics LoupeBrowser (v 7.0) according to capture
985 frame marker regions. For gene expression data, SoupX [125] was applied to remove
986 background ambient RNA. For cellranger-arc called matrices that contained >16,000
987 droplets (exceeding the number expected from targeted droplet recovery) we
988 increased the estimated global rho value by 0.1 to account for the potential of
989 additional ambient RNA. Scrublet [126] was applied to estimate doublet probability and
990 a score of >0.3 was used as a cutoff value. Droplets were filtered for >200 genes, and
991 <5% mitochondrial and ribosomal reads.
992
993 For scATAC-seq, we applied ArchR[127] (v1.0.2) to process the outputs from Cellranger-
994 atac. Initial per-cell quality control was performed considering the number of unique
995 nuclear fragments, signal-to-background ratio and the fragment size distribution.

Moreover, cells with TSSenrichment score<7 and nFrags< 1000 were removed. Doublets were discarded using the default settings. Initial clustering was performed at resolution = 0.2 with top 40 dimensions from Iterative Latent Semantic Indexing (LSI). Then, pseudo-bulk replicates were made for each broad cell type per region from the initial clustering results. Peak calling (501 bp fixed-width peaks) was performed based on pseudo-bulk coverages by macs2. Then, a cell-by-peak count matrix was obtained and exported. We applied muon[128] (0.1.2) for normalization, LSI dimension reduction, and clustering analysis with BBKNN[129] to correct the batch effects from anatomical regions and donors for obtaining ATAC embedding. Gene scores based on chromatin accessibility around gene bodies were calculated. Moreover, we applied MultiVI[130] to get a joint embedding for snRNA-seq and scATAC-seq.

Doublet detection

To interrogate doublet capture, we conducted an adapted scrublet workflow that was previously described[131]. Briefly, per-droplet scrublet scores were first determined for cellranger-arc count-matrices from each 10x multiome (gene-expression) lane independently. The droplets were then overclustered through the standard scanpy workflow using default parameters up to leiden clustering. Each individual cluster was further clustered. A per-cluster median of scrublet scores were computed. A normal distribution of doublet score, centred at the score median with a standard deviation estimated from the median absolute deviation (MAD) was used to compute p-values for each of the clusters. After false-discovery rate adjustment using benjamini-hochberg correction, a p-value of >0.6 was deemed as a cutoff value of good quality cells as doublets are significant outliers.

Cell cluster annotation

We adopted a hierarchical clustering approach by first conducting Leiden clustering on the global integrated scVI (hidden layers=256, latent variables=52, dispersion='gene-batch') RNA embeddings to obtain broad clusters. To validate these we utilised Celltypist to train a model on cell states in the embryonic limb bud [11,132,133], and transferred labels onto our embedding for inspection. We utilise this information in addition to canonical marker genes, to annotate broad clusters and subset sub-lineages. For sub-lineages e.g. (Chondrocytes, Fibroblasts, Bone-related populations, Schwann Cells, Immune cells, Endothelial cells), we further embedded these using

1030 scVI (hidden layers=256, latent variables=52, dispersion='gene-batch') and conducted
1031 Leiden clustering (resolution = 0.6), followed by DEG analyses (method='wilcoxon') to
1032 obtain cluster markers. We additionally utilised the inferred spatial location of cell
1033 states (described below) to inform annotations.
1034
1035 **Spatial mapping using Cell2location**
1036 We performed Cell2location for deconvolution of visium cytassist using annotated
1037 snRNA-seq data from skull base and calvaira cell states as input. The donor was used
1038 as batch variables, and the libraries were considered as covariates in the regression
1039 model. For spatial mapping, we estimated 30 cells per voxel based on histological
1040 data, and hyperparameter detection_alpha of 20 were applied for per-voxel
1041 normalisation.
1042
1043 **ISS Patcher**
1044 ISS_patcher is a simple package for approximating features not experimentally
1045 captured in low-dimensional data based on related, high-dimensional data. It was
1046 developed as an approach to approximate expression signatures for genes missing in
1047 ISS data using matched snRNA-Seq data as a reference in this study. First, a shared
1048 feature space between both datasets was identified by subsetting the 155-158 genes
1049 present in the ISS pool, followed by separate normalisation to median total cell counts,
1050 log-transformation and z-scoring for both modalities. Then, the 15 nearest neighbours
1051 in scRNA-Seq space were identified for each ISS cell with the Annoy python package,
1052 and the genes absent from ISS were imputed as the average raw counts of the scRNA-
1053 Seq neighbours.
1054 **Visium axis annotation using OrganAxis**
1055 Our Visium cranium sample was annotated with TissueTag by semi-automatic mode
1056 to generate a one-dimensional maturation axis. Regions of the developing bone were
1057 first manually annotated based on H&E features. Tissue regions that did not include
1058 bone-forming niches were excluded from annotation. The annotation categories that
1059 were stored included multiple regions of the coronal suture (level_0 to level_2
1060 annotation), stemming from the central-most portion, an osteogenic front (level_3
1061 annotation) harbouring histological features of osteoprogenitors and osteogenic zones
1062 (level_4 to level_7 annotation) from the emergence of histological osteoblasts. All
1063 annotations were saved as TissueTag output format, which logs the annotation

1064 resolution, the pixels per µm (ppm) as well as pixel value interpretation of annotation
1065 names (e.g. 0 = “Suture”) and colours (e.g. “Osteogenic Front”: “red”). To robustly and
1066 efficiently migrate TissueTag annotations to the Visium objects, we first transferred
1067 TissueTag annotations from pixel space to a high-resolution hexagonal grid space (15
1068 µm spot diameter and 15 µm point-to-point centre distance with no gap between spots)
1069 based on the median pixel value of the centre of the spot (radius/4) in the annotated
1070 image. Next, to generate continuous annotations for Visium data we measured for
1071 each spot in the hexagonal high-resolution grid the mean Euclidean distance to the 10
1072 nearest points from each annotated structure in the level_0 annotation as well as the
1073 distance from the closest point for structures in annotation_level_1. All annotations
1074 were mapped to the Visium spots by proximity of the spot annotation grid to the nearest
1075 corresponding spot in the Visium array.

1076

1077 **Gene regulatory network analysis**

1078 The SCENIC+ [134] pipeline was used to predict transcription factors and putative target
1079 genes as well as regulatory genomic regions harbouring binding sites. The fragment
1080 matrix of peaks called with macs2 and processed within ArchR [127] together with the
1081 corresponding RNA count matrix were used as inputs. Meta-cells were created by
1082 clustering cells into groups of around 10-15 cells based on their RNA profiles and
1083 subsequent aggregation of counts and fragments. The pipeline was applied to subsets
1084 of the dataset corresponding to individual lineages: First, CisTopic was applied to
1085 identify region topics and differentially accessible regions from the fragment counts as
1086 candidate regions for TF binding. CisTarget was then run to scan the regions for
1087 transcription factor binding sites and GRNBoost2 [135] was used to link TFs and regions
1088 to target genes based on co-expression/accessibility. Enriched TF motifs in the
1089 regions linked to target genes were used to construct TF-region and TF-gene
1090 regulons. Finally, regulon activity scores (AUC) were computed with AUCell based on
1091 target-gene expression and target region accessibility, and regulon specificity scores
1092 (RSS) derived from them. Networks of TFs, regions and target genes (eGRNs) were
1093 constructed by linking individual regulons. TF-enhancer-gene links for all subsets
1094 (osteogenesis, chondrogenesis, fibrogenesis, early joint progenitors, immune,
1095 schwann) can be found in Supplementary Table 9.

1096

1097 **Trajectory analysis**

1098 For trajectory construction in the osteogenic subcompartment, non-cycling droplets
1099 were subsetted and X_scVI were used as projections for palantir in order to obtain
1100 multiuscale diffusion space. A neighbourhood graph was generated on the diffusion
1101 space using sc.pp.neighbors, and the first two PCs were used as initial positions to
1102 create ForceAtlas2 embeddings using sc.tl.draw_graph. The FA embeddings were
1103 exported into R, and monocle3 [136,137] was used to find a principal graph and define
1104 pseudotime. Differentially expressed genes were then computed along pseudotime
1105 using a graph-based test (morans' I) [138,139] and the principal graph in monocle3, which
1106 allows identification of genes upregulated at any point in pseudotime. The results were
1107 visualised with heatmaps using the complexHeatmap [140] and seriation [141] packages,
1108 after smoothing gene expression with smoothing splines in R (smooth.spline, df=12).
1109 Velocity analysis[142] was performed using scvelo[143] version 0.2.3. Spliced and
1110 unspliced read counts were computed with velocyto from the unprocessed data,
1111 before using scvelo.pp.moments, scvelo.tl.velocity and scvelo.tl.velocity_graph to
1112 compute velocities for the preprocessed droplets.

1113

1114 **In silico TF perturbations**

1115 CellOracle [144] was used with the osteogenesis trajectory created with scFates [145] and
1116 the regulons predicted with SCENIC+ [134] for the same cells were imported into
1117 CellOracle as a base GRN. Cells were aggregated into meta-cells of 10-15 cells and
1118 linear models explaining TF- from target gene expression were fitted with CellOracle
1119 per cell cluster. Regulon-based TF perturbation vectors were inferred using the cell
1120 cluster-specific models to predict effects of TF overexpression and knockout. Diffusion
1121 pseudotime [146] was then computed for IM and EC lineages separately by selecting
1122 corresponding starting points. The pseudotime gradients were used to derive
1123 pseudotime-based differentiation vectors, and the pseudotime-perturbation vector
1124 cross-product was computed to obtain perturbation scores (PS). These PS indicate,
1125 whether the *in silico* perturbation of a TF is consistent with or opposes differentiation
1126 along a lineage (osteogenesis). The simulations were carried out systematically,
1127 overexpressing and knocking out all TFs in the GRN. For each TF and condition the
1128 PS were then averaged per cell cluster and summarised in a table to screen for TFs
1129 promoting or inhibiting osteogenesis.

1130

1131 **fGWAS**

1132 Functional GWAS analysis (fGWAS) [147] was applied to identify disease relevant cell
1133 clusters as described in detail in [148] (https://github.com/natsuhiko/PHM). The model
1134 makes use of full summary statistics from GWAS studies, linking SNPs to genes and
1135 captures a general trend between gene expression and disease association of linked
1136 loci for each cell cluster. At the same time, it also corrects for linkage disequilibrium
1137 (LD) and other relevant factors. We used full GWAS summary statistics obtained from
1138 the EBI GWAS catalog, open targets and knee and hip osteoarthritis (OA) as well as
1139 total knee (TKR) and hip (THR) replacement from [149] (Supplementary Table 8).

1140

1141 **SNP2Cell**

1142 We used a network propagation [92] approach to integrate GWAS summary statistics
1143 and cell cluster marker gene based scores for prioritising disease relevant and cell
1144 cluster specific subunits of our TF-network. First, scores per SNP were computed from
1145 downloaded summary statistics and weighted by linkage-disequilibrium. Then the
1146 scores were mapped to a gene regulatory network, here an eGRN computed with
1147 SCENIC+ for the corresponding lineage. Since the used networks contain TFs and
1148 target genes, and also regions with TF-binding sites as nodes, SNP scores were
1149 mapped both to genes and regions, representing distal regulatory elements. The
1150 scores were then propagated across the network using a random walk with restart (or
1151 personalised page-rank) process. This integrates the contribution of individual SNPs,
1152 with signals converging around relevant network nodes. The procedure was repeated
1153 with 1000 randomly permuted scores to compute permutation-test results and z-
1154 scores. Next, differential expression based marker gene scores for each cell cluster
1155 were propagated in the same way, resulting in cell-cluster specificity scores for each
1156 network node. The SNP and expression based scores were then combined per cell
1157 cluster as in [150] by using the minimum for each node. The final scores were
1158 thresholded and the resulting connected components obtained as enriched sub-
1159 networks. The method has been compiled into a tool we called SNP2CELL and is
1160 available at (https://github.com/Teichlab/snp2cell).

1161

1162 **Mutation calling at single cell resolution**

1163 Scomatic was applied to detect somatic single-nucleotide mutations in snRNA-seq and
1164 snATAC-seq datasets. We first split alignment files into broad cell cluster-specific bam

1165 files per donor per modality. Then a base count matrix indexed by cell cluster and
1166 genomic coordinates was created and merged. We applied somatic variant calling for
1167 each cell cluster based on the matrix to get more mutations for each modality with
1168 snATAC. Then beta binomial test was employed to discount sites affected by recurrent
1169 technical artefacts as somatic mutations. We also applied additional filters based on
1170 Panel of Normals (PON)) to remove germline polymorphisms and recurrent artefacts.
1171 At least 5 reads were covered to consider a callable site and for mutated sites, it is
1172 supported by at least three reads from at least two cells from the same cell cluster as
1173 mutations in cell cluster level. We set max_cell_types to 10000 to disable the filter cell
1174 cluster number filters as we applied Genome Aggregation Database (gnomAD) to
1175 remove suspicious germline mutations as a second criterion. We kept all the other
1176 filtered mutations instead of using PASS mutations to keep more mutations for
1177 downstream analysis. Mutations from RNA and ATAC were aggregated together. We
1178 also applied the function 'SingleCellGenotype.py' to compute the genotype for each
1179 cell at the variant sites with default parameters.
1180 In the chondro-lineages and the Schwann compartments, we filtered mutations based
1181 on the cell cluster level first. Shared mutations per donor require them to mutate in
1182 only both chondro-lineages and the Schwann compartment but also to be callable in
1183 at least one reference cell clusters across all regions (Immune, Endothelial, Epithelial,
1184 Erythroid, PArchMes, Lymphatic endothelial, *PAX7*$^{+}$ Myo, *MYH3*$^{+}$ Myo, Synovial Fibro,
1185 *PI16*$^{+}$ FibroPro, DermFibro, MyoFibro, HIC1^{+} Mes, Tenocyte, LimbMes, FacialMes,
1186 Meningeal Dura, Periosteal Dura, LEPR Mese, Perisuture, SutureMes1, SutureMes2,
1187 CranialMes), which rules out the germline mutations again. Then we exact mutated
1188 single cells based on the shared mutations. Cells with at least 2 reads that exhibited
1189 mutations are treated as mutated cells. Finally, we aggregated all the mutated cells
1190 across all donors and calculated the density of cells in an embedded space using
1191 'tl.embedding_density' function in scanpy to get normalised density values.
1192 For shared mutations between *PAX7*$^{+}$ Chondrocyte and muscle lineages, we filtered
1193 mutations that only mutated in *PAX7*$^{+}$ chondrocyte and the whole muscle compartment
1194 and the reference cell clusters were Immune, Endothelial, Epithelial, Erythroid, IZ,
1195 PArchMes, Lymphatic endothelial, Synovial Fibro, PI16hiFibroPro, DermFibro,
1196 MyoFibro, HIC1hiMes, Tenocyte, LimbMes, FacialMes, Meningeal Dura, Periosteal
1197 Dura, LEPR Mes, Perisuture, SutureMes1, SutureMes2, CranialMes, Neural,
1198 Schwann, and Perichondrium.

1199 **Cell-cell interactions**

1200 Ligand-receptor interactions were inferred using 'cpdb_analysis_method.call' in
1201 CellPhoneDB (v4.0.0). We only considered genes with more than 10% of cells
1202 demonstrating expression within each cell cluster considered and interactions with P
1203 >0.001 were removed. We used NicheNet (v1.1.1) to identify different interactions
1204 between endochondral ossification and intramembranous niches. We first calculate
1205 differentially expressed genes (DEGs) of osteoblastic cell clusters and tip cells
1206 between two niches using standard Seurat Wilcoxon test and min LFC per cell cluster
1207 were used to summarise the differentially expressed ligands and receptors. Then, top
1208 1000 DEGs were applied for calculating the ligand activities. We prioritised the ligand-
1209 receptor links with default settings. Top 10 ligands and their top scoring receptors were
1210 visualised in heatmaps.

References

1. O'Rahilly, R. & Gardner, E. The initial appearance of ossification in staged human embryos. *Am. J. Anat.* **134**, 291–307 (1972).
2. Faro, C., Benoit, B., Wegrzyn, P., Chaoui, R. & Nicolaides, K. H. Three-dimensional sonographic description of the fetal frontal bones and metopic suture. *Ultrasound Obstet. Gynecol.* **26**, 618–621 (2005).
3. Edwards, J. C. *et al.* The formation of human synovial joint cavities: a possible role for hyaluronan and CD44 in altered interzone cohesion. *J. Anat.* **185**, 355 (1994).
4. Pazzaglia, U. E. *et al.* Long bone human anlage longitudinal and circumferential growth in the fetal period and comparison with the growth plate cartilage of the postnatal age. *Microsc. Res. Tech.* **82**, 190–198 (2019).
5. Pazzaglia, U. E. *et al.* Study of Endochondral Ossification in Human Fetalcartilage Anlagen of Metacarpals: Comparative Morphology of Mineral Deposition in Cartilage and in the Periosteal Bone Matrix. *Anat. Rec.* **301**, 571–580 (2018).
6. Wilkie, A. O. M., Johnson, D. & Wall, S. A. Clinical genetics of craniosynostosis. *Curr. Opin. Pediatr.* **29**, 622 (2017).
7. Wilkie, A. O. M. Craniosynostosis: Genes and Mechanisms. *Hum. Mol. Genet.* **6**, 1647–1656 (1997).
8. Twigg, S. R. & Wilkie, A. O. A Genetic-Pathophysiological Framework for Craniosynostosis. *Am. J. Hum. Genet.* **97**, (2015).
9. Rice, S. J. *et al.* Genetic risk of osteoarthritis operates during human skeletogenesis. *Hum. Mol. Genet.* **32**, (2023).
10. Capellini, T. D. *et al.* Ancient selection for derived alleles at a GDF5 enhancer influencing human growth and osteoarthritis risk. *Nat. Genet.* **49**, (2017).
11. Zhang, B. *et al.* A human embryonic limb cell atlas resolved in space and time. *bioRxiv* 2022.04.27.489800 (2023) doi:10.1101/2022.04.27.489800.
12. Jardine, L. *et al.* Blood and immune development in human fetal bone marrow and Down syndrome. *Nature* **598**, 327–331 (2021).
13. He, P. *et al.* The changing mouse embryo transcriptome at whole tissue and single-cell resolution. *Nature* **583**, 760–767 (2020).
14. Pdgfra regulates multipotent cell differentiation towards chondrocytes via inhibiting Wnt9a/beta-catenin pathway during chondrocranial cartilage development. *Dev. Biol.* **466**, 36–46 (2020).
15. Xi, H. *et al.* A Human Skeletal Muscle Atlas Identifies the Trajectories of Stem and Progenitor Cells across Development and from Human Pluripotent Stem Cells. *Cell*

1246 *Stem Cell* **27**, 181–185 (2020).
1247 16. Hita-Contreras, F. *et al.* Development of the human shoulder joint during the embryonic
1248 and early fetal stages: anatomical considerations for clinical practice. *J. Anat.* **232**, 422
1249 (2018).
1250 17. Feng, C. *et al.* Lgr5 and Col22a1 Mark Progenitor Cells in the Lineage toward Juvenile
1251 Articular Chondrocytes. *Stem cell reports* **13**, (2019).
1252 18. Shwartz, Y., Viukov, S., Krief, S. & Zelzer, E. Joint Development Involves a Continuous
1253 Influx of Gdf5-Positive Cells. *Cell Rep.* **15**, (2016).
1254 19. Bian, Q. *et al.* A single cell transcriptional atlas of early synovial joint development.
1255 *Development* **147**, (2020).
1256 20. Lee, W. *et al.* Synergy between Piezo1 and Piezo2 channels confers high-strain
1257 mechanosensitivity to articular cartilage. *Proc. Natl. Acad. Sci. U. S. A.* **111**, (2014).
1258 21. Beccari, L. *et al.* Dbx2 regulation in limbs suggests interTAD sharing of enhancers. *Dev.*
1259 *Dyn.* **250**, 1280 (2021).
1260 22. Chen, H. *et al.* Runx2 Regulates Endochondral Ossification through Control of
1261 Chondrocyte Proliferation and Differentiation. *J. Bone Miner. Res.* **29**, 2653 (2014).
1262 23. Runx2 is required for hypertrophic chondrocyte mediated degradation of cartilage matrix
1263 during endochondral ossification. *Matrix Biology Plus* **12**, 100088 (2021).
1264 24. Ono, K. *et al.* Dmrt2 promotes transition of endochondral bone formation by linking Sox9
1265 and Runx2. *Communications Biology* **4**, 1–13 (2021).
1266 25. Snelling, S. J., Hulley, P. A. & Loughlin, J. BMP5 activates multiple signaling pathways
1267 and promotes chondrogenic differentiation in the ATDC5 growth plate model. *Growth*
1268 *Factors* **28**, (2010).
1269 26. Craft, A. M. *et al.* Specification of chondrocytes and cartilage tissues from embryonic
1270 stem cells. *Development* **140**, 2597–2610 (2013).
1271 27. Pothiawala, A. *et al.* GDF5+ chondroprogenitors derived from human pluripotent stem
1272 cells preferentially form permanent chondrocytes. *Development* **149**, dev196220 (2022).
1273 28. Sugimoto, Y. *et al.* Scx+/Sox9+ progenitors contribute to the establishment of the
1274 junction between cartilage and tendon/ligament. *Development* **140**, 2280–2288 (2013).
1275 29. Arostegui, M., Scott, R. W., Böse, K. & Underhill, T. M. Cellular taxonomy of Hic1+
1276 mesenchymal progenitor derivatives in the limb: from embryo to adult. *Nat. Commun.*
1277 **13**, 1–20 (2022).
1278 30. Buechler, M. B. *et al.* Cross-tissue organization of the fibroblast lineage. *Nature* **593**,
1279 575–579 (2021).
1280 31. Palumbo-Zerr, K. *et al.* Composition of TWIST1 dimers regulates fibroblast activation
1281 and tissue fibrosis. *Ann. Rheum. Dis.* **76**, 244–251 (2017).
1282 32. Twist2-driven chromatin remodeling governs the postnatal maturation of dermal

1283 fibroblasts. *Cell Rep.* **39**, 110821 (2022).
1284 33. Mascharak, S. *et al.* Preventing Engrailed-1 activation in fibroblasts yields wound
1285 regeneration without scarring. *Science* **372**, (2021).
1286 34. Sontake, V. *et al.* Wilms' tumor 1 drives fibroproliferation and myofibroblast
1287 transformation in severe fibrotic lung disease. *JCI Insight* **3**, (2018).
1288 35. Shi, Y. *et al.* Transcription Factor SOX5 Promotes the Migration and Invasion of
1289 Fibroblast-Like Synoviocytes in Part by Regulating MMP-9 Expression in Collagen-
1290 Induced Arthritis. *Front. Immunol.* **9**, (2018).
1291 36. Foxc1 promotes the proliferation of fibroblast-like synoviocytes in rheumatoid arthritis
1292 via PI3K/AKT signalling pathway. *Tissue and Cell* **53**, 15–22 (2018).
1293 37. Wang, J. *et al.* Forkhead box C1 promotes the pathology of osteoarthritis by
1294 upregulating β-catenin in synovial fibroblasts. *FEBS J.* **287**, (2020).
1295 38. Bobowski-Gerard, M. *et al.* Functional genomics uncovers the transcription factor BNC2
1296 as required for myofibroblastic activation in fibrosis. *Nat. Commun.* **13**, 1–20 (2022).
1297 39. Morriss-Kay, G. M. & Wilkie, A. O. M. Growth of the normal skull vault and its alteration
1298 in craniosynostosis: insights from human genetics and experimental studies. *J. Anat.*
1299 **207**, 637–653 (2005).
1300 40. Zhao, H. *et al.* The suture provides a niche for mesenchymal stem cells of craniofacial
1301 bones. *Nat. Cell Biol.* **17**, 386–396 (2015).
1302 41. Zhao, Q., Behringer, R. R. & de Crombrugghe, B. Prenatal folic acid treatment
1303 suppresses acrania and meroanencephaly in mice mutant for the Cart1 homeobox
1304 gene. *Nat. Genet.* **13**, 275–283 (1996).
1305 42. Pini, J. *et al.* ALX1-related frontonasal dysplasia results from defective neural crest cell
1306 development and migration. *EMBO Mol. Med.* **12**, (2020).
1307 43. Eze, U. C., Bhaduri, A., Haeussler, M., Nowakowski, T. J. & Kriegstein, A. R. Single-cell
1308 atlas of early human brain development highlights heterogeneity of human
1309 neuroepithelial cells and early radial glia. *Nat. Neurosci.* **24**, 584–594 (2021).
1310 44. Rocha, M. *et al.* From head to tail: regionalization of the neural crest. *Development* **147**,
1311 dev193888 (2020).
1312 45. Farmer, D. T. *et al.* The developing mouse coronal suture at single-cell resolution. *Nat.*
1313 *Commun.* **12**, (2021).
1314 46. Debnath, S. *et al.* Discovery of a periosteal stem cell mediating intramembranous bone
1315 formation. *Nature* **562**, 133–139 (2018).
1316 47. Cesario, J. M. *et al.* Anti-osteogenic function of a LIM-homeodomain transcription factor
1317 LMX1B is essential to early patterning of the calvaria. *Dev. Biol.* **443**, 103 (2018).
1318 48. Zhang, Q. *et al.* RUNX2 co-operates with EGR1 to regulate osteogenic differentiation
1319 through Htra1 enhancers. *J. Cell. Physiol.* **235**, (2020).

1320 49. Zaman, G. *et al.* Loading-related Regulation of Transcription Factor EGR2/Krox-20 in
1321 Bone Cells Is ERK1/2 Protein-mediated and Prostaglandin, Wnt Signaling Pathway-,
1322 and Insulin-like Growth Factor-I Axis-dependent. *J. Biol. Chem.* **287**, 3946 (2012).
1323 50. Rached, M. T. *et al.* FoxO1 is a positive regulator of bone formation by favoring protein
1324 synthesis and resistance to oxidative stress in osteoblasts. *Cell Metab.* **11**, (2010).
1325 51. Zalc, A., Rattenbach, R., Auradé, F., Cadot, B. & Relaix, F. Pax3 and Pax7 play
1326 essential safeguard functions against environmental stress-induced birth defects. *Dev.*
1327 *Cell* **33**, (2015).
1328 52. Ono, N., Ono, W., Nagasawa, T. & Kronenberg, H. M. A subset of chondrogenic cells
1329 provides early mesenchymal progenitors in growing bones. *Nat. Cell Biol.* **16**, 1157–
1330 1167 (2014).
1331 53. Peters, H., Neubüser, A., Kratochwil, K. & Balling, R. Pax9-deficient mice lack
1332 pharyngeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities.
1333 *Genes Dev.* **12**, (1998).
1334 54. Holmes, G. *et al.* Single-cell analysis identifies a key role for Hhip in murine coronal
1335 suture development. *Nat. Commun.* **12**, 1–16 (2021).
1336 55. Orestes-Cardoso, S. M. *et al.* Postnatal Msx1 expression pattern in craniofacial, axial,
1337 and appendicular skeleton of transgenic mice from the first week until the second year.
1338 *Dev. Dyn.* **221**, 1–13 (2001).
1339 56. Zanotti, S. & Canalis, E. Activation of Nfatc2 in osteoblasts causes osteopenia. *J. Cell.*
1340 *Physiol.* **230**, (2015).
1341 57. Ranger, A. M. *et al.* The Nuclear Factor of Activated T Cells (Nfat) Transcription Factor
1342 Nfatp (Nfatc2) Is a Repressor of Chondrogenesis. *J. Exp. Med.* **191**, 9–22 (2000).
1343 58. Loss of Iroquois homeobox transcription factors 3 and 5 in osteoblasts disrupts cranial
1344 mineralization. *Bone Reports* **5**, 86–95 (2016).
1345 59. Maes, C. *et al.* Osteoblast Precursors, but Not Mature Osteoblasts, Move into
1346 Developing and Fractured Bones along with Invading Blood Vessels. *Dev. Cell* **19**, 329–
1347 344 (2010).
1348 60. Percival, C. J. & Richtsmeier, J. T. Angiogenesis and intramembranous osteogenesis.
1349 *Dev. Dyn.* **242**, 909–922 (2013).
1350 61. Zhu, K. *et al.* ATF4 promotes bone angiogenesis by increasing VEGF expression and
1351 release in the bone environment. *J. Bone Miner. Res.* **28**, 1870 (2013).
1352 62. Kronenberg, H. M. Developmental regulation of the growth plate. *Nature* **423**, 332–336
1353 (2003).
1354 63. Kusumbe, A. P., Ramasamy, S. K. & Adams, R. H. Coupling of angiogenesis and
1355 osteogenesis by a specific vessel subtype in bone. *Nature* **507**, 323–328 (2014).
1356 64. Wang, Y. *et al.* Ephrin-B2 controls VEGF-induced angiogenesis and

1357 lymphangiogenesis. *Nature* **465**, (2010).
1358 65. Wang, F. *et al.* RACK1 regulates VEGF/Flt1-mediated cell migration via activation of a
1359 PI3K/Akt pathway. *J. Biol. Chem.* **286**, (2011).
1360 66. Ruehle, M. A. *et al.* Mechanical Regulation of Microvascular Angiogenesis. *bioRxiv*
1361 2020.01.14.906354 (2020) doi:10.1101/2020.01.14.906354.
1362 67. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB:
1363 inferring cell-cell communication from combined expression of multi-subunit ligand-
1364 receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).
1365 68. Ramasamy, S. K., Kusumbe, A. P., Wang, L. & Adams, R. H. Endothelial Notch activity
1366 promotes angiogenesis and osteogenesis in bone. *Nature* **507**, 376–380 (2014).
1367 69. Miraoui, H. *et al.* Fibroblast growth factor receptor 2 promotes osteogenic differentiation
1368 in mesenchymal cells via ERK1/2 and protein kinase C signaling. *J. Biol. Chem.* **284**,
1369 (2009).
1370 70. Kevorkova, O. *et al.* Low-Bone-Mass Phenotype of Deficient Mice for the Cluster of
1371 Differentiation 36 (CD36). *PLoS One* **8**, e77701 (2013).
1372 71. Baron, R. & Kneissel, M. WNT signaling in bone homeostasis and disease: from human
1373 mutations to treatments. *Nat. Med.* **19**, 179–192 (2013).
1374 72. Yu, X., Huang, Y., Collin-Osdoby, P. & Osdoby, P. CCR1 Chemokines Promote the
1375 Chemotactic Recruitment, RANKL Development, and Motility of Osteoclasts and Are
1376 Induced by Inflammatory Cytokines in Osteoblasts. *J. Bone Miner. Res.* **19**, 2065–2077
1377 (2004).
1378 73. DPP-4 inhibitor impedes lipopolysaccharide-induced osteoclast formation and bone
1379 resorption in vivo. *Biomed. Pharmacother.* **109**, 242–253 (2019).
1380 74. Tomlinson, R. E. *et al.* NGF-TrkA Signaling by Sensory Nerves Coordinates the
1381 Vascularization and Ossification of Developing Endochondral Bone. *Cell Rep.* **16**,
1382 (2016).
1383 75. Hu, B. *et al.* Sensory nerves regulate mesenchymal stromal cell lineage commitment by
1384 tuning sympathetic tones. *J. Clin. Invest.* **130**, (2020).
1385 76. Tower, R. J. *et al.* Spatial transcriptomics reveals a role for sensory nerves in preserving
1386 cranial suture patency through modulation of BMP/TGF-β signaling. *Proc. Natl. Acad.*
1387 *Sci. U. S. A.* **118**, (2021).
1388 77. Shum, L. & Nuckolls, G. The life cycle of chondrocytes in the developing skeleton.
1389 *Arthritis Res.* **4**, 94–106 (2002).
1390 78. Taïhi, I., Nassif, A., Isaac, J., Fournier, B. P. & Ferré, F. Head to Knee: Cranial Neural
1391 Crest-Derived Cells as Promising Candidates for Human Cartilage Repair. *Stem Cells*
1392 *Int.* **2019**, 9310318 (2019).
1393 79. Chilton, J. K. & Guthrie, S. Cranial expression of class 3 secreted semaphorins and their

1394 neuropilin receptors. *Dev. Dyn.* **228**, 726–733 (2003).
1395 80. Berndt, J. D. & Halloran, M. C. Semaphorin 3d promotes cell proliferation and neural
1396 crest cell development downstream of TCF in the zebrafish hindbrain. *Development*
1397 **133**, 3983–3992 (2006).
1398 81. DeTomaso, D. & Yosef, N. Hotspot identifies informative gene modules across
1399 modalities of single-cell genomics. *Cell Syst.* **12**, 446–456.e9 (2021).
1400 82. Wang, Z., Liu, B., Lin, K., Duan, C. & Wang, C. The presence and degradation of nerve
1401 fibers in articular cartilage of neonatal rats. *J. Orthop. Surg. Res.* **17**, 331 (2022).
1402 83. Yueh, Y. G., Gardner, D. P. & Kappen, C. Evidence for regulation of cartilage
1403 differentiation by the homeobox gene Hoxc-8. *Proc. Natl. Acad. Sci. U. S. A.* **95**, (1998).
1404 84. Muyas, F. *et al.* De novo detection of somatic mutations in high-throughput single-cell
1405 profiling data sets. *Nat. Biotechnol.* (2023) doi:10.1038/s41587-023-01863-z.
1406 85. Cartilage repair using bone morphogenetic protein 4 and muscle-derived stem cells.
1407 http://dx.doi.org/10.1002/art.21632 doi:10.1002/art.21632.
1408 86. Yin, Z. *et al.* Atlas of Musculoskeletal Stem Cells with the Soft and Hard Tissue
1409 Differentiation Architecture. *Adv. Sci. Lett.* **7**, (2020).
1410 87. Grimaldi, A., Comai, G., Mella, S. & Tajbakhsh, S. Identification of bipotent progenitors
1411 that give rise to myogenic and connective tissues in mouse. *Elife* **11**, (2022).
1412 88. He, J. *et al.* Dissecting human embryonic skeletal stem cell ontogeny by single-cell
1413 transcriptomic and functional analyses. *Cell Res.* **31**, 742–757 (2021).
1414 89. Xie, M. *et al.* Schwann cell precursors contribute to skeletal formation during embryonic
1415 development in mice and zebrafish. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 15068–15073
1416 (2019).
1417 90. Richard, D. *et al.* Evolutionary Selection and Constraint on Human Knee Chondrocyte
1418 Regulation Impacts Osteoarthritis Risk. *Cell* **181**, 362–381.e28 (2020).
1419 91. Wilkinson, J. M. & Zeggini, E. The Genetic Epidemiology of Joint Shape and the
1420 Development of Osteoarthritis. *Calcif. Tissue Int.* **109**, 257–276 (2021).
1421 92. Cowen, L., Ideker, T., Raphael, B. J. & Sharan, R. Network propagation: a universal
1422 amplifier of genetic associations. *Nat. Rev. Genet.* **18**, 551–562 (2017).
1423 93. Heinonen, J. *et al.* Defects in chondrocyte maturation and secondary ossification in
1424 mouse knee joint epiphyses due to Snorc deficiency. *Osteoarthritis Cartilage* **25**, 1132–
1425 1142 (2017).
1426 94. Lee, H.-S. *et al.* Activation of Integrin-RACK1/PKCalpha signalling in human articular
1427 chondrocyte mechanotransduction. *Osteoarthritis Cartilage* **10**, 890–897 (2002).
1428 95. Zhang, F. *et al.* NFATc1 marks articular cartilage progenitors and negatively determines
1429 articular chondrocyte differentiation. *Elife* **12**, (2023).
1430 96. Greenblatt, M. B. *et al.* NFATc1 and NFATc2 repress spontaneous osteoarthritis. *Proc.*

1431 *Natl. Acad. Sci. U. S. A.* **110**, 19914–19919 (2013).
1432 97. Ren, R. *et al.* The role of Ca2+/Calcineurin/NFAT signalling pathway in
1433 osteoblastogenesis. *Cell Prolif.* **54**, e13122 (2021).
1434 98. Genetics of developmental dysplasia of the hip. *Eur. J. Med. Genet.* **63**, 103990 (2020).
1435 99. Zhang, X. *et al.* Lipid peroxidation in osteoarthritis: focusing on 4-hydroxynonenal,
1436 malondialdehyde, and ferroptosis. *Cell Death Discovery* **9**, 1–13 (2023).
1437 100. Mathijssen, I. M. *et al.* Tracing craniosynostosis to its developmental stage through
1438 bone center displacement. *J. Craniofac. Genet. Dev. Biol.* **19**, (1999).
1439 101. Lajeunie, E., Le Merrer, M., Bonaïti-Pellie, C., Marchac, D. & Renier, D. Genetic study of
1440 nonsyndromic coronal craniosynostosis. *Am. J. Med. Genet.* **55**, (1995).
1441 102. Wilkie, A. O. *et al.* Prevalence and complications of single-gene and chromosomal
1442 disorders in craniosynostosis. *Pediatrics* **126**, (2010).
1443 103. Sharma, V. P. *et al.* Mutations in TCF12, encoding a basic helix-loop-helix partner of
1444 TWIST1, are a frequent cause of coronal craniosynostosis. *Nat. Genet.* **45**, 304–307
1445 (2013).
1446 104. Yoon, W. J. *et al.* The Boston-type craniosynostosis mutation MSX2 (P148H) results in
1447 enhanced susceptibility of MSX2 to ubiquitin-dependent degradation. *J. Biol. Chem.*
1448 **283**, (2008).
1449 105. James, A. W., Levi, B., Xu, Y., Carre, A. L. & Longaker, M. T. Retinoic acid enhances
1450 osteogenesis in cranial suture-derived mesenchymal cells: potential mechanisms of
1451 retinoid-induced craniosynostosis. *Plast. Reconstr. Surg.* **125**, 1352–1361 (2010).
1452 106. Clouthier, D. E. *et al.* Cranial and cardiac neural crest defects in endothelin-A receptor-
1453 deficient mice. *Development* **125**, 813–824 (1998).
1454 107. Spence, S., Anderson, C., Cukierski, M. & Patrick, D. Teratogenic effects of the
1455 endothelin receptor antagonist L-753,037 in the rat. *Reprod. Toxicol.* **13**, 15–29 (1999).
1456 108. Bruderer, M., Richards, R. G., Alini, M. & Stoddart, M. J. Role and regulation of RUNX2
1457 in osteogenesis. *Eur. Cell. Mater.* **28**, 269–286 (2014).
1458 109. van Lierop, A. H. *et al.* Van Buchem disease: clinical, biochemical, and densitometric
1459 features of patients and disease carriers. *J. Bone Miner. Res.* **28**, 848–854 (2013).
1460 110. Collette, N. M. *et al.* Targeted deletion of Sost distal enhancer increases bone formation
1461 and bone mass. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 14092–14097 (2012).
1462 111. Liu, N. *et al.* DNA binding-dependent and -independent functions of the Hand2
1463 transcription factor during mouse embryogenesis. *Development* **136**, 933 (2009).
1464 112. Xiong, W. *et al.* Hand2 is required in the epithelium for palatogenesis in mice. *Dev. Biol.*
1465 **330**, 131 (2009).
1466 113. Zhou, C. *et al.* Lhx8 Mediated Wnt and TGFβ Pathways in Tooth Development and
1467 Regeneration. *Biomaterials* **63**, 35 (2015).

114. Pijuan-Sala, B. *et al.* A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature* **566**, 490–495 (2019).
115. Kameneva, P. *et al.* Single-cell transcriptomics of human embryos identifies multiple sympathoblast lineages with potential implications for neuroblastoma origin. *Nat. Genet.* **53**, 694–706 (2021).
116. Crow, M., Paul, A., Ballouz, S., Huang, Z. J. & Gillis, J. Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. *Nat. Commun.* **9**, 1–12 (2018).
117. Lee, H., Marco Salas, S., Gyllborg, D. & Nilsson, M. Direct RNA targeted in situ sequencing for transcriptomic profiling in tissue. *Sci. Rep.* **12**, 1–9 (2022).
118. Li, T. *et al.* WebAtlas pipeline for integrated single cell and spatial transcriptomic data. *bioRxiv* 2023.05.19.541329 (2023) doi:10.1101/2023.05.19.541329.
119. GitHub - VasylVaskivskyi/microaligner: Image registration (alignment) software for large microscopy images. *GitHub* https://github.com/VasylVaskivskyi/microaligner.
120. Pachitariu, M. & Stringer, C. Cellpose 2.0: how to train your own model. *Nat. Methods* **19**, 1634–1641 (2022).
121. Gataric, M. *et al.* PoSTcode: Probabilistic image-based spatial transcriptomics decoder. *bioRxiv* (2021) doi:10.1101/2021.10.12.464086.
122. Virshup, I., Rybakov, S., Theis, F. J., Angerer, P. & Wolf, F. A. anndata: Annotated data. *bioRxiv* (2021) doi:10.1101/2021.12.16.473007.
123. Litviňuková, M. *et al.* Cells of the adult human heart. *Nature* **588**, (2020).
124. Heaton, H. *et al.* Souporcell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. *Nat. Methods* **17**, 615–620 (2020).
125. Young, M. D. & Behjati, S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *Gigascience* **9**, giaa151 (2020).
126. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell systems* **8**, (2019).
127. Granja, J. M. *et al.* ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat. Genet.* **53**, 403–411 (2021).
128. Bredikhin, D., Kats, I. & Stegle, O. MUON: multimodal omics analysis framework. *Genome Biol.* **23**, 1–12 (2022).
129. Polański, K. *et al.* BBKNN: fast batch alignment of single cell transcriptomes. *Bioinformatics* **36**, (2020).
130. Ashuach, T. *et al.* MultiVI: deep generative model for the integration of multimodal data. *Nat. Methods* **20**, 1222–1231 (2023).
131. Popescu, D.-M. *et al.* Decoding human fetal liver haematopoiesis. *Nature* **574**, 365–371 (2019).

132. Xu, C. *et al.* Automatic cell type harmonization and integration across Human Cell Atlas datasets. *bioRxiv* 2023.05.01.538994 (2023) doi:10.1101/2023.05.01.538994.
133. Domínguez Conde, C. *et al.* Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science* **376**, eabl5197 (2022).
134. Bravo González-Blas, C. *et al.* SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. *Nat. Methods* 1–13 (2023).
135. Moerman, T. *et al.* GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics* **35**, 2159–2161 (2019).
136. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
137. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* vol. 566 496–502 Preprint at https://doi.org/10.1038/s41586-019-0969-x (2019).
138. Moran, P. A. P. Notes on continuous stochastic phenomena. *Biometrika* **37**, 17–23 (1950).
139. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* vol. 566 496–502 Preprint at https://doi.org/10.1038/s41586-019-0969-x (2019).
140. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* vol. 32 2847–2849 Preprint at https://doi.org/10.1093/bioinformatics/btw313 (2016).
141. Hahsler, M., Hornik, K. & Buchta, C. Getting Things in Order: An Introduction to theRPackageseriation. *Journal of Statistical Software* vol. 25 Preprint at https://doi.org/10.18637/jss.v025.i03 (2008).
142. La Manno, G. *et al.* RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
143. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020).
144. Kamimoto, K. *et al.* Dissecting cell identity via network inference and in silico gene perturbation. *Nature* **614**, 742–751 (2023).
145. Faure, L., Soldatov, R., Kharchenko, P. V. & Adameyko, I. scFates: a scalable python package for advanced pseudotime and bifurcation analysis from single-cell data. *Bioinformatics* **39**, (2023).
146. Haghverdi, L., Büttner, M., Wolf, F. A., Buettner, F. & Theis, F. J. Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods* **13**, 845–848 (2016).
147. Pickrell, J. K. Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. *Am. J. Hum. Genet.* **94**, 559–573 (2014).
148. Elmentaite, R. *et al.* Cells of the human intestinal tract mapped across space and time. *Nature* **597**, 250–255 (2021).

1542 149. Boer, C. G. *et al.* Deciphering osteoarthritis genetics across 826,690 individuals from 9
1543 populations. *Cell* **184**, 6003–6005 (2021).
1544 150. Paull, E. O. *et al.* Discovering causal pathways linking genomic events to transcriptional
1545 states using Tied Diffusion Through Interacting Events (TieDIE). *Bioinformatics* **29**,
1546 2757–2764 (2013).

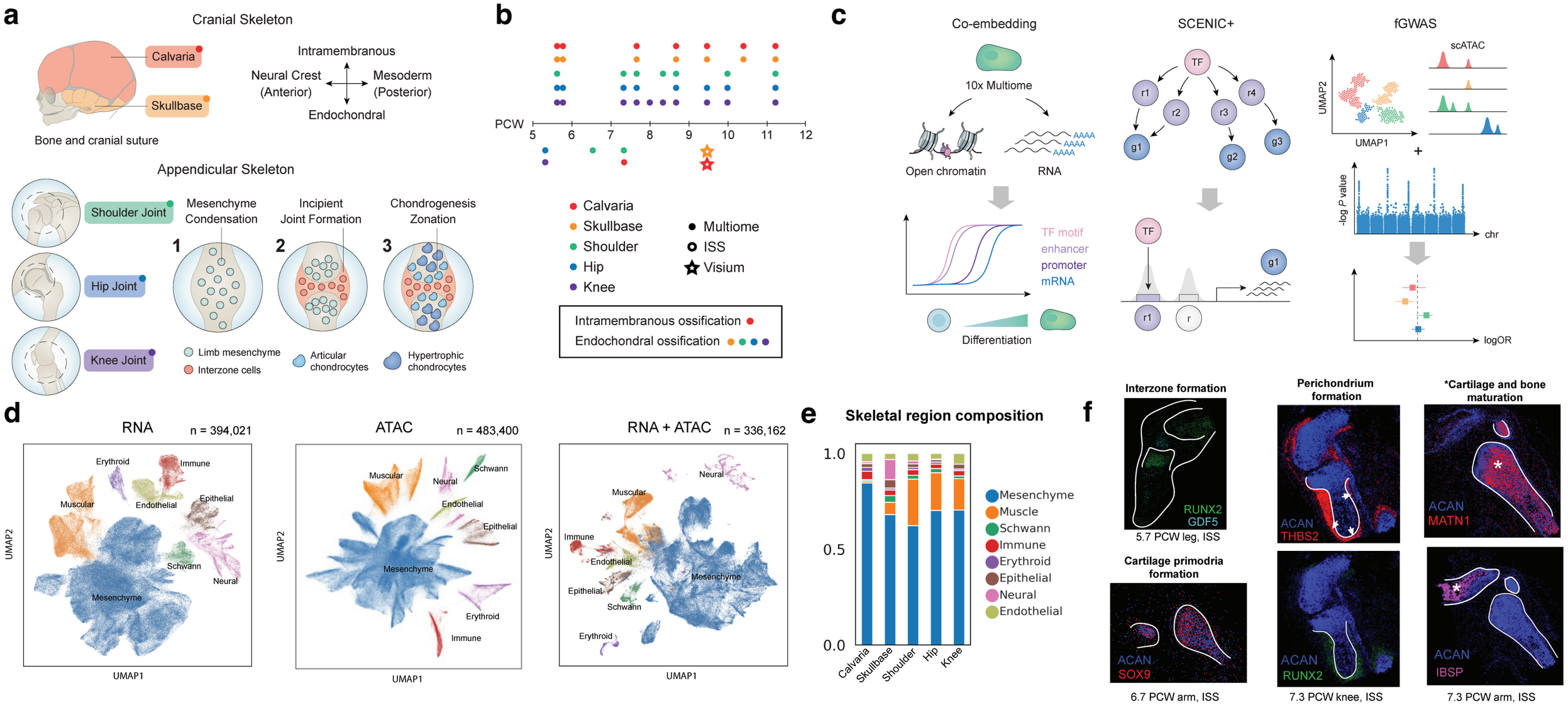

a
Cranial Skeleton
Calvaria
Skullbase
Bone and cranial suture
Intramembranous
Neural Crest (Anterior)
Mesoderm (Posterior)
Endochondral
Appendicular Skeleton
Shoulder Joint
Hip Joint
Knee Joint
Mesenchyme Condensation
Incipient Joint Formation
Chondrogenesis Zonation
1
2
3
Limb mesenchyme
Interzone cells
Articular chondrocytes
Hypertrophic chondrocytes
b
PCW
5
6
7
8
9
10
11
12
Calvaria
Skullbase
Shoulder
Hip
Knee
Multiome
ISS
Visium
Intramembranous ossification
Endochondral ossification
c
Co-embedding
10x Multiome
Open chromatin
RNA
AAAA
TF motif
enhancer
promoter
mRNA
Differentiation
SCENIC+
TF
r1
r2
r3
r4
g1
g2
g3
r
fGWAS
scATAC
UMAP2
UMAP1
-log P value
chr
logOR
d
RNA
n = 394,021
ATAC
n = 483,400
RNA + ATAC
n = 336,162
Erythroid
Immune
Muscular
Endothelial
Epithelial
Schwann
Neural
Mesenchyme
UMAP2
UMAP1
e
Skeletal region composition
1.0
0.5
0.0
Calvaria
Skullbase
Shoulder
Hip
Knee
Mesenchyme
Muscle
Schwann
Immune
Erythroid
Epithelial
Neural
Endothelial
f
Interzone formation
RUNX2
GDF5
5.7 PCW leg, ISS
Perichondrium formation
ACAN
THBS2
*Cartilage and bone maturation
ACAN
MATN1
Cartilage primodria formation
ACAN
SOX9
6.7 PCW arm, ISS
ACAN
RUNX2
7.3 PCW knee, ISS
ACAN
IBSP
7.3 PCW arm, ISS

Figure. 1: Human fetal skeletal atlas overview. a) Anatomical sampling approach to dissect five main anatomical sites. Cell types of origin within the cranium are determined by the anterior-posterior axis. **b)** Donor overview across age and anatomical regions sampled; atlasing modalities represented in the legend. **c)** Analysis approaches applied to integrated RNA-ATAC seq data. **d)** UMAP embedding of dataset using RNA-seq only, ATAC-seq only and integrated datasets. **e)** Cellular compartment composition of the skeletal region across anatomical regions. **f)** Relative cell type abundance across anatomical locations.

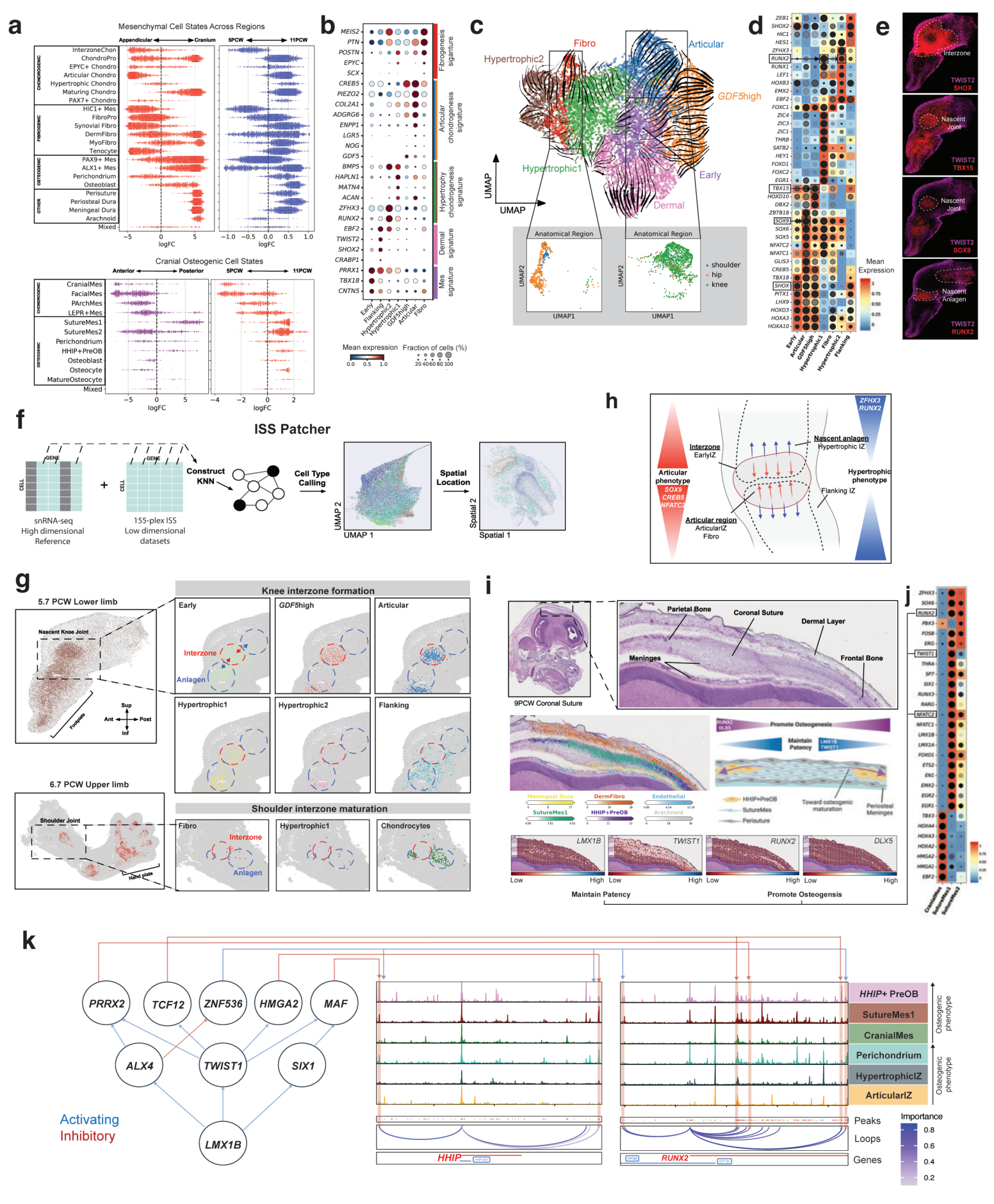
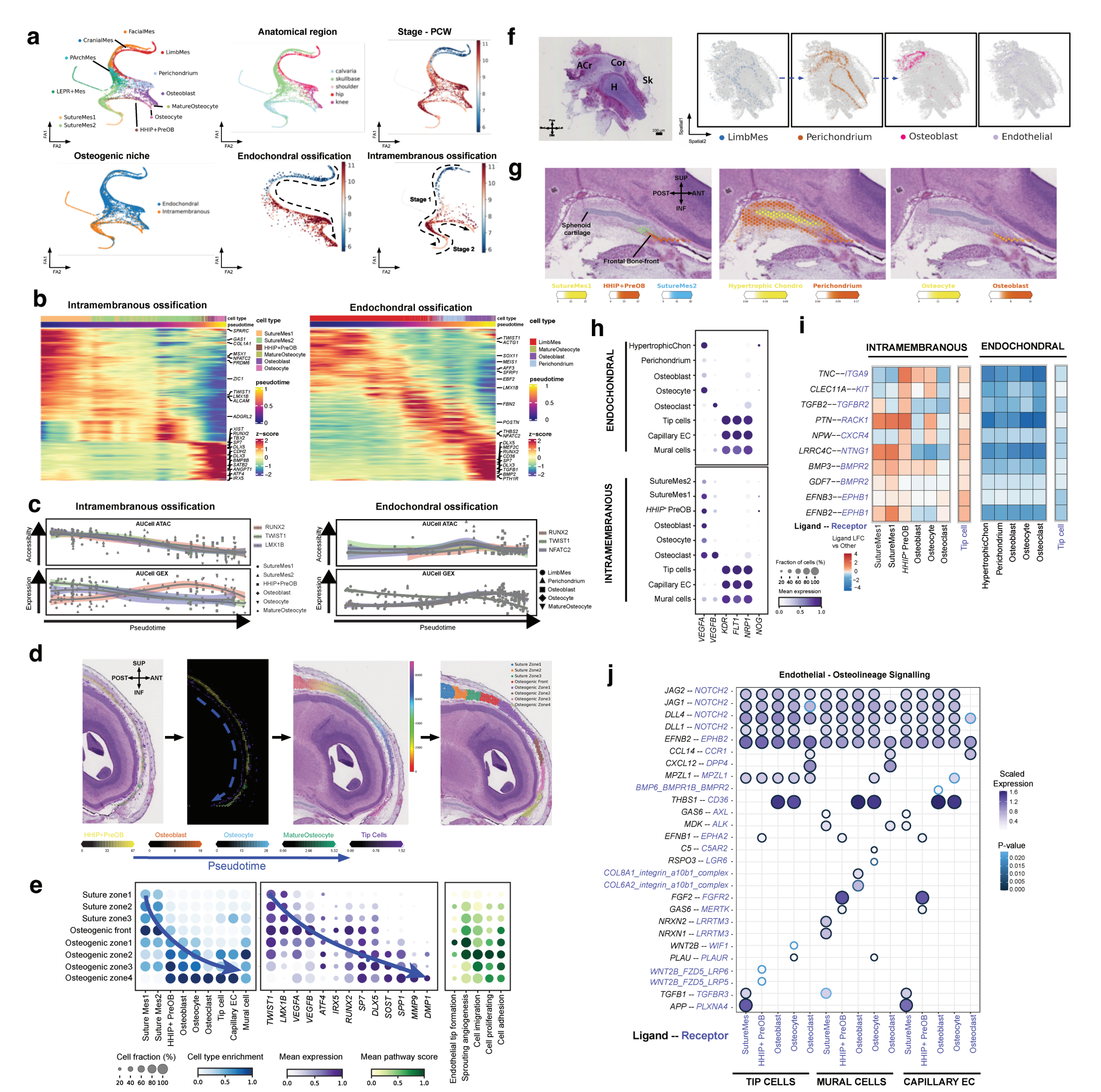

a
Mesenchymal Cell States Across Regions
Appendicular
Cranium
5PCW
11PCW
CHONDROGENIC
InterzoneChon
ChondroPro
EPYC+ Chondro
Articular Chondro
Hypertrophic Chondro
Maturing Chondro
PAX7+ Chondro
FIBROGENIC
HIC1+ Mes
FibroPro
Synovial Fibro
DermFibro
MyoFibro
Tenocyte
OSTEOGENIC
PAX9+ Mes
ALX1+ Mes
Perichondrium
Osteoblast
OTHER
Perisuture
Periosteal Dura
Meningeal Dura
Arachnoid
Mixed
logFC
Cranial Osteogenic Cell States
Anterior
Posterior
CHONDROGENIC
CranialMes
FacialMes
PArchMes
LEPR+Mes
OSTEOGENIC
SutureMes1
SutureMes2
Perichondrium
HHIP+PreOB
Osteoblast
Osteocyte
MatureOsteocyte
Mixed
b
MEIS2
PTN
POSTN
EPYC
SCX
CREB5
PIEZO2
COL2A1
ADGRG6
ENPP1
LGR5
NOG
GDF5
BMP5
HAPLN1
MATN4
ACAN
ZFHX3
RUNX2
EBF2
TWIST2
SHOX2
CRABP1
PRRX1
TBX18
CNTN5
Fibrogenesis signature
Articular chondrogenesis signature
Hypertrophy chondrogenesis signature
Dermal signature
Mes signature
Early
Flanking
Hypertrophic2
Hypertrophic1
GDF5high
Articular
Fibro
Mean expression
Fraction of cells (%)
c
Fibro
Articular
Hypertrophic2
GDF5high
Hypertrophic1
Early
Dermal
UMAP
Anatomical Region
shoulder
hip
knee
UMAP1
UMAP2
d
Mean Expression
e
Interzone
TWIST2
SHOX
Nascent Joint
TWIST2
TBX15
Nascent Joint
TWIST2
SOX9
Nascent Anlagen
TWIST2
RUNX2
f
ISS Patcher
GENE
CELL
snRNA-seq
High dimensional
Reference
155-plex ISS
Low dimensional
datasets
Construct KNN
Cell Type Calling
UMAP 1
UMAP 2
Spatial Location
Spatial 1
Spatial 2
h
ZFHX3
RUNX2
Interzone
EarlyIZ
Nascent anlagen
Hypertrophic IZ
Articular phenotype
SOX9
CREB5
NFATC2
Hypertrophic phenotype
Flanking IZ
Articular region
ArticularIZ
Fibro
g
5.7 PCW Lower limb
Nascent Knee Joint
Footplate
Sup
Ant
Post
Inf
Knee interzone formation
Early
GDF5high
Articular
Interzone
Anlagen
Hypertrophic1
Hypertrophic2
Flanking
6.7 PCW Upper limb
Shoulder Joint
Hind plate
Shoulder interzone maturation
Fibro
Interzone
Anlagen
Hypertrophic1
Chondrocytes
i
9PCW Coronal Suture
Parietal Bone
Coronal Suture
Dermal Layer
Meninges
Frontal Bone
Meningeal Dura
DermFibro
Endothelial
SutureMes1
HHIP+PreOB
Arachnoid
Promote Osteogenesis
Maintain Patency
HHIP+PreOB
SutureMes
Perisuture
Toward osteogenic maturation
Periosteal Meninges
LMX1B
TWIST1
RUNX2
DLX5
Low
High
Maintain Patency
Promote Osteogensis
j
k
PRRX2
TCF12
ZNF536
HMGA2
MAF
ALX4
TWIST1
SIX1
LMX1B
Activating
Inhibitory
HHIP+ PreOB
SutureMes1
CranialMes
Perichondrium
HypertrophicIZ
ArticularIZ
Osteogenic phenotype
Osteogenic phenotype
Peaks
Loops
Genes
Importance
HHIP
RUNX2

Figure. 2: Formation of the embryonic joints across space and time. a) Differential abundance (MILO) of cell states across lineages against multiple comparators. Top: Broad clusters of non-myogenic mesenchyme. Bottom: Cranial osteogenic subclusters. ChondroPro: Chondroprogenitors, *EPYC*$^{+}$ Chondro: *EPYC*$^{+}$ Chondrocyte, Articular Chondro: Chondrocyte, Hypertrophic Chondro: Hypertrophic Chondrocyte, Maturing Chondro: Maturing Chondrocyte, *PAX7*$^{+}$ Chondro: *PAX7*$^{+}$ Chondrocyte, *HIC1*$^{+}$ Mes: *HIC1*$^{+}$ Mesenchyme, FibroPro: Fibroblast progenitors, Synovial Fibro: Synovial fibroblast, DermFibro: Dermal fibroblast, MyoFibro: Myofibroblast, *PAX9*$^{+}$ Mes: *PAX9*$^{+}$ Mesenchyme, *ALX1*$^{+}$ Mes, *ALX1*$^{+}$ Mesenchyme, CranialMes: Cranial mesenchyme, FacialMes: Facial mesenchyme, PArchMes: Pharyngeal mesenchyme, *LEPR*$^{+}$ Mes: *LEPR*$^{+}$ Mesenchyme, SutureMes1: Suture mesenchyme1, SutureMes2: Suture mesenchyme2, *HHIP*$^{+}$ PreOB: *HHIP*$^{+}$ Pre-Osteoblast. Mixed: other related cell types that are not included in each plot respectively. **b)** Normalised gene expression of genes in modules associated with skeletal development within InterzoneChon subclusters. **c)** RNA-velocity on UMAP of RNA-sub-clustered cell states of the broad InterzoneChon cluster. **d)** SCENIC+ predicted TF expression (box color), dot size shows target gene accessibility (AUCell) and dot shade (grayscale) shows target gene expression (GEX AUCell). **e)** Demultiplexed marker gene expression from ISS of the 5.7 PCW lower limb. **f)** ISS-Patcher workflow (see methods). **g)** Spatial plots of ISS-Patcher cell cluster imputation. **h)** Schematic of knee joint interzone formation, gradients demonstrating genes associated with zonated hypertrophic and articular phenotypes. **i)** Histological view and annotations of adjacent sections of 10x cytassist visium data. Cell2location results of coronal suture and schematic demonstrating TF gradient across regions of the coronal suture. **j)** SCENIC+ predicted TF expression (box color), dot size shows target gene accessibility (AUCell) and dot shade (grayscale) shows target gene expression (GEX AUCell) in suture progenitors. **k)** Coverage plots showing aggregated scATAC signals around the *HHIP* and *RUNX2* loci for osteoprogenitor cell states with increasing osteogenic phenotype in IM (CranialMes to HHIP+ ProOB) and EC (ArticularIZ to Perichondrium). Below each coverage plot, loops predicted by Scenic+ between TSS and enhancers are shown (colored by importance score). Selected upstream TFs predicted by Scenic+ to bind and regulate via some of the enhancers are shown on the left. The network links inhibitors of osteogenesis (such as TWIST1, LMX1B) to pro-osteogenic genes (RUNX2, HHIP) via overall inhibitory connections.


a
Anatomical region
Stage - PCW
Osteogenic niche
Endochondral ossification
Intramembranous ossification
b
Intramembranous ossification
Endochondral ossification
c
Intramembranous ossification
Endochondral ossification
AUCell ATAC
AUCell GEX
Pseudotime
d
Pseudotime
e
f
LimbMes
Perichondrium
Osteoblast
Endothelial
g
SUP
POST
ANT
INF
Sphenoid cartilage
Frontal Bone-front
SutureMes1
HHIP+PreOB
SutureMes2
Hypertrophic Chondro
Perichondrium
Osteocyte
Osteoblast
h
ENDOCHONDRAL
INTRAMEMBRANOUS
i
INTRAMEMBRANOUS
ENDOCHONDRAL
Ligand -- Receptor
j
Endothelial - Osteolineage Signalling
Scaled Expression
P-value
Ligand -- Receptor
TIP CELLS
MURAL CELLS
CAPILLARY EC

1585 **Figure 3: Intramembranous and endochondral osteogenesis. a)** Force directed (FA)
1586 embedding showing trajectories per anatomical region that converge on osteoblasts and
1587 osteocytes. In intramembranous (IM) ossification the suture mesenchyme constitutes a
1588 distinct state that maintains progenitor pools. A two-stage IM process forms the sutures then
1589 subsequently bone-lineage. LimbMes: Limb mesenchyme. **b)** Heatmaps showing the
1590 expression of differentially expressed genes along pseudotime for IM and endochondral (EC)
1591 ossification. Differential expression has been tested using a spatial autocorrelation test in
1592 monocle3. **c)** TF accessibility (AUCell ATAC) and Target gene accessibility (AUCell GEX) of
1593 selected TFs along pseudotime. Differing patterns are observed for osteogenic (RUNX2) and
1594 inhibitory TFs (TWIST1, LMX1B, NFATC2), pointing to a complex regulatory mechanism. **d)**
1595 Spatial binning of Visium Cytassist image of frontal bone (sagittal section) using OrganAxis,
1596 axis begins from coronal suture to the ossifying frontal bone (posterior-anterior) Left-Right:
1597 Histological image of frontal bone; Cell-cluster enrichment from Cell2location; Axis values
1598 (rainbow); manual bins based on histological features **e)** Enrichment within spatial bins for IM
1599 cell states (left), selected marker gene expression (middle) and pathway enrichment (right). **f)**
1600 7.3 PCW humerus (Coronal section) and imputed cell clusters showing EC ossification as a
1601 sequence of mapped cell-states. ACr, Acromion; Cor, Coracoid; H, Humerus; Sk, Skin. **g)**
1602 Cross-section of Visium Cytassist at the meeting point between calvaria and skull base
1603 (anterior border of the sphenoid), chondrogenic cell states enrichment from Cell2location **h)**
1604 Predicted ligand-receptor interactions per IM and EC cell clusters. **i)** Gene expression along
1605 IM and EC cell clusters in a) as well as endothelial cells. **j)** Ligand-receptor interactions
1606 predicted between endothelial cells and cells of the osteogenic lineage.

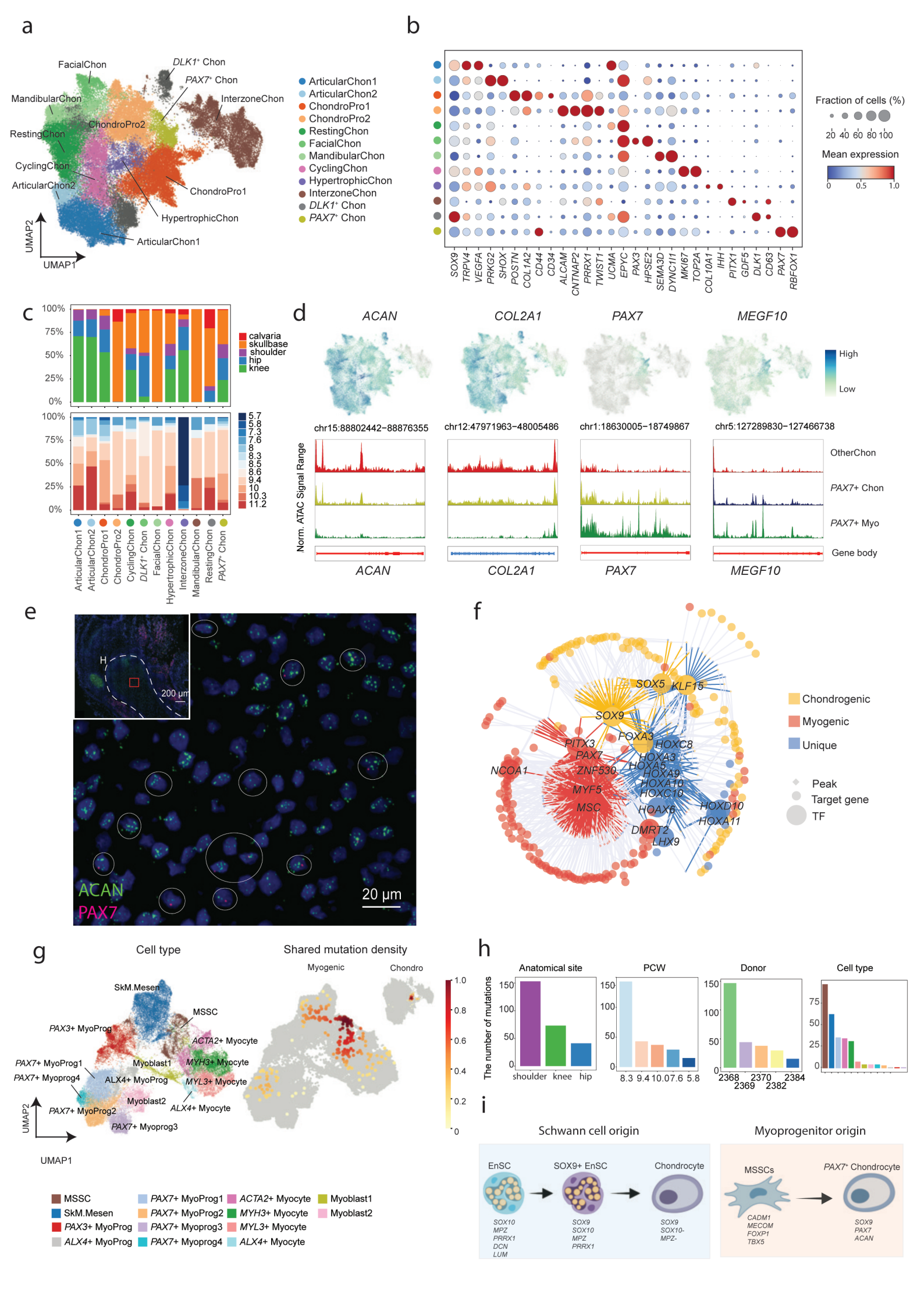

a
FacialChon
DLK1+ Chon
PAX7+ Chon
MandibularChon
InterzoneChon
ChondroPro2
RestingChon
CyclingChon
ArticularChon2
ChondroPro1
HypertrophicChon
ArticularChon1
UMAP2
UMAP1
ArticularChon1
ArticularChon2
ChondroPro1
ChondroPro2
RestingChon
FacialChon
MandibularChon
CyclingChon
HypertrophicChon
InterzoneChon
DLK1+ Chon
PAX7+ Chon
b
Fraction of cells (%)
20 40 60 80 100
Mean expression
0 0.5 1.0
SOX9 TRPV4 VEGFA PRKG2 SHOX POSTN COL1A2 CD44 CD34 ALCAM CNTNAP2 PRRX1 TWIST1 UCMA EPYC PAX3 HPSE2 SEMA3D DYNC1I1 MKI67 TOP2A COL10A1 IHH PITX1 GDF5 DLK1 CD63 PAX7 RBFOX1
c
100% 75% 50% 25% 0%
calvaria
skullbase
shoulder
hip
knee
5.7 5.8 7.3 7.6 8 8.3 8.5 8.6 9.4 10 10.3 11.2
ArticularChon1 ArticularChon2 ChondroPro1 ChondroPro2 CyclingChon DLK1+ Chon FacialChon HypertrophicChon InterzoneChon MandibularChon RestingChon PAX7+ Chon
d
ACAN
COL2A1
PAX7
MEGF10
High
Low
chr15:88802442–88876355
chr12:47971963–48005486
chr1:18630005–18749867
chr5:127289830–127466738
Norm. ATAC Signal Range
OtherChon
PAX7+ Chon
PAX7+ Myo
Gene body
ACAN
COL2A1
PAX7
MEGF10
e
H
200 µm
ACAN
PAX7
20 µm
f
SOX5
KLF15
SOX9
FOXA3
HOXC8
PITX3
PAX7
HOXA3
NCOA1
ZNF530
HOXA5
HOXA9
HOXA10
MYF5
HOXC10
MSC
HOXA6
HOXD10
HOXA11
DMRT2
LHX9
Chondrogenic
Myogenic
Unique
Peak
Target gene
TF
g
Cell type
Shared mutation density
Myogenic
Chondro
SkM.Mesen
MSSC
PAX3+ MyoProg
ACTA2+ Myocyte
PAX7+ MyoProg1
Myoblast1
MYH3+ Myocyte
PAX7+ Myoprog4
ALX4+ MyoProg
MYL3+ Myocyte
Myoblast2
PAX7+ MyoProg2
ALX4+ Myocyte
PAX7+ Myoprog3
UMAP2
UMAP1
1.0 0.8 0.6 0.4 0.2 0
MSSC
SkM.Mesen
PAX3+ MyoProg
ALX4+ MyoProg
PAX7+ MyoProg1
PAX7+ MyoProg2
PAX7+ Myoprog3
PAX7+ Myoprog4
ACTA2+ Myocyte
MYH3+ Myocyte
MYL3+ Myocyte
ALX4+ Myocyte
Myoblast1
Myoblast2
h
Anatomical site
PCW
Donor
Cell type
The number of mutations
150 100 50 0
shoulder knee hip
8.3 9.4 10.0 7.6 5.8
2368 2369 2370 2382 2384
75 50 25 0
i
Schwann cell origin
Myoprogenitor origin
EnSC
SOX9+ EnSC
Chondrocyte
SOX10
MPZ
PRRX1
DCN
LUM
SOX9
SOX10
MPZ
PRRX1
SOX9
SOX10-
MPZ-
MSSCs
PAX7+ Chondrocyte
CADM1
MECOM
FOXP1
TBX5
SOX9
PAX7
ACAN

Figure. 4: Chondrogenesis across different anatomical regions. **a**) RNA UMAP embeddings of chondrocytes coloured by subtypes. **b**) Dot plot showing the mean expression (dot colour) and fraction of expressing cells (dot size) of selected marker genes. **c**) Distribution of different chondrogenic sub types across anatomical regions. a, b, c share the same legends. **d**) RNA UMAPs (the upper panel) of representative chondrogenic genes (*ACAN*, *COL2A1*) and myogenic (*PAX7*, *MEGF10*) for *PAX7*$^+$ chondrocytes. Genome browser tracks (the lower panel) from scATAC-seq analyses displaying the sum signal within the indicated gene loci. OtherChon represents the whole chondro-lineages except *PAX7*$^+$ Chon. **e**) RNA scope for cells that co-expressed myogenic markers (*PAX7*, red) and chondrogenic markers (ACAN, green) in the shoulder at 7.3PCW. The blue marks the cell nucleus. Scale bar, 20/200 μm. **f**) Enhancer-GRN showing enriched TFs in *PAX7*$^+$ Chon, inferred using SCENIC+. Circles and diamonds represent genes and regions with TF-binding sites, respectively. Colour indicates whether the TF of *PAX7*$^+$ Chon is also differentially expressed in chondro-lineages or *PAX7*$^+$ myocytes. TF-region links are coloured by TF. **g**) The left panel shows the UMAP of myogenic lineages, while the right panel shows the gaussian kernel density estimates of shared somatic mutations over myogenic lineages and chondrocyte lineages. Myoprog: Myogenic progenitor. **h**) Barplots showing the number of shared mutations in myogenic cells per anatomical site/pcw/donor/celltype. **i**)**.** Schematic representation of the putative non-classical lineages that contribute to chondrogenesis.

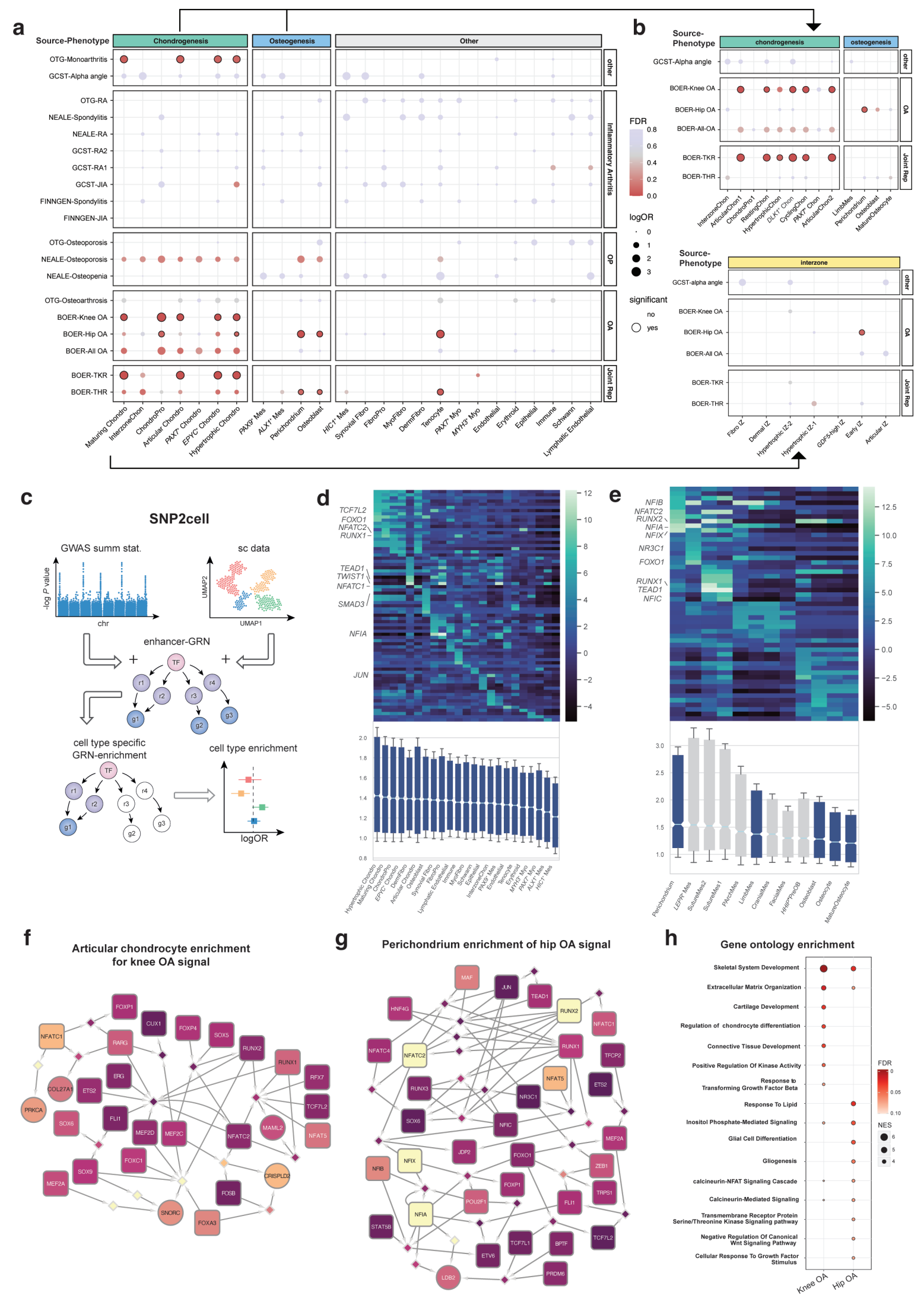
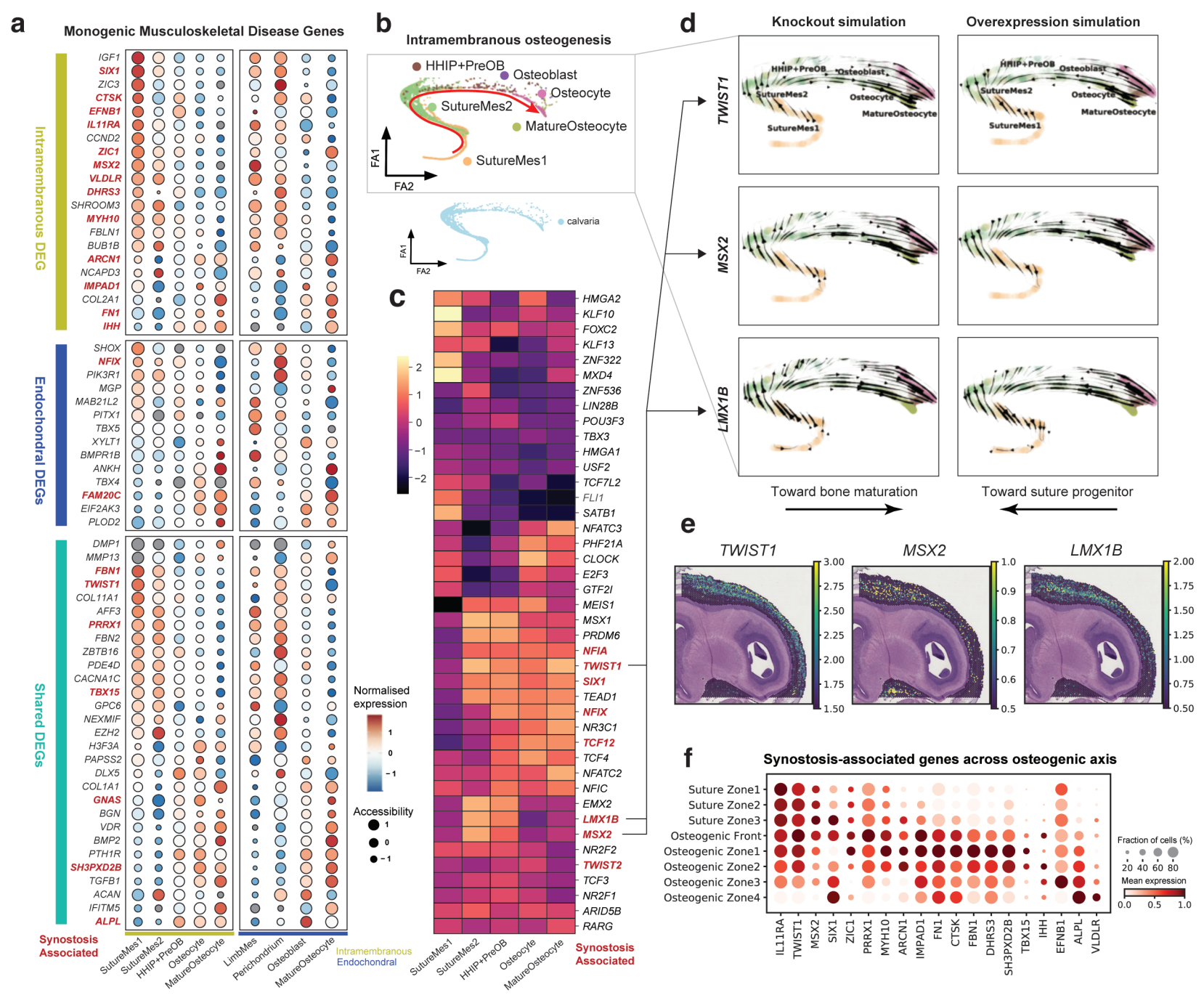

a
Source-Phenotype
Chondrogenesis
Osteogenesis
Other
OTG-Monoarthritis
GCST-Alpha angle
OTG-RA
NEALE-Spondylitis
NEALE-RA
GCST-RA2
GCST-RA1
GCST-JIA
FINNGEN-Spondylitis
FINNGEN-JIA
OTG-Osteoporosis
NEALE-Osteoporosis
NEALE-Osteopenia
OTG-Osteoarthrosis
BOER-Knee OA
BOER-Hip OA
BOER-All OA
BOER-TKR
BOER-THR
other
Inflammatory Arthritis
OP
OA
Joint Rep
Maturing Chondro
InterzoneChon
ChondroPro
Articular Chondro
PAX7+ Chondro
EPYC+ Chondro
Hypertrophic Chondro
PAX9+ Mes
ALX1+ Mes
Perichondrium
Osteoblast
HIC1+ Mes
Synovial Fibro
FibroPro
MyoFibro
DermFibro
Tenocyte
PAX7+ Myo
MYH3+ Myo
Endothelial
Erythroid
Epithelial
Immune
Schwann
Lymphatic Endothelial
b
Source-Phenotype
chondrogenesis
osteogenesis
GCST-Alpha angle
BOER-Knee OA
BOER-Hip OA
BOER-All-OA
BOER-TKR
BOER-THR
FDR
0.8
0.6
0.4
0.2
0.0
logOR
0
1
2
3
significant
no
yes
InterzoneChon
ArticularChon1
ChondroPro1
RestingChon
HypertrophicChon
DLK1+ Chon
CyclingChon
PAX7+ Chon
ArticularChon2
LimbMes
Perichondrium
Osteoblast
MatureOsteocyte
Source-Phenotype
interzone
GCST-alpha angle
BOER-Knee OA
BOER-Hip OA
BOER-All OA
BOER-TKR
BOER-THR
Fibro IZ
Dermal IZ
Hypertrophic IZ-2
Hypertrophic IZ-1
GDF5-high IZ
Early IZ
Articular IZ
c
SNP2cell
GWAS summ stat.
-log P value
chr
sc data
UMAP2
UMAP1
enhancer-GRN
TF
r1
r2
r3
r4
g1
g2
g3
cell type specific
GRN-enrichment
cell type enrichment
logOR
d
TCF7L2
FOXO1
NFATC2
RUNX1
TEAD1
TWIST1
NFATC1
SMAD3
NFIA
JUN
Hypertrophic Chondro
Maturing Chondro
ChondroPro
EPYC+ Chondro
DermFibro
Articular Chondro
Osteoblast
Synovial Fibro
FibroPro
Lymphatic Endothelial
Immune
MyoFibro
Schwann
Epithelial
InterzoneChon
PAX9+ Mes
Endothelial
Tenocyte
Erythroid
MYH3+ Myo
PAX7+ Myo
ALX1+ Mes
HIC1+ Mes
e
NFIB
NFATC2
RUNX2
NFIA
NFIX
NR3C1
FOXO1
RUNX1
TEAD1
NFIC
Perichondrium
LEPR+ Mes
SutureMes2
SutureMes1
PArchMes
LimbMes
CranialMes
FacialMes
HHIP+PreOB
Osteoblast
Osteocyte
MatureOsteocyte
f
Articular chondrocyte enrichment
for knee OA signal
FOXP1
NFATC1
CUX1
FOXP4
SOX5
RARG
RUNX2
RUNX1
ERG
RFX7
COL27A1
ETS2
TCF7L2
PRKCA
SOX6
FLI1
MEF2D
MEF2C
MAML2
NFAT5
NFATC2
FOXC1
SOX9
CRISPLD2
MEF2A
FOSB
SNORC
FOXA3
g
Perichondrium enrichment of hip OA signal
MAF
JUN
TEAD1
HNF4G
RUNX2
NFATC1
NFATC4
NFATC2
RUNX1
TFCP2
RUNX3
NFAT5
ETS2
NR3C1
SOX6
NFIC
MEF2A
JDP2
FOXO1
NFIX
ZEB1
NFIB
FOXP1
TRPS1
POU2F1
FLI1
NFIA
STAT5B
TCF7L2
TCF7L1
BPTF
ETV6
PRDM6
LDB2
h
Gene ontology enrichment
Skeletal System Development
Extracellular Matrix Organization
Cartilage Development
Regulation of chondrocyte differentiation
Connective Tissue Development
Positive Regulation Of Kinase Activity
Response to
Transforming Growth Factor Beta
Response To Lipid
Inositol Phosphate-Mediated Signaling
Glial Cell Differentiation
Gliogenesis
calcineurin-NFAT Signaling Cascade
Calcineurin-Mediated Signaling
Transmembrane Receptor Protein
Serine/Threonine Kinase Signaling pathway
Negative Regulation Of Canonical
Wnt Signaling Pathway
Cellular Response To Growth Factor
Stimulus
FDR
0
0.05
0.10
NES
6
5
4
Knee OA
Hip OA

Figure 5: Links to complex diseases of the skeleton. a) fGWAS analysis results for enrichment of GWAS signals for bone-related complex traits in broadly annotated cell states. **b)** Enrichment of selected complex traits in finely annotated osteo- and chondrogenic cell states (top) and InterzoneChon subclusters (bottom). **c)** Schematic representation of SNP2cell: scores derived from GWAS summary statistics and cell cluster marker scores are mapped and integrated across a gene regulatory network, highlighting enriched modules predicted to play a cluster-specific role in disease. **d)** Enrichment scores for knee OA across all cell clusters (heatmap, top) showing a cluster of similar enriched genes for chondrocytes and median enrichment scores (boxplot, bottom). **e)** Enrichment scores for hip OA across the osteogenic trajectory (heatmap, top) and median enrichment scores (boxplot, bottom). **f)** Articular chondrocyte and knee OA specific enriched sub-network. Brighter colours correspond to a larger enrichment score, relative to scores obtained from random permutations. **g)** Perichondrium and hip OA specific enriched sub-network. Brighter colours correspond to larger enrichment scores as in d. **h)** Gene Set Enrichment Analysis for GO Biological Process terms across Articular Chondro-knee OA and Perichondrium-hip OA enrichment scores.


a
Monogenic Musculoskeletal Disease Genes
Intramembranous DEG
Endochondral DEGs
Shared DEGs
IGF1
SIX1
ZIC3
CTSK
EFNB1
IL11RA
CCND2
ZIC1
MSX2
VLDLR
DHRS3
SHROOM3
MYH10
FBLN1
BUB1B
ARCN1
NCAPD3
IMPAD1
COL2A1
FN1
IHH
SHOX
NFIX
PIK3R1
MGP
MAB21L2
PITX1
TBX5
XYLT1
BMPR1B
ANKH
TBX4
FAM20C
EIF2AK3
PLOD2
DMP1
MMP13
FBN1
TWIST1
COL11A1
AFF3
PRRX1
FBN2
ZBTB16
PDE4D
CACNA1C
TBX15
GPC6
NEXMIF
EZH2
H3F3A
PAPSS2
DLX5
COL1A1
GNAS
BGN
VDR
BMP2
PTH1R
SH3PXD2B
TGFB1
ACAN
IFITM5
ALPL
Synostosis Associated
SutureMes1
SutureMes2
HHIP+PreOB
Osteocyte
MatureOsteocyte
LimbMes
Perichondrium
Osteoblast
Intramembranous
Endochondral
b
Intramembranous osteogenesis
HHIP+PreOB
Osteoblast
Osteocyte
SutureMes2
MatureOsteocyte
SutureMes1
FA1
FA2
calvaria
c
HMGA2
KLF10
FOXC2
KLF13
ZNF322
MXD4
ZNF536
LIN28B
POU3F3
TBX3
HMGA1
USF2
TCF7L2
FLI1
SATB1
NFATC3
PHF21A
CLOCK
E2F3
GTF2I
MEIS1
MSX1
PRDM6
NFIA
TWIST1
SIX1
TEAD1
NFIX
NR3C1
TCF12
TCF4
NFATC2
NFIC
EMX2
LMX1B
MSX2
NR2F2
TWIST2
TCF3
NR2F1
ARID5B
RARG
Normalised expression
Accessibility
Synostosis Associated
d
Knockout simulation
Overexpression simulation
TWIST1
MSX2
LMX1B
Toward bone maturation
Toward suture progenitor
e
TWIST1
MSX2
LMX1B
f
Synostosis-associated genes across osteogenic axis
Suture Zone1
Suture Zone2
Suture Zone3
Osteogenic Front
Osteogenic Zone1
Osteogenic Zone2
Osteogenic Zone3
Osteogenic Zone4
IL11RA
TWIST1
MSX2
SIX1
ZIC1
PRRX1
MYH10
ARCN1
IMPAD1
FN1
CTSK
FBN1
DHRS3
SH3PXD2B
TBX15
IHH
EFNB1
ALPL
VLDLR
Fraction of cells (%)
Mean expression
g
LMX1B
HNF4G
RUNX2
TCF12
TWIST1
NFATC2
MAFB
JUN
MAF
NFIX
SIX1
TEAD1
TEAD3
ZIC5
ALX4
MSX1
PLXNA4
EMX2
PRRX2
TYK2
LMX1A
MSX2
ZFHX3
h
Importance
KLF10
DLX3
PURA
DLX5
ATF4
SOST
KLF2
FOSL2
KLF13
JUNB
KLF4
NFATC1
MEF2C
SATB1

1642 **Figure 6: Links to Monogenic conditions affecting bone development. a)** Known genes
1643 mutated in musculoskeletal monogenic conditions with their expression and accessibility along
1644 the intramembranous and endochondral osteogenesis trajectory. **b)** Force-directed
1645 embedding of the endochondral trajectory showing cell cluster annotations. **c)** Heatmap of *in*
1646 *silico* TF knockout perturbation scores per cell cluster. A higher score indicates transcriptomic
1647 changes induced by the perturbation are promoting osteogenesis. **d)** TF perturbation vectors
1648 showing the direction of induced transcriptomic changes on the FA embedding for three
1649 selected TFs: TWIST1, MSX2 and LMX1B. KO simulation of all three genes promotes
1650 osteogenesis, while overexpression inhibits it. **e)** TF expression on Visium Cytassist of the
1651 frontal region of skull, enrichment of TFs in coronal suture and skull-forming layers. **f)**
1652 Expression of synostosis-associated genes along spatial zones defined using OrganAxis. **g)**
1653 Enhancer-GRN showing inter-regulation between TWIST1, MSX2 and LMX1B, inferred using
1654 SCENIC+. Circles and diamonds represent genes and regions with TF-binding sites,
1655 respectively. Region-gene links are coloured and scaled according to peak2gene importance,
1656 while TF-region links are coloured by TF. Blue circles represent regulated targets, orange
1657 circles regulators and white intermediate genes. **h)** Enhancer-GRN showing predicted
1658 regulation of SOST. Region-gene links are coloured and scaled according to peak2gene
1659 importance. Highlighted in red is a region containing the enhancer known to be mutated in van
1660 Buchem disease.