

1 **Epigenetic mechanisms of osteoarthritis risk in**

2 **human skeletal development**

3 Euan McDonnell¹, Sarah E Orr², Matthew J Barter², Danielle Rux³, Abby Brumwell²,
4 Nicola Wrobel⁴, Lee Murphy⁴, Lynne M Overmann⁵, Antony K Sorial², David A Young²,
5 Jamie Soul^{7*}, Sarah J Rice^{2*}

6

7 1. Computational Biology Facility, University of Liverpool, MerseyBio, Crown Street,
8 United Kingdom

9 2. Biosciences Institute, Newcastle University, Central Parkway, Newcastle upon Tyne,
10 United Kingdom

11 3. Orthopaedic Surgery, UConn Health, Farmington, Connecticut, USA

12 4. Edinburgh Clinical Research Facility, University of Edinburgh, Edinburgh, United
13 Kingdom.

14 5. Human Developmental Biology Resource, Newcastle University, International Centre
15 for Life, Central Parkway, Newcastle upon Tyne, United Kingdom.

16 6. Institute of Systems, Molecular and Integrative Biology, University of Liverpool,
17 Liverpool, United Kingdom.

18

19 *These authors made an equal contribution

20 Address correspondence to:

21 Dr Sarah Rice sarah.rice@ncl.ac.uk (lead contact)

22 Dr Jamie Soul Jamie.soul@liverpool.ac.uk

23

24 Word Count: 5199

25 Running Title: Articular cartilage methylome in human development

26 **Abstract**

27 The epigenome, including the methylation of cytosine bases at CG dinucleotides, is
28 intrinsically linked to transcriptional regulation. The tight regulation of gene expression during
29 skeletal development is essential, with ~1/500 individuals born with skeletal abnormalities.
30 Furthermore, increasing evidence is emerging to link age-associated complex genetic
31 musculoskeletal diseases, including osteoarthritis (OA), to developmental factors including
32 joint shape. Multiple studies have shown a functional role for DNA methylation in the genetic
33 mechanisms of OA risk using articular cartilage samples taken from aged patients. Despite
34 this, our knowledge of temporal changes to the methylome during human cartilage
35 development has been limited.

36 We quantified DNA methylation at ~700,000 individual CpGs across the epigenome of
37 developing human articular cartilage in 72 samples ranging from 7-21 post-conception weeks,
38 a time period that includes cavitation of the developing knee joint. We identified significant
39 changes in 8% of all CpGs, and >9400 developmental differentially methylated regions
40 (dDMRs). The largest hypermethylated dDMRs mapped to transcriptional regulators of early
41 skeletal patterning including *MEIS1* and *IRX1*. Conversely, the largest hypomethylated
42 dDMRs mapped to genes encoding extracellular matrix proteins including *SPON2* and *TNXB*
43 and were enriched in chondrocyte enhancers. Significant correlations were identified between
44 the expression of these genes and methylation within the hypomethylated dDMRs. We further
45 identified 811 CpGs at which significant dimorphism was present between the male and
46 female samples, with the majority (68%) being hypermethylated in female samples.

47 Following imputation, we captured the genotype of these samples at >5 million variants
48 and performed epigenome-wide methylation quantitative trait locus (mQTL) analysis.
49 Colocalization analysis identified 26 loci at which genetic variants exhibited shared impacts
50 upon methylation and OA genetic risk. This included loci which have been previously reported
51 to harbour OA-mQTLs (including *GDF5* and *ALDH1A2*), yet the majority (73%) were novel
52 (including those mapping to *CHST3*, *FGF1* and *TEAD1*).

53 To our knowledge, this is the first extensive study of DNA methylation across human
54 articular cartilage development. We identify considerable methylomic plasticity within the
55 development of knee cartilage and report active epigenomic mediators of OA risk operating in
56 prenatal joint tissues.

57

58 **Introduction**

59 The development of human limbs and synovial joints commences between four and eight
60 weeks post conception (pcw). During limb bud outgrowth, mesenchymal condensations form
61 the cartilage anlagen (driven by the transcriptional master-regulator, SOX9) to establish the
62 rudimentary embryonic skeletal elements. The limb synovial joints are established through
63 coordinated dedifferentiation of the nascent cartilage anlagen at the presumptive joint
64 locations to form a unique pool of progenitors at each location known as the interzone. From
65 these progenitors, all structures of the mature synovial joint are established including articular
66 cartilage, synovial lining, meniscus and intrajoint ligaments. For the knee joint,
67 chondrogenesis within the femoral condyles is advanced by the start of seven pcw (Carnegie
68 Stage (CS) 19), and a clear interzone has formed between the femur and the tibia, with full
69 cavitation of the joint and formation of the synovial cavity commencing at eight pcw¹. The first
70 appearance of menisci and ligaments also occurs within a similar developmental window,
71 whilst the formation of the infrapatellar fat pad does not appear until much later in development
72 (18pcw)¹. This dynamic process of limb and synovial joint development requires the
73 orchestrated expression genes essential for the development of defined and differentiated
74 tissues². The transcriptome is primarily regulated through the spatiotemporal expression of
75 transcription factors (TFs), yet epigenetic processes, including DNA methylation (DNAm) both
76 underlie and reinforce transitional plasticity during development³.

77
78 To date, DNAm remains the most extensively studied mammalian epigenetic mechanism.
79 Methylation of DNA occurs at Cytosine-phosphate-Guanine (CpG) dinucleotides, of which
80 there are approximately 28 million within the human genome. CpGs are recognised by DNA
81 methyltransferase (DNMT) enzymes, which actively catalyse the addition of a methyl group
82 from a S-adenosyl-L-methionine (SAM) donor to form 5-methyl cytosine (5mC)⁴. DNAm is
83 intrinsically linked to transcriptional regulation, ostensibly to gene repression, by preventing
84 binding of transcriptional activators to promoter regions, and further through the recruitment

85 of repressive methyl-binding proteins. However, the relationship between DNAm and gene
86 expression is far from straightforward, with gene body methylation often being associated with
87 active transcription⁵. It is generally considered that DNAm within *cis*-regulatory elements
88 (CREs) is repressive to the expression of the gene target^{6,7}.

89

90 The methylome of human articular cartilage has been extensively studied, primarily in the
91 context of the chronic joint disease osteoarthritis (OA). OA is a global leading cause of
92 disability amongst older adults, hallmarked by the degradation of articular cartilage in the
93 joints, most commonly the hip or knee. Studies of the aged cartilage methylome have revealed
94 distinct epigenomic signatures between disease states^{8,9}, and between joint sites
95 themselves^{10,11}, increasing our understanding of the joint specificity of disease. OA is
96 multifactorial, with genetic risk factors contributing to ~30% of the lifetime risk of developing
97 knee OA. The integration of DNAm data into genetic studies, such as the statistical fine
98 mapping of GWAS signals, has identified the co-localisation of methylation quantitative trait
99 loci (mQTLs) with genetic risk signals^{8,11-13}. This interplay between DNA sequence and CpG
100 methylation status¹⁴ has further been shown to underpin tissue-specific molecular
101 mechanisms of gene expression within the joint^{15,16}.

102

103 Developmental factors also play a role in the risk of OA onset and progression in older age¹⁷.
104 This includes both the shape of the joint¹⁸⁻²⁰ (which affects the biomechanical properties and
105 weight-bearing capacity) and the biochemical composition of the articular cartilage (which
106 affects the resilience of the tissue to withstand stresses throughout the life course). Our recent
107 targeted study of 39 CpGs investigated the presence of OA mQTLs in human foetal limbs, at
108 seven well-characterised OA genetic risk loci²¹. We identified that at 85% of the CpGs, the
109 significant OA-mQTLs replicated in the foetal tissues, demonstrating for the first time that
110 functional epigenetic mechanisms associated with a musculoskeletal disease of older age
111 operate from the start of life²¹.

112

113 Murine knockout or inactivation of the three known DNA methyltransferases (DNMTs) is lethal
114 embryonically (*Dnmt1* and *Dnmt3b*) or postnatally (*Dnmt3a*), demonstrating the vital role of
115 DNAm throughout development^{22,23}. DNA methylating enzymes are expressed in proliferating
116 chondrocytes during embryonic development and persist in articular chondrocytes during
117 post-natal development²⁴. Investigation of the methylome of human developmental cartilage
118 has been limited to two studies to date, which utilised small sample numbers (n=8) falling
119 within a narrow developmental window (14-19pcw)²⁵, or, in the case of our previous study,
120 applying targeted approaches which captured only a handful of CpGs²¹. Here, we present a
121 comprehensive study of the methylomic trajectory of human articular cartilage in a broader
122 window of development to capture novel risk loci for joint disease.

123 In this study, we examine the DNA methylome of 72 foetal cartilage samples from the
124 developing knee (distal femur) between 7 (CS 23) and 21pcw. We identify the presence of
125 differentially methylated regions across this developmental timeframe and examine their
126 enrichment in CREs and TF binding sites. Furthermore, we investigate sexual dimorphism
127 within the cartilage methylome during human development. Importantly, we identify the
128 presence of developmental epigenomic changes contributing to the risk of OA in older life
129 through the identification of foetal cartilage mQTLs co-localising with OA genetic risk signals.

130

131

132

133

134

135 **Methods**

136 **Sample Collection**

137 Cartilaginous tissue from the distal end of the developing human femur was supplied from the
138 MRC and Wellcome Trust-funded Human Developmental Biology Resource (HDBR) at
139 Newcastle University (<http://www.hdbbr.org>, project number 200363), Newcastle upon Tyne.
140 Tissues were obtained with appropriate maternal written consent and approval from the
141 Newcastle and North Tyneside NHS Health Authority Joint Ethics Committee. HDBR is
142 regulated by the UK Human Tissue Authority (HTA; www.hta.gov.uk) and operates by the
143 relevant HTA Codes of Practice.

144

145 **Generation of methylation data and genotyping**

146 DNA was genotyped using Illumina Global Screening Array v3 BeadChip at the Edinburgh
147 Clinical Research Facility. Bisulphite conversion of DNA (500ng) was conducted using the EZ
148 DNA methylation kit (Zymo) with Illumina-advised adjustments to the standard
149 protocol. Methylation data was generated using the Illumina Methylation850 EPIC v1 array.
150 The arrays were imaged on the Illumina iScan platform. Detected and autosomal probes were
151 retained while those detected by less than 3 beads, with low detection P-values ($P < 0.001$),
152 present at known SNP sites, or known to be cross-reactive were removed from the analysis^{26–}
153 ²⁸.

154

155 **Transcript expression analysis measured by RT-qPCR.**

156 RNA (1µg) was reverse transcribed with the SuperScript IV cDNA synthesis kit (Invitrogen).
157 The cDNA product was then used for quantitative real-time PCR (RT-qPCR) analysis following
158 1:20 dilution. Gene expression was quantified using TaqMan chemistry (QuantStudio3,
159 Thermo Fisher). Pre-designed TaqMan assays (Integrated DNA Technologies, Belgium) were
160 used to measure the expression of the target genes (Table S1). Gene expression was

161 analysed relative to the expression of three housekeeping genes: *18S*,
162 *GAPDH* and *HPRT1* using the $2^{-\Delta\text{Ct}}$ method as previously described¹¹.

163

164 **mQTL analysis**

165 mQTL analysis was performed using the matrixEQTL R package (v2.3), including Sentrix ID
166 (batch), sex and gestation stage as covariates and testing for *cis*-associations based on a
167 distance threshold of 500kb²⁹. To account for ancestry, zero to four ancestry principal
168 components (PC) from the 1000 genomes PC analysis (PCA) were trialled for inclusion in the
169 mQTL analysis. The number of nominally significant mQTLs was compared between the
170 different models and two PCs (79.86% of the total variance) were included in the final model.
171 Resultant mQTLs were filtered based on an FDR < 0.05 or a conservative Bonferroni
172 correction calculated using the matrixEQTL-reported total number of statistical tests.
173 Significant mQTLs were associated with the protein-coding genes that they were closest to,
174 based on EnsDb (v75) annotations³⁰.

175

176 **Colocalisation analysis**

177 The previously reported 100 independent genome-wide significant SNVs associated with 11
178 OA skeletal site phenotypes were used to assess colocalisation with the mQTL data³¹.
179 Summary statistics for each phenotype were obtained from the musculoskeletal knowledge
180 portal (<https://msk.hugeamp.org/>). MAF was calculated from the effect allele frequency, and
181 case and control sizes for each phenotype were extracted from the respective publication. To
182 assess mQTLs within a range of $\pm 500\text{kb}$ from independent OA risk loci signal SNVs, we
183 repeated the mQTL analysis while retaining all mQTL results (`pvOutputThreshold=1`). For
184 methylation probes within $\pm 500\text{kb}$ of the risk loci, we tested associated SNVs for colocalisation
185 with genetic variants found in both the *cis* mQTL and the GWAS data. Colocalisation analyses
186 were performed separately for each GWAS phenotype using the `coloc` R package³² (v 5.1.0)
187 and the `coloc.abf` function, which calculates posterior probabilities for four hypotheses:

188

189 H0: Neither trait has a genetic association in the region.

190 H1: Only trait 1 has a genetic association in the region.

191 H2: Only trait 2 has a genetic association in the region.

192 H3: Both traits are associated but with different causal variants.

193 H4: Both traits are associated and share a single causal variant.

194

195 A posterior probability of $H4 > 0.8$ was considered as evidence for colocalisation. We further

196 filtered results to retain SNVs present in our imputed genotype data and where the mQTLs

197 was nominally significant ($P < 0.05$).

198

199

200

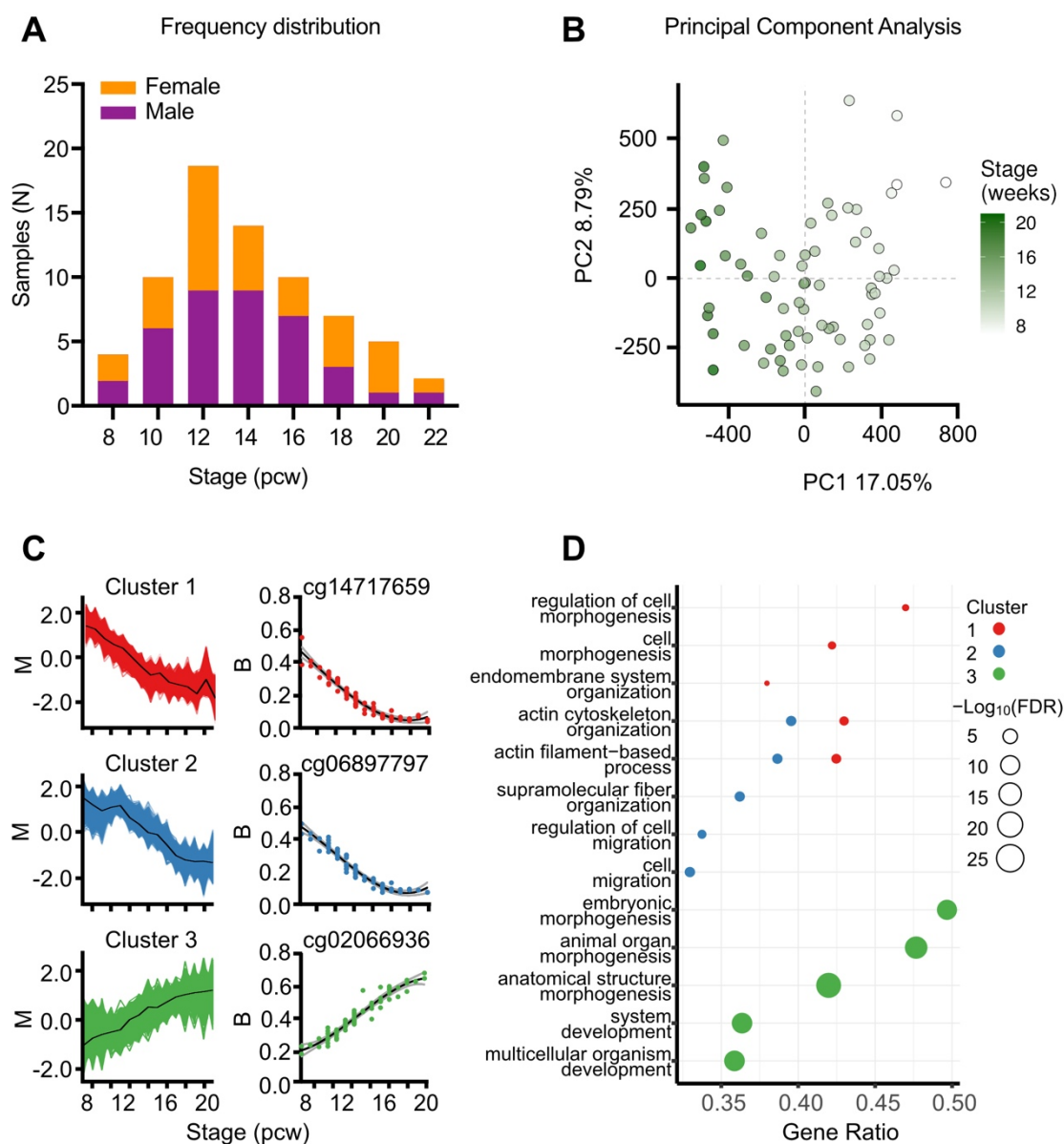
201 **Results**

202 **The DNA methylome remains dynamic throughout cartilage development.**

203 We isolated articular cartilage from the distal femur of 72 samples ranging from 7-21pcw
204 (Fig.S1A). The mean(\pm SEM) gestational stage was 13.3 ± 0.4 pcw (Fig.1A). Within the isolated
205 cartilage, gene expression analysis of chondrocyte progenitor markers *GDF5* and *SOX9*
206 showed a significant decrease in expression with increasing developmental stage ($P=0.011$
207 and 0.006 , respectively; Fig.S1B-C). Conversely, the expression of articular chondrocyte
208 marker *PRG4* significantly increased with stage ($P=0.003$; Fig.S1D). The Type II collagen
209 gene, *COL2A1*, was highly expressed throughout the stages and did not significantly change
210 ($P=0.219$; Fig.S1E).

211 We profiled the DNA methylation (DNAm) of the 72 samples using Illumina
212 HumanMethylation850 EPIC v1 microarrays. One sample was dropped due to mismatch
213 between predicted and labelled sex. The remaining 71 samples (Table S2) had a minimum
214 bisulphite conversion rate of 89.69% and following QC we retained data at 678,267 CpGs.
215 Data were normalised via quantile-normalisation, followed by principal component analysis
216 (PCA), where developmental stage was found to strongly correlate with PC1, which accounted
217 for 17% of the total variance in global methylation (Fig. 1B). Across the captured
218 developmental window, we identified 53,800 significantly differentially methylated probes
219 (DMPs; $FDR<0.001$, $\log_2fc>0.1$; Table S3), with 48% becoming significantly hypomethylated
220 through skeletogenesis. We analysed the trend in methylation changes across the
221 developmental stages, identifying three clusters. Clusters 1 and 2 became significantly
222 hypomethylated across the developmental window (Fig. 1C; red and blue, respectively). The
223 primary difference separating the two groups was an observable lag in decreased methylation
224 in Cluster 2, where DNAm remained stable until >10 pcw. Conversely, the CpGs belonging to
225 Cluster 3 (Fig. 1C; green) became hypermethylated throughout development. GO analysis
226 revealed similar and shared significant terms ($FDR<0.05$) for Clusters 1 and 2, including
227 biological processes such as “actin filament-based processes” and “cell morphogenesis”

228 (Fig.1D) mapping to cellular compartments including “actin cytoskeleton” and “focal adhesion”
 229 (Table S4). Analysis of Cluster 3 identified significant enrichment of biological processes
 230 including “embryonic morphogenesis” and “anatomical structure development” (Fig.1D) and
 231 the cellular compartments “collagen-containing extracellular matrix” and “transcriptional
 232 regulator complex” (Table S4).



233

234 **Figure 1 DNA methylation significantly changes at individual CpGs in foetal knee**
 235 **cartilage across the 7-21pcw developmental window. A**, Histogram of samples (N=71)
 236 showing the distribution of developmental stage. Purple, Male; Orange, female. **B**, Principal
 237 component analysis of the samples revealed that PC1 is associated with the developmental
 238 stage. **C**, The DMPs fell into three distinct clusters based upon the methylation trends across

239 the developmental window. The mfuzz plot for each cluster is displayed, along with the most
240 significantly differentially methylated CpG belonging to each cluster. **D**, Gene ontology (GO)
241 analysis of the top significant biological processes enriched in each of the three clusters. Gene
242 ratio is the ratio of the proportion of genes in a GO term annotated with at least one significant
243 CpG relative to the proportion of genes in a GO term annotated with a detected CpG.

244
245

246 To gain further insight into the biological role of the DMPs, we next investigated the presence
247 of differentially methylated regions (DMRs) based upon regions of DNAm co-regulation. We
248 identified 9,462 developmental DMRs (dDMRs) ranging in size between 2 and 59 CpGs (mean
249 = 2.9) consisting of 27,447 individual sites (FDR<0.001, log₂fc>0.1; Fig.2A, Table S5). In
250 Figure 2A the most significant dDMRs (irrespective of size) are annotated with the nearest
251 protein coding gene. The largest hypermethylated dDMRs (n=21-59) mapped to genes
252 encoding known transcriptional regulators of early development including the *HOXA* gene
253 locus and *TBX3* (Fig.2B). GO analysis revealed 326 significant (FDR<0.01) biological
254 processes associated with hypermethylated dDMRs (Table S6). The top 20 most significant
255 biological process terms included “skeletal system development”, “anatomical structure
256 development”, and “embryonic morphogenesis” (Fig.2C). Just 20 biological processes were
257 significantly associated with hypomethylated regions (FDR<0.01; Table S7), which mapped to
258 genes encoding ECM proteins including *TNXB* (Tenascin-X) and *SPON2* (Spondin-2; Fig.2B0.
259 “Supramolecular fiber organisation” was the highest enriched biological process (gene ratio,
260 0.22; FDR = 0.006) along with multiple terms relating to the innate immune system (Fig.2C).

261

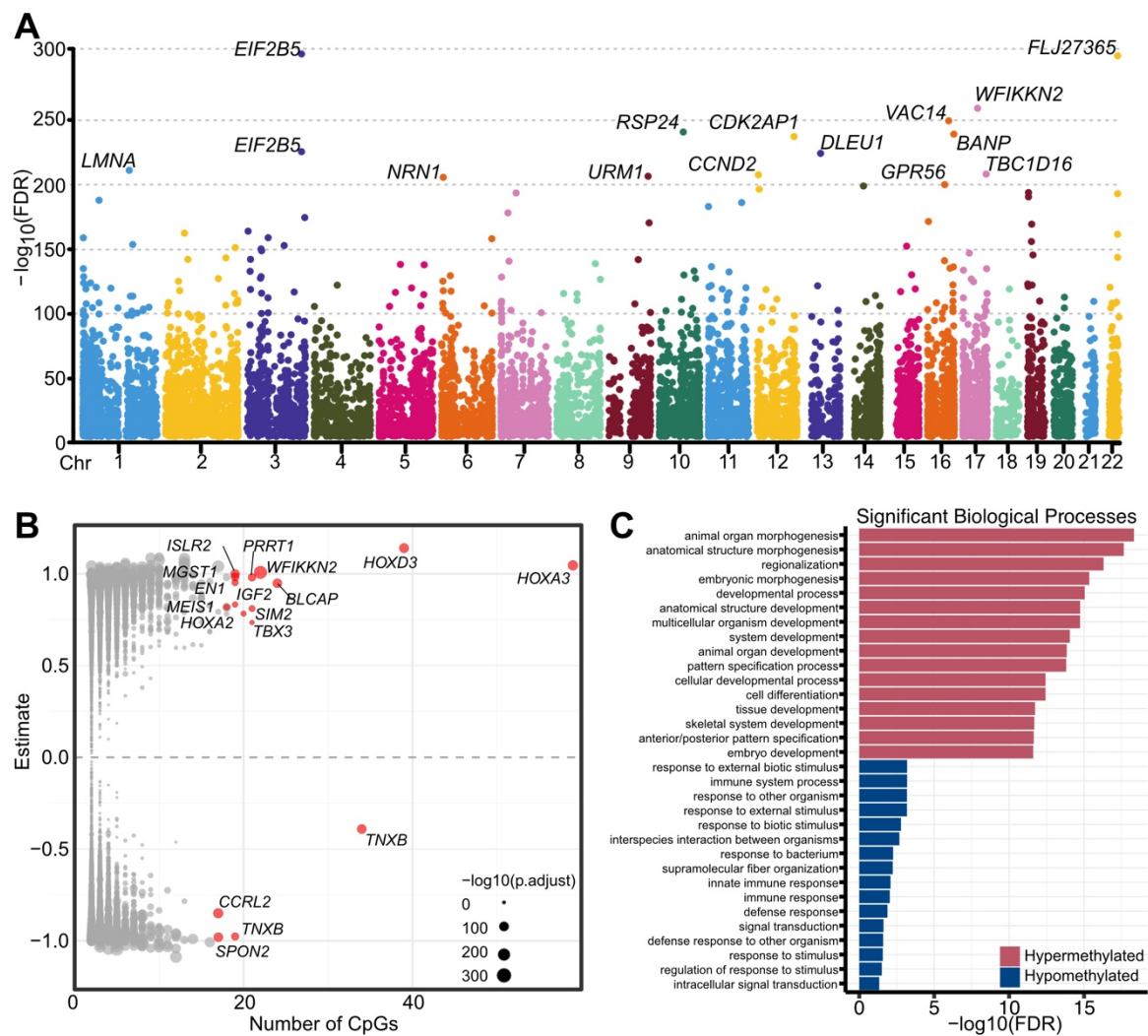
262

263

264

265

266



267

268 **Figure 2. Co-regulation of DNAm at individual CpGs results in developmental**
 269 **differentially methylated regions (dDMRs) in foetal knee cartilage. A, Manhattan plot of**
 270 **dDMRs across the epigenome. The nearest protein-coding gene to the most significant**
 271 **dDMRs is labelled. B, Volcano plot of dDMRs by size (number of CpGs). The most significant**
 272 **are highlighted in red with the nearest protein-coding gene labelled. C, Gene ontology (GO)**
 273 **analysis of the most significant biological processes enriched in hypermethylated (red) and**
 274 **hypomethylated (blue) dDMRs.**

275

276

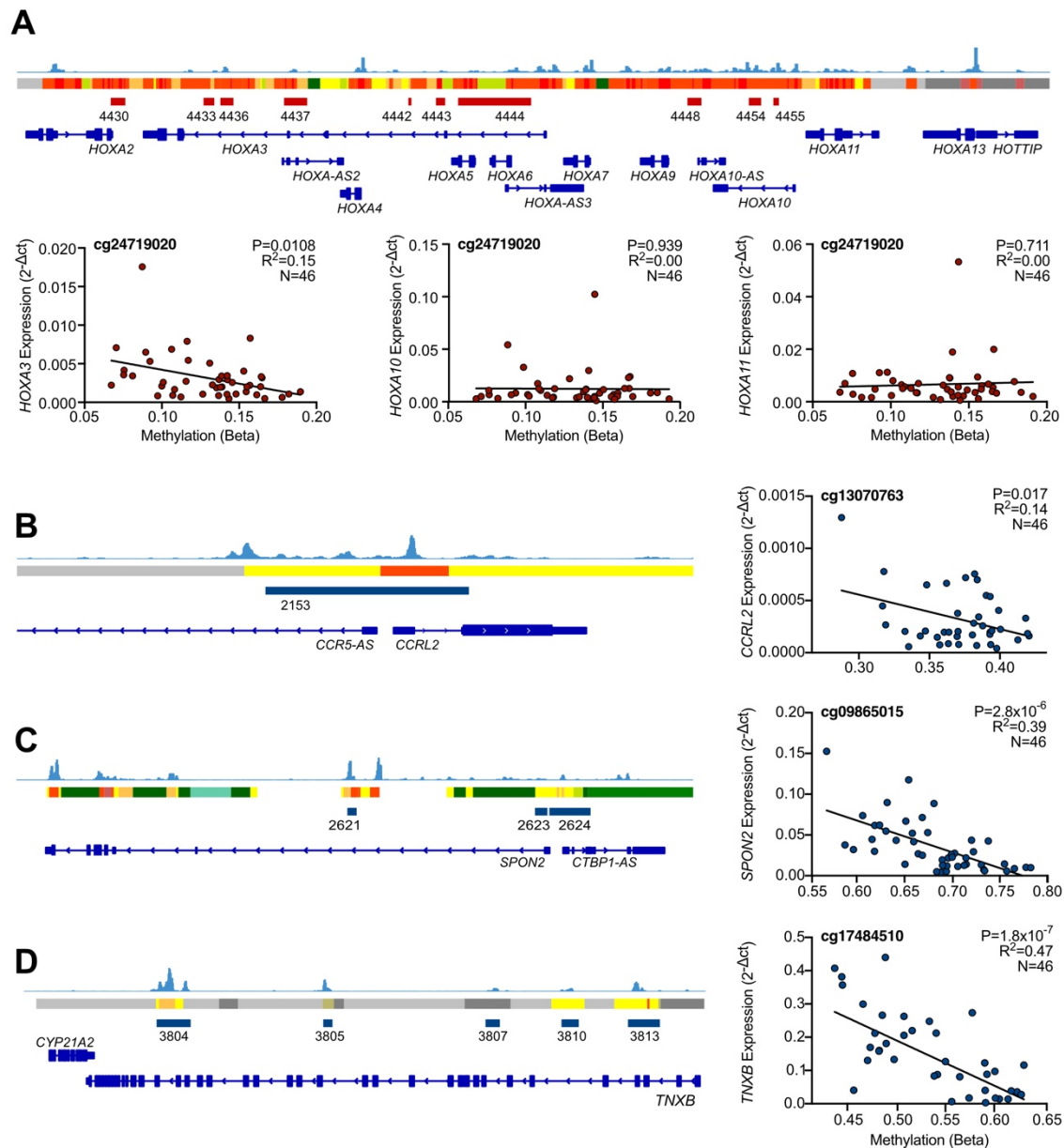
277 **Developmental DMRs overlap with open chromatin regions and correlate with**
278 **gene expression.**

279 We next intersected the physical location of the dDMRs with open chromatin peaks from
280 12pcw foetal distal femoral cartilage. We identified overlap between 3792 dDMRs (40%),
281 mapping to 2558 genes, further highlighting those most likely to impart regulatory function
282 during development (Table S8). The largest hyper- and hypomethylated regions overlapping
283 with foetal knee ATAC peaks are displayed in Table 1.

284 To investigate the putative functional role of these DMRs in developmental gene
285 regulation, we measured the expression of eight of the genes mapping to the dDMRs presented
286 in Table 1. The six transcripts mapping to hypermethylated regions (*MEIS1*, *EN1*, *HOXD3*,
287 *IRX1*, *HOXA3*, *SIM2*) were expressed at low levels throughout the captured developmental
288 window (Fig.S2, red). The expression of *HOXA3* negatively correlated with DNAm at 25/59
289 individual CpGs within the dDMR ($P=0.048-0.008$; Table S9, Fig.3A LHS). Interestingly, the
290 *HOX* genes mapping to this dDMR do not have reported functions in limb development. We
291 compared their expression to that of *HOXA10* and *HOXA11*, which have recently been
292 reported to be expressed specifically in the zone between the proximal and distal human
293 hindlimb at 5.6pcw, in the region where the knee cavity will form³³. These genes are also
294 located at the *HOXA* gene locus, a region of tightly controlled co-regulation (Fig.3A upper
295 panel). However, their expression showed no correlation with DNAm within this region
296 ($P>0.05$; Fig.3A middle and RHS; Table S9).

297 All three genes mapping to the largest hypomethylated regions, *CCRL2* (labelled
298 dDMR 2153), *SPON2* (dDMR 2624) and *TNXB* (dDMRs 3804 and 3813), fell within putative
299 chondrocyte enhancers and regions of open chromatin in foetal chondrocytes (Fig.3B-D). The
300 expression of all three genes significantly increased in correlation with decreasing methylation
301 within the mapped dDMRs at 3/17 ($P=0.017 - 0.041$), 6/17 ($P=2.8\times 10^{-6} - 0.002$) and 42/53
302 ($P=1.82\times 10^{-7} - 0.017$) CpGs, respectively (Fig.3B-D).

303
304



305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320

Figure 3. The expression of genes mapping to dDMRs significantly correlates with methylation within the regions. A, (Upper) Screenshot of the Integrated Genome Viewer (IGV) displaying the HOXA gene locus on chromosome 7. Foetal knee cartilage ATAC-seq peaks (light blue) show regions of open chromatin. Below the peaks, the ROADMAP chromatin state track from human cultured chondrocytes (E049) is displayed. Red, active transcription start site; yellow/orange, enhancer; green, transcribed; grey, repressed. Hypermethylated dDMRs are shown in dark red and annotated with the dDMR ID. Genes within the region are shown in bright blue with the gene names labelled. (Lower) Gene expression of HOXA3 (LHS), HOXA10 (middle) and HOXA11 (RHS) measured by RT-qPCR regressed against DNAm at cg24719020, which falls within dDMR 4444. **B-D, (LHS)** IGV screenshots of chromosomes 3, 4, and 6, respectively. Foetal ATAC-seq data is displayed as described in (A). Hypomethylated dDMRs are shown in blue. **RHS,** gene expression of CCRL2, SPON2, and TNXB regressed against DNAm at a single CpG within each of the dDMRs mapping to the respective genes. Statistical analysis was performed by multiple linear regression using methylation M-values. Beta values are displayed for ease of interpretation.

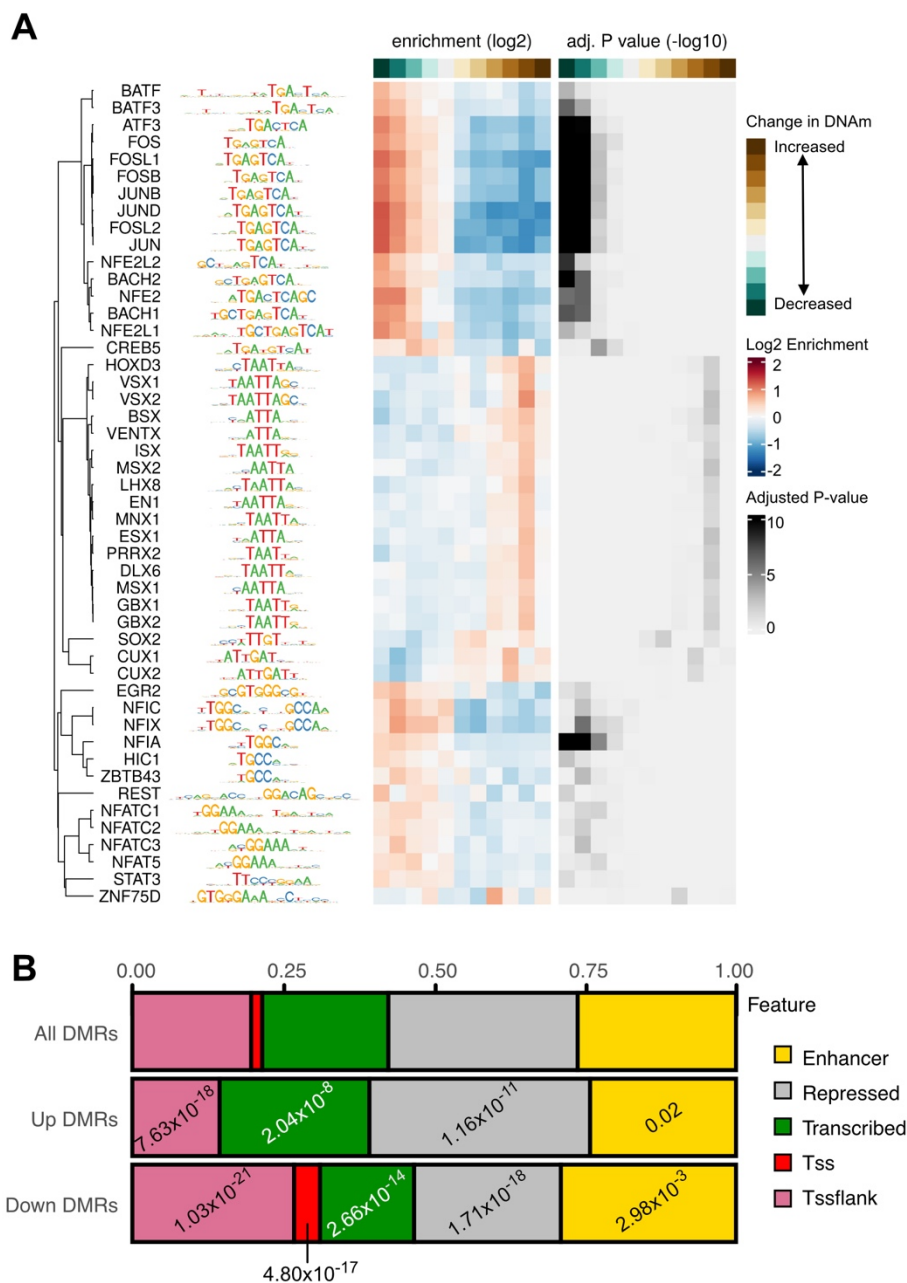
321 **Developmental DMRs are enriched in active enhancers and transcription factor**
322 **binding sites.**

323 We performed motif enrichment analysis on the hyper- and hypomethylated dDMRs for
324 enrichment of known transcription factor (TF) binding motifs. We identified significant
325 enrichment ($FDR \leq 0.05$) of 48 known TF motifs on the HOCOMOCO database (Fig.4A, Table
326 S10). Notably, the regions that decreased in methylation across the developmental window
327 were enriched for TFs with known roles in chondrocyte biology and synovial joint formation
328 including NFATC1 ($P_{adj}=0.02$)³⁴, CREB5 ($P_{adj}=9.7 \times 10^{-5}$)³⁵, and the FOS/JUN families of
329 TFs ($P_{adj} < 8.1 \times 10^{-19}$). The regions that increased in methylation across development were
330 enriched for motifs of early transcriptional regulators of skeletogenesis including HOXD3
331 ($P_{adj}=0.01$) and CUX1/2 ($P_{adj} < 0.04$; Fig.4A)³⁶.

332 Finally, we overlapped the dDMRs with Roadmap chromatin state data generated in
333 human chondrocytes (E049) to test for overrepresentation amongst distinct categories of
334 CREs. Hypermethylated dDMRs were significantly overrepresented in transcribed (Fig.4B,
335 green; $P_{adj}=2.0 \times 10^{-8}$) and repressed regions (Fig.4B, grey; $P_{adj}=1.2 \times 10^{-11}$) when compared
336 to all identified significant dDMRs. No hypermethylated dDMRs overlapped directly with
337 transcriptional start sites (TSS; Fig.4B, red). Conversely, hypomethylated dDMRs were
338 overrepresented in enhancers ($P_{adj}=2.9 \times 10^{-3}$) and TSS ($P_{adj}=4.8 \times 10^{-17}$) but
339 underrepresented in repressed regions ($P_{adj}=1.7 \times 10^{-18}$), indicating that across development,
340 regions becoming hypomethylated were predominantly regulating the function of CREs.

341

342



343

344 **Figure 4 Developmental DMRs are enriched in regulatory elements.** **A**, Transcription
 345 factor (TF) motifs taken from the Hocomoco database which are enriched in dDMRs **Left**,
 346 name and logo plot of enriched TF motif within the dDMRs, **Centre**, Heatmap of Log2
 347 enrichment value of the motifs within hypermethylated (brown) or hypomethylated (green)
 348 dDMRs. **Right**, Heatmap of $-\text{Log}_{10}$ adjusted P-value. Darker colour indicates higher statistical
 349 significance. **B**, Frequency of identified dDMRs overlapping with human cultured chondrocyte
 350 (Roadmap E049) chromatin states. Enhancer, yellow; repressed region, grey; transcribed
 351 region, green; transcription start site (TSS), red; region flanking TSS, pink. Statistical
 352 significance was calculated using Fisher's exact test with Bonferroni correction with all
 353 significant DMRs as background. Significant under- or overrepresentation is labelled with the
 354 adjusted P-value.

355

356 **Chondrocyte DNA methylation in articular cartilage development exhibits**
357 **sexual dimorphism.**

358 Previously, methylome analysis performed in 358 blood samples from extremely premature
359 neonates (at 23-27 weeks' gestation, which relates to ~21-25pcw) showed that a large
360 proportion of CpGs exhibited sexually dimorphic levels of DNAm³⁷. Using the EPIC array,
361 Santos *et al.*, identified 5595 DMPs (FDR<0.00001), 95% of which were hypermethylated in
362 females. We tested for significant sex-specific differences in DNAm amongst our
363 developmental cartilage samples. Epigenome-wide analysis revealed 811 sex-specific DMPs
364 (sDMPs; FDR<0.05), mapping to 534 genes (Table S11, Fig.S3A). Of these sites, 523 (64.5%)
365 were hypermethylated in the female samples. Fifty-seven percent (461) of the sDMPs we
366 identified in developing articular cartilage also showed significant sex effects in the neonatal
367 blood dataset (Fig.S3B). There was a direct overlap between all of the top 20 reported probes,
368 with methylation effects occurring in the same direction (Table 2, Fig.S3C). Amongst the
369 genes mapping to these CpGs are *NAB1*, *RFTN1*, and *CEP170*.

370
371 We further identified 60 significant sDMRs (FDR<0.01), consisting of 178 CpGs (Table S12).
372 At 46 identified regions (77%), DNAm was significantly higher in the female samples than in
373 male. No significant GO terms (all FDR>0.05) were associated with either the sDMPs or
374 sDMRs, consistent with the findings of Santos *et al.*, in their analysis of neonatal blood
375 samples³⁷.

376
377 **mQTLs co-localizing with OA risk signals (OA-mQTLs) are active during**
378 **articular cartilage development.**

379 We next investigated the presence of mQTLs in the developing human knee cartilage.
380 Epigenome-wide analysis revealed 399,952 significant (Bonferroni-corrected P<0.05) *cis*-
381 mQTLs (Table S13). The signals were widespread across the epigenome (Fig.5A) and
382 consisted of 11,299 individual CpGs and 257,908 SNVs. The most significant mQTLs mapped

383 to the genes *ELMOD1* ($P_{\text{adj}}=3.7 \times 10^{-50}$), *ATPIF1* ($P=1.4 \times 10^{-48}$), and *PHGDH* ($P_{\text{adj}}=4.0 \times 10^{-47}$).

385 To identify shared causal impacts of SNV genotype upon OA risk and DNAm, we
386 performed a colocalisation analysis of 100 risk signals from 11 OA phenotypes with the foetal
387 cartilage mQTLs. In total, we identified 104 colocalisations with foetal cartilage mQTLs
388 (posterior probability (PP) >0.8; Table S14) amongst 49 CpGs spanning 26 genomic loci
389 (Table 3). At 7 of the loci, OA-mQTLs were identified which have previously been reported in
390 investigations of aged adult cartilage. We overlapped the physical location of the CpGs with
391 open chromatin regions in foetal and OA knee chondrocytes²¹, regions of H3K27ac
392 enrichment in foetal limb tissues³⁸, and with designated chromatin states in mesenchymal
393 stem cells (MSCs) and chondrocytes³⁹. Through this, we compiled *in silico* evidence of
394 functional impact for the identified mQTLs restricted to development (D) or throughout the life
395 course (L). Evidence supporting function was identified at 19 loci, with around half (9)
396 consisting of mQTLs which fell within open chromatin or H3K27ac-enriched regions solely in
397 developmental samples.

398 One example of OA-mQTLs operating in both foetal and aged human cartilage is on
399 chromosome 15, where mQTLs at cg12031962 and cg12382153 map to *ALDH1A2*
400 ($P=7.4 \times 10^{-13}$ and 1.3×10^{-4} , respectively). These mQTLs co-localise with the primary
401 phenotypes “Hand OA” (PP=0.99, Fig. 5B), “Thumb OA” (PP=0.99), “Total Knee
402 Replacement” (PP=0.98) and “Finger OA” (PP=0.93; Table S14) marked by the risk SNV
403 rs11071366 (A>T). At both CpGs, OA-mQTLs have previously been reported in adult cartilage
404 (Table 3). The methylation site cg12382153 falls within an open chromatin peak flanking the
405 *ALDH1A2* transcriptional start site (Fig.5C). Here, the major and OA effect allele, A, correlates
406 with a decrease in DNAm at both CpGs, consistent with reports in aged cartilage⁸.

407 Conversely, on chromosome 20, a single intronic mQTL was identified which maps to
408 a previously unreported site at a known OA risk locus, mapping to the gene *GDF5* (Fig.5D).
409 Here, DNAm at cg12548730 co-localised with the OA risk SNV rs143384 in the “All OA” (PP
410 = 0.916), “Total Joint Replacement” (PP = 0.915) and “Knee and Hip OA” (PP = 0.891) primary

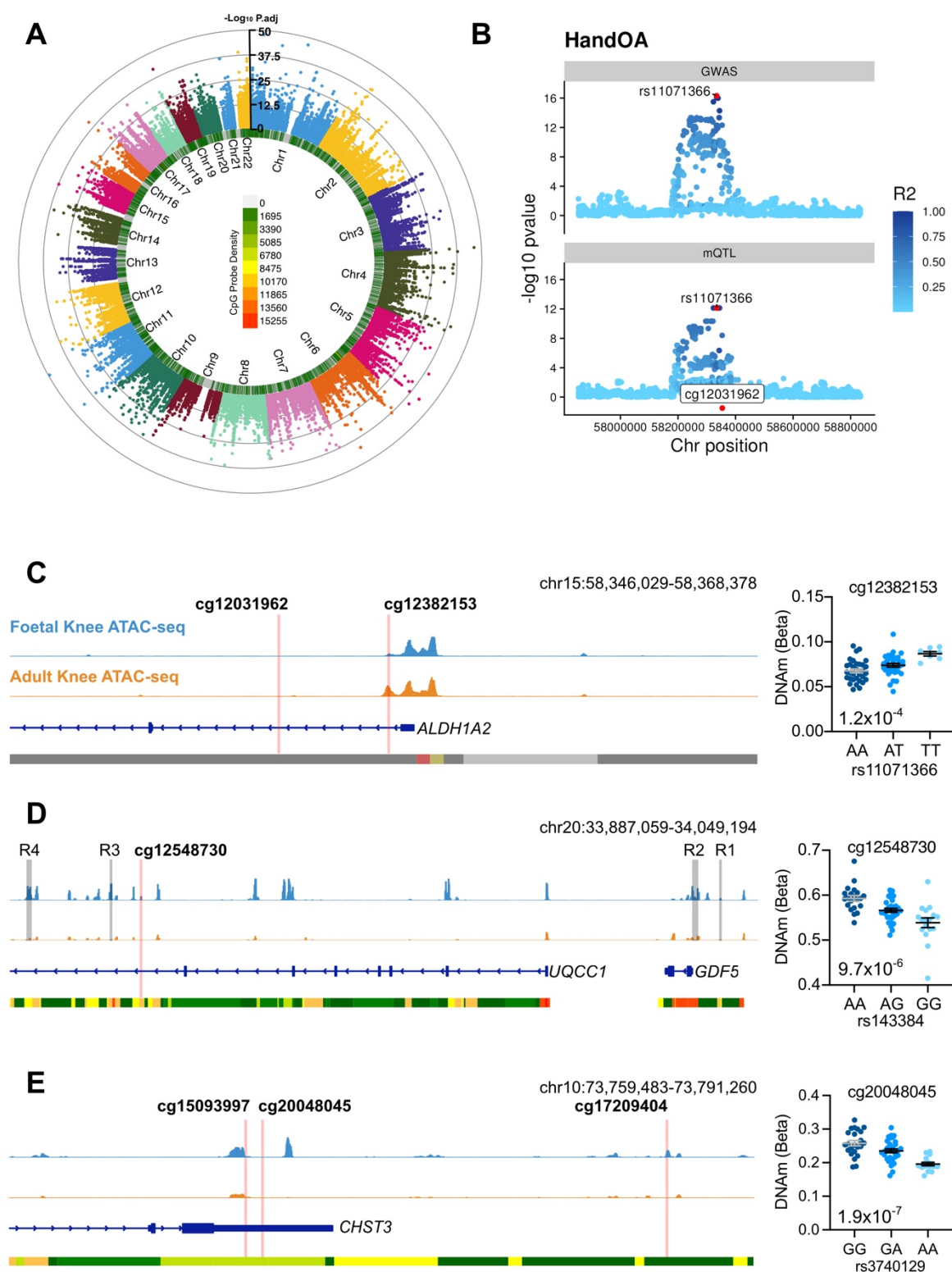
411 phenotypes (Table S14). The CpG is located within an intron of *UQCC1*, downstream of
412 previously reported *GDF5* cartilage enhancers R3 and R4^{40,41}. The OA effect allele (G)
413 significantly correlated with lower levels of DNAm ($P=9.7 \times 10^{-6}$; Fig.5D). Analysis of our existing
414 knee cartilage ATAC-seq dataset in human foetal (distal femur) and adult cartilage samples
415 showed that the CpG is located adjacent to an open chromatin peak in foetal cartilage (blue),
416 which is not an accessible region in OA cartilage (orange). This region has not previously been
417 investigated as a regulatory element for *GDF5*. A previous mQTL at this locus has been
418 reported in adult OA cartilage, located adjacent to the promoter of *UQCC1* (cg14752227)¹¹.

419 At the majority of the loci (19), no previous reports of OA-mQTLs have been reported
420 through analyses of adult cartilage. On chromosome 10, three OA-mQTLs were identified that
421 fell within putative regulatory elements of *CHST3*, encoding carbohydrate sulfotransferase 3
422 (Fig.5E). Two of the CpGs, cg15093997 and cg20048045, located within the 3'UTR of the
423 gene, near to open chromatin regions (Fig.5E). At both CpGs, the effect allele, A, at rs3740129
424 (G>A) correlated with significantly decreased levels of DNAm ($P=1.6 \times 10^{-12}$ and 1.9×10^{-7} ,
425 respectively). Conversely, at the more distal CpG (cg17209404), located within an intergenic
426 open chromatin region only in foetal cartilage, DNAm was significantly increased in the
427 presence of the OA effect allele ($P=3.9 \times 10^{-4}$; Table S14).

428

429 **Figure 5. Epigenome-wide mQTL analysis reveals co-localisation with OA**
430 **genetic risk signals in developmental cartilage. A**, Manhattan plot of significant
431 (Bonferroni adjusted P -value <0.05) epigenome-wide mQTL signals. The inner
432 heatmap represents the Illumina probe density in the respective regions across the
433 genome. **B**, An example of a colocalisation event. The mQTL for the methylation site
434 cg02900766 (bottom) colocalised with the GWAS signal rs11071366 for hand OA in
435 the same genomic region (top). Here, we observed a posterior probability (PP) for a
436 shared causal variant of 99%. **C-E**, Integrated Genome Viewer (IGV) screenshot of
437 the genetic region mapping to foetal cartilage mQTLs colocalising with OA GWAS
438 signals rs11071366 (C), rs143384 (D), and rs3740129 (E). Light blue, foetal knee
439 ATAC-seq data with peaks representing open chromatin regions; orange, OA knee
440 ATAC-seq data; dark blue, gene transcript; bottom track is Roadmap chromatin state
441 data from E049 cultured chondrocytes. Red, active transcription start site;
442 yellow/orange, enhancer; green, transcribed; grey, repressed.

443



446 **Genetic and epigenetic effects influence distinct subsets of CpGs comprising**
447 **the foetal cartilage methylome.**

448 Finally, we tested whether the 11,299 CpGs which comprise the identified significant mQTLs
449 across the epigenome are amenable to external influences occurring between sexes (811
450 sDMPs) and through articular cartilage development (80,536 dDMPs). Between the mQTLs
451 and dDMPs, an overlap of just 744 CpGs was identified, indicating that only 6.6% of CpGs
452 that are significantly regulated by DNA sequence are impacted by developmental stage
453 ($P=5.2 \times 10^{-80}$). Similarly, between the mQTL CpGs and the 811 sDMPs, an overlap of only 6
454 CpGs (0.74%) was observed ($P= 0.0184$). This indicates that the CpGs at which DNAm is
455 heavily influenced by the genetic architecture of the region are less amenable to
456 environmental influences on the epigenome.

457

458 **Discussion**

459 In this study, we present the epigenomic trajectory across the development of human articular
460 cartilage. To our knowledge, this provides the first insight into methylomic plasticity across a
461 large window of human skeletogenesis. We identified that developmental stage is the
462 predominant determinant of methylation status within our samples, with ~8% of CpGs
463 significantly changing across the captured timeframe. We further identified >9400 dDMRs,
464 40% of which intersected with open chromatin regions in foetal knee cartilage (12pcw),
465 providing evidence for the potential functionality of the regions. We demonstrated that 811 of
466 the captured CpGs exhibited sexually dimorphic effects, with the majority (64%) being
467 hypermethylated in female samples. Finally, we conducted an epigenome-wide mQTL
468 analysis and found strong evidence of co-localizations of these molecular phenotypes with OA
469 GWAS signals, building upon existing evidence of developmental origins of a complex genetic
470 disease of older age.

471 In 2023, Zhang *et al* published the first single-cell and spatial transcriptomic atlas of
472 the developing human hindlimb³³. They provided evidence of nascent articular chondrocytes

473 that expressed high levels of *PRG4* in the distal femur by 8pcw. In our data, the overwhelming
474 impact of the developmental stage upon the cartilage methylome partly reflects the changing
475 nature of the cartilage cellular composition captured in our bulk analysis. In earlier stages
476 (<12pcw) we have likely captured epiphyseal chondrocytes along with the nascent articular
477 cartilage that will continue to develop postnatally, whereas after 12pcw, we were able to isolate
478 the nascent articular cartilage more specifically. Our gene expression profiling data provide
479 strong evidence that the primary population of cells analysed in this study are nascent
480 epiphyseal and articular chondrocytes. Underscoring this, our data also align with studies of
481 embryonic development of epiphyseal and articular cartilage in mouse models where time-
482 course expression and lineage tracing studies have shown that both *Gdf5* and *Sox9*
483 expression decreases, *Col2a1* remains relatively stable and *Prg4* increases as joint
484 development proceeds through cavitation and beyond⁴²⁻⁴⁴. In 2015, Spiers *et al.* applied the
485 Illumina HumanMethylation450 array to DNA extracted from 179 human foetal brain samples
486 between 3 and 26pcw⁴⁵. Comparable to our dataset, they also identified that DNAm at ~7% of
487 captured CpGs significantly changed across the developmental window, despite the
488 development within brain tissue of a more heterogenous cell population.

489 The largest hypermethylated dDMRs identified in this study predominantly map to
490 transcriptional regulators of early embryogenesis and skeletal patterning, including *MEIS1*,
491 *EN1*, *IRX1*, and *HOXA3*. Except for *HOXA3*, no correlations were identified between the
492 expression of any of these genes and DNAm in the identified regions. However, the expression
493 of these genes was low in all samples from 7pcw onwards and we postulate that we have
494 captured a developmental window beyond the timeframe in which these genes impart their
495 functional roles in cartilage development. Here, we appear to be observing hypermethylation
496 of DNA as these gene regions become repressed, potentially to reinforce chromatin
497 condensation. This is supported by the results of the GO analysis, which are enriched for a
498 plethora of biological process terms relating to embryogenesis and skeletal development.
499 Furthermore, the hypermethylated dDMRs were overrepresented in repressed regions of
500 mature chondrocytes (36%, compared to 30% of all dDMRs) whereas none overlapped with

501 gene transcriptional start sites. These observations also align with functional studies in murine
502 models of limb development and the recent spatial maps of gene expression in human limb
503 development³³. *Meis1* is a regulator of proximodistal identity^{46,47} and was further shown by
504 Zhang *et al* to be expressed specifically in the proximal hindlimb (developing femur) at
505 5.6pcw³³. *En1* and *Irxf1*, which orchestrate early murine limb bud patterning and digit formation,
506 respectively are most highly expressed earlier and in the most distal locations in the limb⁴⁸.

507 Conversely, several large hypomethylated regions mapped to genes encoding ECM
508 proteins including Spondin-2 and Tenascin-X, highlighting the role of the DNA methylome in
509 underlying transcriptional changes. We investigated the expression of the three genes
510 mapping to the largest significant hypomethylated dDMRs, identifying that all significantly
511 increased with decreasing DNAm. Across the developmental window, hypomethylated
512 dDMRs were enriched in CREs, particularly annotated chondrocyte enhancers (29%
513 overlapped with enhancers, compared to 25% of all dDMRs). This is consistent with earlier
514 reports using an MSC *in vitro* model that, throughout differentiation, hypomethylation of DNA
515 is enriched in enhancer regions⁴⁹.

516 Dimorphism in DNAm levels between the biological sexes has been identified in a
517 multitude of tissues across the human lifespan and is believed to contribute to the discrepancy
518 in disease trajectories between males and females⁵⁰. A recent study demonstrated that
519 developmental sexual dimorphism in DNAm is tissue-specific, with a trend for hypomethylation
520 in the placental tissues of female premature neonates, and hypermethylation in blood³⁷. Due
521 to the known discrepancies in the prevalence of knee OA in men and women, we investigated
522 sDMRs at autosomal loci to look for sex-specific differences within the developing knee joint.
523 The most significant sDMR mapped to the gene *NAB1*, encoding nerve growth factor induced
524 (NGFI)-A-binding protein 1, a transcriptional repressor⁵¹ which is also the most significant
525 DMP identified in neonatal blood. We found no significant interactions in our samples between
526 sex and developmental stage upon DNAm. Sex differences in methylation profiles could
527 potentially be reflective of different cell type populations in the same tissues, or subsets of
528 articular chondrocytes between sexes. However, it is more likely that a yet-unknown biological

529 impact is resulting in the observed sexual dimorphism, independent of cellular context, as
530 many of the identified loci have also been identified in neonatal blood³⁷.

531 Finally, we investigated the presence of epigenome-wide mQTLs within our foetal
532 cartilage samples, discovering ~400,000 significant SNV-CpG associations. Our co-
533 localisation analysis revealed a shared impact of SNVs on DNAm and OA genetic risk. This
534 confirms and builds upon our earlier report that the genetic and epigenetic interplay underlying
535 OA risk can operate within chondrocytes from the start of life²¹. The majority (73%) of the co-
536 localisations had not previously been identified in studies of aged cartilage, supporting our
537 earlier hypothesis that some epigenetic mechanisms identified in the foetal samples may
538 operate throughout the life course, whilst others may be constrained within the developmental
539 period. Our enrichment analysis also demonstrated that the CpGs which are stringently
540 regulated by SNVs in a tissue specific manner from the very start of life, are less amenable to
541 modulation by their environment, highlighting their ability to differentially regulate a target gene
542 throughout the life course. Whilst the genes mapping to the mQTL loci remain to be
543 functionally validated, our *in-silico* analysis reveals putative regulatory mechanisms.

544 We identified novel OA-mQTLs mapping to the 3'UTR (2 CpGs) and downstream (1
545 CpG) of *CHST3*, within and adjacent to open chromatin regions. *CHST3* codes for chondroitin-
546 6-O-sulfotransferase 1, an enzyme responsible for the sulfation of extracellular matrix (ECM)
547 glycosaminoglycans (GAGs)⁵². Sulfation profiles of ECM GAGs within cartilage change
548 throughout development and ageing and play a role in cell differentiation and tissue
549 morphogenesis⁵². Mutations in *CHST3* lead to skeletal dysplasia, highlighting the important
550 function of this protein in skeletogenesis⁵³. Any conferred changes to the expression of the
551 gene and its encoded protein, mediated through differential methylation, could potentially
552 confer a subtle phenotypic shift within cartilage, decreasing its integrity to withstand
553 mechanical forces over the life course.

554 We further identified OA-mQTLs in foetal cartilage that map to the genes *ALDH1A2*
555 and *GDF5*, loci at which such effects have previously been identified in aged cartilage. On
556 chromosome 15, two OA-mQTLs were identified falling within the first intron on *ALDH1A2*,

557 encoding RALDH2, integral for all-*trans* retinoic acid (atRA) synthesis. Genetic studies of hand
558 OA have been conducted on *ALDH1A2* based on the lead variant rs11071366, first identified
559 in the Icelandic population⁵⁴. Allelic expression imbalance at the locus showed a decrease in
560 gene expression driven by the OA risk allele, which was present in osteochondral tissue from
561 the trapezium, along with cartilage from the hip and knee^{55,56}. This effect was also replicated
562 in our targeted analysis of human foetal cartilage²¹. Furthermore, an OA-mQTL was identified
563 at cg12031962, where the minor and effect allele, C, correlated with a decrease in DNAm¹¹.
564 In this study, we colocalised the OA risk SNV rs11071366 (A>T)³¹, which is in moderate
565 pairwise linkage disequilibrium in European populations ($r^2=0.56$). In our analysis of the foetal
566 tissues, we also identified an OA-mQTL at the same site, where the effect allele, A, was also
567 associated with decreased DNAm. Recent studies have shown that atRA has an anti-
568 inflammatory effect within the joint, and that pharmacologically blocking its cellular metabolism
569 exhibited a chondroprotective effect⁵⁶. As an example of an epigenetic OA risk mechanism
570 occurring throughout the life course, this knowledge could influence the key therapeutic
571 window for future targeted interventions.

572 Due to the nature of the samples used in this study, it was not possible to use a
573 replication cohort to validate our findings. One further limitation to this analysis is that it has
574 not been possible to directly compare our findings to those in adult cartilage samples, nor
575 studies of foetal methylation in blood and brain tissues due to batch effects and differences in
576 the applied array technology. We hypothesise that such analyses would potentially reveal OA-
577 mQTLs which become functionally manifest in later life, as opposed to exerting
578 developmental, or lifelong functional effects.

579 In summary, we report the first comprehensive analysis of DNA methylation during
580 human skeletal development. Our findings complement existing evidence that genetic factors
581 linked to OA risk exert regulatory influences during the prenatal stage, underscoring the
582 significance of early developmental processes in joint homeostasis and disease in later life.
583 Moreover, our study showcases the effectiveness of mQTL mapping in pinpointing potential
584 causal genetic sites associated with complex diseases across complex genomic regions.

585 Future studies including both aged and developmental samples will identify potential subsets
586 of genetic risk loci which may be active across the life course, requiring earlier interventions
587 to prevent the onset and progression of disease. Our data supports our hypothesis of three
588 categories of functional OA risk: 1) loci active only during development; 2) loci active
589 throughout the life-course; 3) loci active only in mature/aged tissue. When this is known for a
590 locus/pathway, intelligent timing of a pharmacological intervention can then be undertaken,
591 and those which act in later life can be prioritised. Investigations that directly compare multiple
592 joint tissues have also recently proven insightful into the molecular basis of the disease^{8,57}.
593 Such studies undertaken throughout the life course will be vital to further elucidate the tissue-
594 specific gene regulatory networks contributing to disease, paving the way for therapeutic
595 intervention.

596

597

598 **Acknowledgements**

599 SJR is funded by The Royal Society (RGS\R1\231319), the JGW Patterson Foundation, the
600 MRC-Versus Arthritis Centre for Integrated Research into Musculoskeletal Ageing (CIMA;
601 MR/P020941/1 and MR/R502182/1), and a Versus Arthritis Career Development Fellowship
602 (22615). DR is funded by NIH K99AR078352. AKS received research fellowship funding from
603 the Wellcome Trust, Royal College of Surgeons of England, and the UK-US Fulbright
604 commission.

605

606 **Author Contributions**

607 Conceptualization, SJR; Methodology, EM, JS, AKS, SJR; Formal Analysis, EM, JS;
608 Investigation, EM, JS, SEO, MJB, NW, LM, LO, AKS, SJR; Resources, NW, LM, LO; Data
609 Curation, EM, JS, SJR; Writing- Original Draft, SJR; Writing- Reviewing and Editing; EM, JS,
610 DAY, SJR; Visualization, SEO, SJR; Supervision, JS, SJR, DAY; Project administration, SJR;
611 Funding Acquisition, SJR.

612

613 **Declaration of Interests**

614 LM has received speaker and consultancy fees from Illumina. We have no other conflicting
615 interest to declare.

616

617 **Data Availability**

618 Code used during this project are available on [https://github.com/CBFLivUni/Epigenetics-OA-](https://github.com/CBFLivUni/Epigenetics-OA-Risk-Human-Skeletal-Dev)
619 [Risk-Human-Skeletal-Dev](https://github.com/CBFLivUni/Epigenetics-OA-Risk-Human-Skeletal-Dev). Processed DNA methylation data for the 72 samples have been
620 uploaded to the Gene Expression Omnibus (GEO) and will be made available on the
621 publication of this article.

622

623 **Supplemental information**

624 Supplementary Figures: Figures S1–S5

625 Supplementary Tables: Tables S1-14

626 Supplementary Methods: Additional methodological data which is too long to fit in the

627 main body of the text. Contains the majority of the information regarding the

628 bioinformatic analyses.

629

630

- 631 1. Finnegan, M. A. & Uthoff, H. K. The Development of the Knee. *The Embryology of*
632 *the Human Locomotor System* 129–140 (1990). doi:10.1007/978-3-642-75310-7_14
- 633 2. Funari, V. A., Day, A., Krakow, D., Cohn, Z. A., Chen, Z., Nelson, S. F. & Cohn, D. H.
634 Cartilage-selective genes identified in genome-scale analysis of non-cartilage and
635 cartilage gene expression. *BMC Genomics* **8**, 1–13 (2007).
- 636 3. Geiman, T. M. & Muegge, K. DNA methylation in early development. *Mol Reprod Dev*
637 **77**, 105–113 (2010).
- 638 4. Cheng, X. DNA modification by methyltransferases. *Curr Opin Struct Biol* **5**, 4–10
639 (1995).
- 640 5. Wolf, S. F., Jolly, D. J., Lunnen, K. D., Friedmann, T. & Migeon, B. R. Methylation of
641 the hypoxanthine phosphoribosyltransferase locus on the human X chromosome:
642 implications for X-chromosome inactivation. *Proc Natl Acad Sci U S A* **81**, 2806–2810
643 (1984).
- 644 6. Schübeler, D. Function and information content of DNA methylation. *Nature* **517**, 321–
645 326 (2015).
- 646 7. Roberts, J. B. & Rice, S. J. Osteoarthritis as an Enhanceropathy: Gene Regulation in
647 Complex Musculoskeletal Disease. *Curr Rheumatol Rep* 1–13 (2024).
648 doi:10.1007/S11926-024-01142-Z/TABLES/2
- 649 8. Kreitmaier, P., Suderman, M., Southam, L., Coutinho de Almeida, R., Hatzikotoulas,
650 K., Meulenbelt, I., Steinberg, J., Relton, C. L., Wilkinson, J. M. & Zeggini, E. An
651 epigenome-wide view of osteoarthritis in primary tissues. *Am J Hum Genet* **109**,
652 1255–1271 (2022).
- 653 9. Jeffries, M. A., Donica, M., Baker, L. W., Stevenson, M. E., Annan, A. C., Beth
654 Humphrey, M., James, J. A. & Sawalha, A. H. Genome-Wide DNA Methylation Study
655 Identifies Significant Epigenomic Changes in Osteoarthritic Subchondral Bone and
656 Similarity to Overlying Cartilage. *Arthritis and Rheumatology* **68**, 1403–1414 (2016).
- 657 10. Den Hollander, W., Ramos, Y. F. M., Bos, S. D., Bomer, N., Van Der Breggen, R.,
658 Lakenberg, N., De Dijkker, W. J., Duijnisveld, B. J., Slagboom, P. E., Nelissen, R. G.
659 H. H. & Meulenbelt, I. Knee and hip articular cartilage have distinct epigenomic
660 landscapes: Implications for future cartilage regeneration approaches. *Ann Rheum*
661 *Dis* **73**, 2208–2212 (2014).
- 662 11. Rushton, M. D., Reynard, L. N., Young, D. A., Shepherd, C., Aubourg, G., Gee, F.,
663 Darlay, R., Deehan, D., Cordell, H. J. & Loughlin, J. Methylation quantitative trait locus
664 analysis of osteoarthritis links epigenetics with genetic risk. (2015).
665 doi:10.1093/hmg/ddv433
- 666 12. Prioritization of PLEC and GRINA as Osteoarthritis Risk Genes Through the
667 Identification and Characterization of Novel Methylation Quantitative Trait Loci - Rice -
668 2019 - Arthritis & Rheumatology - Wiley Online Library. at
669 <<https://onlinelibrary.wiley.com/doi/full/10.1002/art.40849>>
- 670 13. Rice, S. J., Cheung, K., Reynard, L. N. & Loughlin, J. Discovery and analysis of
671 methylation quantitative trait loci (mQTLs) mapping to novel osteoarthritis genetic risk
672 signals. *Osteoarthritis Cartilage* **27**, 1545–1556 (2019).
- 673 14. Rice, S. J., Beier, F., Young, D. A. & Loughlin, J. Interplay between genetics and
674 epigenetics in osteoarthritis. *Nat Rev Rheumatol* 1–14 (2020). doi:10.1038/s41584-
675 020-0407-3
- 676 15. Kehayova, Y. S., Wilkinson, J. M., Rice, S. J. & Loughlin, J. Osteoarthritis genetic risk
677 acting on the galactosyltransferase gene COLGALT2 has opposing functional effects
678 in articulating joint tissues. *Arthritis Res Ther* **25**, 83 (2023).
- 679 16. Rice, S. J., Roberts, J. B., Tselepi, M., Brumwell, A., Falk, J., Steven, C. & Loughlin,
680 J. Genetic and Epigenetic Fine-Tuning of TGFB1 Expression Within the Human
681 Osteoarthritic Joint. *Arthritis Rheumatol* **73**, 1866–1877 (2021).
- 682 17. Pitsillides, A. A. & Beier, F. Cartilage biology in osteoarthritis - Lessons from
683 developmental biology. *Nat Rev Rheumatol* **7**, 654–663 Preprint at
684 <https://doi.org/10.1038/nrrrheum.2011.129> (2011)

- 685 18. Frysz, M., Faber, B. G., Ebsim, R., Saunders, F. R., Lindner, C., Gregory, J. S.,
686 Aspden, R. M., Harvey, N. C., Cootes, T. & Tobias, J. H. Machine Learning–Derived
687 Acetabular Dysplasia and Cam Morphology Are Features of Severe Hip
688 Osteoarthritis: Findings From UK Biobank. *Journal of Bone and Mineral Research* **37**,
689 1720–1732 (2022).
- 690 19. Faber, B. G., Baird, D., Gregson, C. L., Gregory, J. S., Barr, R. J., Aspden, R. M.,
691 Lynch, J., Nevitt, M. C., Lane, N. E., Orwoll, E. & Tobias, J. H. DXA-derived hip shape
692 is related to osteoarthritis: findings from in the MrOS cohort. *Osteoarthritis Cartilage*
693 **25**, 2031–2038 (2017).
- 694 20. Baird, D. A., Evans, D. S., Kamanu, F. K., Gregory, J. S., Saunders, F. R., Giuraniuc,
695 C. V., Barr, R. J., Aspden, R. M., Jenkins, D., Kiel, D. P., Orwoll, E. S., Cummings, S.
696 R., Lane, N. E., Mullin, B. H., Williams, F. M. K., Richards, J. B., Wilson, S. G.,
697 Spector, T. D., Faber, B. G., Lawlor, D. A., Grundberg, E., Ohlsson, C., Pettersson-
698 Kymmer, U., Capellini, T. D., Richard, D., Beck, T. J., Evans, D. M., Paternoster, L.,
699 Karasik, D. & Tobias, J. H. Identification of Novel Loci Associated With Hip Shape: A
700 Meta-Analysis of Genomewide Association Studies. *Journal of Bone and Mineral*
701 *Research* **34**, 241–251 (2019).
- 702 21. Rice, S. J., Brumwell, A., Falk, J., Kehayova, Y. S., Casement, J., Parker, E., Hofer, I.
703 M. J., Shepherd, C. & Loughlin, J. Genetic risk of osteoarthritis operates during
704 human skeletogenesis. *Hum Mol Genet* (2022). doi:10.1093/HMG/DDAC251
- 705 22. Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase
706 gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
- 707 23. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and
708 Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**,
709 247–257 (1999).
- 710 24. Shen, J., Wang, C., Li, D., Xu, T., Myers, J., Ashton, J. M., Wang, T., Zuscik, M. J.,
711 McAlinden, A. & O’Keefe, R. J. DNA methyltransferase 3b regulates articular cartilage
712 homeostasis by altering metabolism. *JCI Insight* **2**, (2017).
- 713 25. Sarkar, A., Liu, N. Q., Magallanes, J., Tassej, J., Lee, S., Shkhyan, R., Lee, Y., Lu, J.,
714 Ouyang, Y., Tang, H., Bian, F., Tao, L., Segil, N., Ernst, J., Lyons, K., Horvath, S. &
715 Evseenko, D. STAT3 promotes a youthful epigenetic state in articular chondrocytes.
716 *Aging Cell* **22**, (2023).
- 717 26. Pidsley, R., Zotenko, E., Peters, T. J., Lawrence, M. G., Risbridger, G. P., Molloy, P.,
718 Van Dijk, S., Muhlhäusler, B., Stirzaker, C. & Clark, S. J. Critical evaluation of the
719 Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation
720 profiling. *Genome Biol* **17**, 1–17 (2016).
- 721 27. McCartney, D. L., Walker, R. M., Morris, S. W., McIntosh, A. M., Porteous, D. J. &
722 Evans, K. L. Identification of polymorphic and off-target probe binding sites on the
723 Illumina Infinium MethylationEPIC BeadChip. *Genom Data* **9**, 22–24 (2016).
- 724 28. Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P.,
725 Hansen, K. D. & Irizarry, R. A. Minfi: a flexible and comprehensive Bioconductor
726 package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**,
727 1363–1369 (2014).
- 728 29. Shabalin, A. A. Matrix eQTL: ultra fast eQTL analysis via large matrix operations.
729 *Bioinformatics* **28**, 1353–1358 (2012).
- 730 30. Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I.
731 M., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Bennett, R., Berry, A., Bhai, J.,
732 Bignell, A., Billis, K., Boddu, S., Brooks, L., Charkhchi, M., Cummins, C., Da Rin
733 Fioretto, L., Davidson, C., Dodiya, K., Donaldson, S., El Houdaigui, B., El Naboulsi,
734 T., Fatima, R., Giron, C. G., Genez, T., Martinez, J. G., Gujjarro-Clarke, C., Gymer,
735 A., Hardy, M., Hollis, Z., Hourlier, T., Hunt, T., Juettemann, T., Kaikala, V., Kay, M.,
736 Lavidas, I., Le, T., Lemos, D., Marugán, J. C., Mohanan, S., Mushtaq, A., Naven, M.,
737 Ogeh, D. N., Parker, A., Parton, A., Perry, M., Pilizota, I., Prosovetskaia, I., Sakthivel,
738 M. P., Salam, A. I. A., Schmitt, B. M., Schuilenburg, H., Sheppard, D., Perez-Silva, J.
739 G., Stark, W., Steed, E., Sutinen, K., Sukumaran, R., Sumathipala, D., Suner, M. M.,

- 740 Szpak, M., Thormann, A., Tricomi, F. F., Urbina-Gómez, D., Veidenberg, A., Walsh, T.
741 A., Walts, B., Willhoft, N., Winterbottom, A., Wass, E., Chakiachvili, M., Flint, B.,
742 Frankish, A., Giorgetti, S., Haggerty, L., Hunt, S. E., Ilesley, G. R., Loveland, J. E.,
743 Martin, F. J., Moore, B., Mudge, J. M., Muffato, M., Perry, E., Ruffier, M., Tate, J.,
744 Thybert, D., Trevanion, S. J., Dyer, S., Harrison, P. W., Howe, K. L., Yates, A. D.,
745 Zerbino, D. R. & Flicek, P. Ensembl 2022. *Nucleic Acids Res* **50**, D988–D995 (2022).
746 31. Boer, C. G., Hatzikotoulas, K., Southam, L., Stefánsdóttir, L., Zhang, Y., Coutinho de
747 Almeida, R., Wu, T. T., Zheng, J., Hartley, A., Teder-Laving, M., Skogholt, A. H.,
748 Terao, C., Zengini, E., Alexiadis, G., Barysenka, A., Bjornsdottir, G., Gabrielsen, M.
749 E., Gilly, A., Ingvarsson, T., Johnsen, M. B., Jonsson, H., Kloppenburg, M., Luetge,
750 A., Lund, S. H., Mägi, R., Mangino, M., Nelissen, R. R. G. H. H., Shivakumar, M.,
751 Steinberg, J., Takuwa, H., Thomas, L. F., Tuerlings, M., Loughlin, J., Arden, N.,
752 Birrell, F., Carr, A., Deloukas, P., Doherty, M., McCaskie, A. W., Ollier, W. E. R., Rai,
753 A., Ralston, S. H., Spector, T. D., Wallis, G. A., Martinsen, A. E., Willer, C., Fors, E.
754 A., Mundal, I., Hagen, K., Nilsen, K. B., Lie, M. U., Børte, S., Brumpton, B., Nielsen, J.
755 B., Fritsche, L. G., Zhou, W., Heuch, I., Storheim, K., Tyrpenou, E., Koukakis, A.,
756 Chytas, D., Evangelopoulos, D. S., Efstathios, C., Pneumaticsos, S., Nikolaou, V. S.,
757 Malizos, K., Anastasopoulou, L., Abecasis, G., Baras, A., Cantor, M., Coppola, G.,
758 Deubler, A., Economides, A., Lotta, L. A., Overton, J. D., Reid, J. G., Shuldiner, A.,
759 Karalis, K., Siminovitch, K., Beechert, C., Forsythe, C., Fuller, E. D., Gu, Z., Lattari,
760 M., Lopez, A., Schleicher, T. D., Padilla, M. S., Widom, L., Wolf, S. E., Pradhan, M.,
761 Manoochchri, K., Bai, X., Balasubramanian, S., Boutkov, B., Eom, G., Habegger, L.,
762 Hawes, A., Krasheninina, O., Lanche, R., Mansfield, A. J., Maxwell, E. K., Nafde, M.,
763 O’Keeffe, S., Orelus, M., Panea, R., Polanco, T., Rasool, A., Salerno, W., Staples, J.
764 C., Li, D., Sharma, D., Banerjee, I., Bovijn, J., Locke, A., Verweij, N., Haas, M., Hindy,
765 G., De, T., Akbari, P., Sosina, O., Ferreira, M. A. R., Jones, M. B., Mighty, J., LeBlanc,
766 M. G., Mitnau, L. J., Babis, G. C., Cheung, J. P. Y., Kang, J. H., Kraft, P., Lietman, S.
767 A., Samartzis, D., Slagboom, P. E., Stefansson, K., Thorsteinsdottir, U., Tobias, J. H.,
768 Uitterlinden, A. G., Winsvold, B., Zwart, J. A., Davey Smith, G., Sham, P. C.,
769 Thorleifsson, G., Gaunt, T. R., Morris, A. P., Valdes, A. M., Tsezou, A., Cheah, K. S.
770 E., Ikegawa, S., Hveem, K., Esko, T., Wilkinson, J. M., Meulenbelt, I., Lee, M. T. M.,
771 van Meurs, J. B. J., Styrkársdóttir, U. & Zeggini, E. Deciphering osteoarthritis genetics
772 across 826,690 individuals from 9 populations. *Cell* **184**, 4784–4818.e17 (2021).
773 32. Wang, G., Sarkar, A., Carbonetto, P. & Stephens, M. A Simple New Approach to
774 Variable Selection in Regression, with Application to Genetic Fine Mapping. *J R Stat*
775 *Soc Series B Stat Methodol* **82**, 1273–1300 (2020).
776 33. Zhang, B., He, P., Lawrence, J. E. G., Wang, S., Tuck, E., Williams, B. A., Roberts,
777 K., Kleshchevnikov, V., Mamanova, L., Bolt, L., Polanski, K., Li, T., Elmentaite, R.,
778 Fasouli, E. S., Prete, M., He, X., Yayon, N., Fu, Y., Yang, H., Liang, C., Zhang, H.,
779 Blain, R., Chedotal, A., FitzPatrick, D. R., Firth, H., Dean, A., Bayraktar, O. A.,
780 Marioni, J. C., Barker, R. A., Storer, M. A., Wold, B. J., Zhang, H. & Teichmann, S. A.
781 A human embryonic limb cell atlas resolved in space and time. *Nature* **2023** 1–11
782 (2023). doi:10.1038/s41586-023-06806-x
783 34. Zhang, F., Wang, Y., Zhao, Y., Wang, M., Zhou, B., Zhou, B. & Ge, X. NFATc1 marks
784 articular cartilage progenitors and negatively determines articular chondrocyte
785 differentiation. *Elife* **12**, (2023).
786 35. Zhang, C. H., Gao, Y., Hung, H. H., Zhuo, Z., Grodzinsky, A. J. & Lassar, A. B. Creb5
787 coordinates synovial joint formation with the genesis of articular cartilage. *Nature*
788 *Communications* **2022** *13*:1 **13**, 1–18 (2022).
789 36. Tavares, A. T., Tsukui, T. & Belmonte, J. C. I. Evidence that members of the
790 Cut/Cux/CDP family may be involved in AER positioning and polarizing activity during
791 chick limb development. *Development* **127**, 5133–5144 (2000).
792 37. Santos, H. P., Enggasser, A. E., Clark, J., Roell, K., Zhabotynsky, V., Gower, W. A.,
793 Yanni, D., Yang, N. G., Washburn, L., Gogcu, S., Marsit, C. J., Kuban, K., O’Shea, T.

- 794 M. & Fry, R. C. Sexually dimorphic methylation patterns characterize the placenta and
795 blood from extremely preterm newborns. *BMC Biol* **21**, 1–12 (2023).
- 796 38. Cotney, J., Leng, J., Yin, J., Reilly, S. K., Demare, L. E., Emera, D., Ayoub, A. E.,
797 Rakic, P. & Noonan, J. P. The evolution of lineage-specific regulatory activities in the
798 human embryonic limb. *Cell* **154**, 185 (2013).
- 799 39. Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky,
800 M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M. J.,
801 Amin, V., Whitaker, J. W., Schultz, M. D., Ward, L. D., Sarkar, A., Quon, G.,
802 Sandstrom, R. S., Eaton, M. L., Wu, Y. C., Pfening, A. R., Wang, X., Claussnitzer,
803 M., Liu, Y., Coarfa, C., Harris, R. A., Shores, N., Epstein, C. B., Gjoneska, E., Leung,
804 D., Xie, W., Hawkins, R. D., Lister, R., Hong, C., Gascard, P., Mungall, A. J., Moore,
805 R., Chuah, E., Tam, A., Canfield, T. K., Hansen, R. S., Kaul, R., Sabo, P. J., Bansal,
806 M. S., Carles, A., Dixon, J. R., Farh, K. H., Feizi, S., Karlic, R., Kim, A. R., Kulkarni,
807 A., Li, D., Lowdon, R., Elliott, G., Mercer, T. R., Neph, S. J., Onuchic, V., Polak, P.,
808 Rajagopal, N., Ray, P., Sallari, R. C., Siebenthall, K. T., Sinnott-Armstrong, N. A.,
809 Stevens, M., Thurman, R. E., Wu, J., Zhang, B., Zhou, X., Beaudet, A. E., Boyer, L.
810 A., De Jager, P. L., Farnham, P. J., Fisher, S. J., Haussler, D., Jones, S. J. M., Li, W.,
811 Marra, M. A., McManus, M. T., Sunyaev, S., Thomson, J. A., Tlsty, T. D., Tsai, L. H.,
812 Wang, W., Waterland, R. A., Zhang, M. Q., Chadwick, L. H., Bernstein, B. E.,
813 Costello, J. F., Ecker, J. R., Hirst, M., Meissner, A., Milosavljevic, A., Ren, B.,
814 Stamatoyannopoulos, J. A., Wang, T. & Kellis, M. Integrative analysis of 111
815 reference human epigenomes. *Nature* **518**, 317–329 (2015).
- 816 40. Chen, H., Capellini, T. D., Schoor, M., Mortlock, D. P., Reddi, A. H. & Kingsley, D. M.
817 Heads, Shoulders, Elbows, Knees, and Toes: Modular Gdf5 Enhancers Control
818 Different Joints in the Vertebrate Skeleton. *PLoS Genet* **12**, (2016).
- 819 41. Capellini, T. D., Chen, H., Cao, J., Doxey, A. C., Kiapour, A. M., Schoor, M. &
820 Kingsley, D. M. Ancient selection for derived alleles at a GDF5 enhancer influencing
821 human growth and osteoarthritis risk. *Nature Genetics* **2017 49:8 49**, 1202–1210
822 (2017).
- 823 42. Rux, D., Decker, R. S., Koyama, E. & Pacifici, M. Joints in the appendicular skeleton:
824 Developmental mechanisms and evolutionary influences. *Curr Top Dev Biol* **133**,
825 119–151 (2019).
- 826 43. Shwartz, Y., Viukov, S., Krief, S. & Zelzer, E. Joint Development Involves a
827 Continuous Influx of Gdf5-Positive Cells. *Cell Rep* **15**, 2577–2587 (2016).
- 828 44. Kim, M., Koyama, E., Saunders, C. M., Querido, W., Pleshko, N. & Pacifici, M.
829 Synovial joint cavitation initiates with microcavities in interzone and is coupled to
830 skeletal flexion and elongation in developing mouse embryo limbs. *Biol Open* **11**,
831 (2022).
- 832 45. Spiers, H., Hannon, E., Schalkwyk, L. C., Smith, R., Wong, C. C. Y., O'Donovan, M.
833 C., Bray, N. J. & Mill, J. Methylomic trajectories across human fetal brain
834 development. *Genome Res* **25**, 338 (2015).
- 835 46. Mercader, N., Leonardo, E., Azplazu, N., Serrano, A., Morata, G., Martínez-A, C. &
836 Torres, M. Conserved regulation of proximodistal limb axis development by
837 Meis1/Hth. *Nature* **1999 402:6760 402**, 425–429 (1999).
- 838 47. Delgado, I., Giovinazzo, G., Temiño, S., Gauthier, Y., Balsalobre, A., Drouin, J. &
839 Torres, M. Control of mouse limb initiation and antero-posterior patterning by Meis
840 transcription factors. *Nat Commun* **12**, (2021).
- 841 48. Loomis, C. A., Harris, E., Michaud, J., Wurst, W., Hanks, M. & Joyner, A. L. The
842 mouse Engrailed-1 gene and ventral limb patterning. *Nature* **382**, 360–363 (1996).
- 843 49. Barter, M. J., Bui, C., Cheung, K., Falk, J., Gómez, R., Skelton, A. J., Elliott, H. R.,
844 Reynard, L. N. & Young, D. A. DNA hypomethylation during MSC chondrogenesis
845 occurs predominantly at enhancer regions. *Sci Rep* **10**, (2020).
- 846 50. Martin, E., Smeester, L., Bommarito, P. A., Grace, M. R., Boggess, K., Kuban, K.,
847 Karagas, M. R., Marsit, C. J., O'Shea, T. M. & Fry, R. C. Sexual epigenetic

- 848 dimorphism in the human placenta: implications for susceptibility during the prenatal
849 period. <http://dx.doi.org/10.2217/epi-2016-0132> **9**, 267–278 (2017).
- 850 51. Thiel, G., Kaufmann, K., Magin, A., Lietz, M., Bach, K. & Cramer, M. The human
851 transcriptional repressor protein NAB1: expression and biological activity. *Biochimica*
852 *et Biophysica Acta (BBA) - Gene Structure and Expression* **1493**, 289–301 (2000).
- 853 52. Kitagawa, H., Tsutsumi, K., Tone, Y. & Sugahara, K. Developmental regulation of the
854 sulfation profile of chondroitin sulfate chains in the chicken embryo brain. *J Biol Chem*
855 **272**, 31377–31381 (1997).
- 856 53. Thiele, H., Sakano, M., Kitagawa, H., Sugahara, K., Rajab, A., Höhne, W., Ritter, H.,
857 Leschik, G., Nürnberg, P. & Mundlos, S. Loss of chondroitin 6-O-sulfotransferase-1
858 function results in severe human chondrodysplasia with progressive spinal
859 involvement. *Proc Natl Acad Sci U S A* **101**, 10155–10160 (2004).
- 860 54. Styrkarsdottir, U., Thorleifsson, G., Helgadóttir, H. T., Bomer, N., Metrustry, S.,
861 Bierma-Zeinstra, S., Strijbosch, A. M., Evangelou, E., Hart, D., Beekman, M.,
862 Jonasdóttir, A., Sigurdsson, A., Eiriksson, F. F., Thorsteinsdóttir, M., Frigge, M. L.,
863 Kong, A., Gudjonsson, S. A., Magnússon, O. T., Masson, G., Hofman, A., Arden, N.
864 K., Ingvarsson, T., Lohmander, S., Kloppenburg, M., Rivadeneira, F., Nelissen, R. G.
865 H. H., Spector, T., Uitterlinden, A., Slagboom, P. E., Thorsteinsdóttir, U., Jonsdóttir, I.,
866 Valdes, A. M., Meulenbelt, I., Van Meurs, J., Jonsson, H. & Stefansson, K. Severe
867 osteoarthritis of the hand associates with common variants within the ALDH1A2 gene
868 and with rare variants at 1p31. *Nat Genet* **46**, 498–502 (2014).
- 869 55. Shepherd, C., Zhu, D., Skelton, A. J., Combe, J., Threadgold, H., Zhu, L., Vincent, T.
870 L., Stuart, P., Reynard, L. N. & Loughlin, J. Functional Characterization of the
871 Osteoarthritis Genetic Risk Residing at ALDH1A2 Identifies rs12915901 as a Key
872 Target Variant. *Arthritis and Rheumatology* **70**, 1577–1587 (2018).
- 873 56. Zhu, L., Kamalathavan, P., Koneva, L. A., Zarebska, J. M., Chanalaris, A., Ismail, H.,
874 Wiberg, A., Ng, M., Muhammad, H., Walsby-Tickle, J., McCullagh, J. S. O., Watt, F.
875 E., Sansom, S. N., Furniss, D., Gardiner, M. D. & Vincent, T. L. Variants in ALDH1A2
876 reveal an anti-inflammatory role for retinoic acid and a new class of disease-modifying
877 drugs in osteoarthritis. *Sci Transl Med* **14**, (2022).
- 878 57. Kreitmaier, P., Park, Y.-C., Swift, D., Gilly, A., Wilkinson, J. M., Zeggini, E. &
879 Wilkinson, † J Mark. Epigenomic profiling of the infrapatellar fat pad in osteoarthritis.
880 *Hum Mol Genet* (2023). doi:10.1093/HMG/DDAD198
- 881 58. Aubourg, G., Rice, S. J., Bruce-Wootton, P. & Loughlin, J. Genetics of osteoarthritis.
882 *Osteoarthritis Cartilage* (2021). doi:10.1016/j.joca.2021.03.002
- 883

885

<i>Top hypermethylated DMRs</i>							
DMR ID (Cluster)	Chr	Start	End	Gene	Classification/Function	FDR	CpGs (N)
1413 (3)	2	66665027	66668012	<i>MEIS1</i>	Transcriptional regulator of proximodistal limb identity	2.03x10 ⁻³⁶	18
1576 (3)	2	119607192	119611880	<i>EN1</i>	Engrailed Homeobox Transcription Factor	2.26x10 ⁻⁴⁴	19
1755 (3)	2	177021702	177030228	<i>HOXD3</i>	Homeobox Transcription Factor	6.31x10 ⁻¹²⁷	39
3054 (3)	5	3597311	3602413	<i>IRX1</i>	Iroquois Homeobox Transcription Factor	1.98x10 ⁻¹⁸	18
3759 (3)	6	30848726	30851086	<i>DDR1</i>	Discoidin Domain Receptor	2.01x10 ⁻⁹⁷	18
3825 (3)	6	32115979	32118204	<i>PRRT1</i>	Transmembrane Protein	1.43x10 ⁻⁶⁹	21
4430 (3)	7	27142100	27143788	<i>HOXA2</i>	Homeobox Transcription Factor	2.76x10 ⁻²²	20
4444 (3)	7	27181067	27189610	<i>HOXA3</i>	Homeobox Transcription Factor	1.48x10 ⁻¹⁴⁰	59
6992 (3)	12	16760056	16763480	<i>MGST1</i>	Glutathione S-Transferase	1.49x10 ⁻¹⁰⁴	19
8319 (3)	15	74423308	74427499	<i>ISLR2</i>	Immunoglobulin Superfamily	8.15x10 ⁻⁷⁶	19
9322 (3)	17	48911141	48913317	<i>WFIKKN2</i>	Serine and metalloprotease inhibitor	8.44x10 ⁻²⁶⁰	22
10141 (3)	20	36145902	36148775	<i>BLCAP</i>	Regulates cell proliferation and apoptosis	1.86x10 ⁻¹¹²	24
10351 (3)	21	38076709	38082613	<i>SIM2</i>	Basic helix-loop-helix transcription factor	7.42x10 ⁻³⁷	21
<i>Top hypomethylated DMRs</i>							
DMR ID (Cluster)	Chr	Start	End	Gene	Classification/Function	FDR	CpGs (N)
2153 (2)	3	46447254	46449636	<i>CCRL2</i>	Chemokine receptor	3.11x10 ⁻¹²⁷	17
2624 (2)	4	1202780	1206150	<i>SPON2</i>	ECM Glycoprotein	3.11x10 ⁻¹⁰⁵	17
3804 (2)	6	32014059	32016535	<i>TNXB</i>	ECM Glycoprotein	3.80x10 ⁻¹¹⁷	34
3813 (2)	6	32048632	32050947	<i>TNXB</i>	ECM Glycoprotein	1.04x10 ⁻⁶⁷	19

886

887 **Table 1. Top Developmental DMRs which intersect with open chromatin regions**
 888 **in developing knee cartilage.** The largest identified dDMRs (N CpGs>16) which
 889 became either hypermethylated (upper) or hypomethylated (lower) across the
 890 captured developmental window. The nearest physical gene to each dDMR is listed.

891

892

893

CpG ID	Chr	Position	FDR-adjusted P-value	Nearest Gene Symbol	Distance (bp)	Cartilage highest mean B	Blood highest mean B
cg03618918	1	160865097	1.59x10 ⁻²¹	<i>ITLN1</i>	10136	M	M
cg26919182	1	202522232	2.41x10 ⁻²²	<i>PPP1R12B</i>	0	F	F
cg15817705	1	209406063	2.86x10 ⁻²³	<i>CAMK1G</i>	350998	M	M
cg25742246	1	214153460	2.22x10 ⁻¹⁹	<i>PROX1</i>	3063	F	F
cg12691488	1	243053673	1.27x10 ⁻³⁹	<i>CEP170</i>	234056	M	M
cg15228509	1	243074717	2.12x10 ⁻⁴²	<i>CEP170</i>	213012	F	F
cg02989351	2	9770584	1.16x10 ⁻²³	<i>YWHAQ</i>	0	F	F
cg03226871	2	128128958	2.88x10 ⁻²⁰	<i>MAP3K2</i>	0	M	M
cg19765154	2	191524409	7.70x10 ⁻³⁶	<i>NAB1</i>	0	M	M
cg20262915	2	191524489	2.02x10 ⁻²²	<i>NAB1</i>	0	M	M
cg00148935	3	16398839	1.19x10 ⁻²⁷	<i>RFTN1</i>	0	F	F
cg07850329	9	33265029	3.04x10 ⁻²¹	<i>CHMP5</i>	0	F	F
cg07852945	9	84303915	9.64x10 ⁻²⁰	<i>TLE1</i>	0	F	F
cg02716779	9	131016143	3.27x10 ⁻²²	<i>DNM1</i>	0	F	F
cg26355737	13	114292172	2.84x10 ⁻²¹	<i>TFDP1</i>	0	M	M
cg11284736	15	83826708	3.77x10 ⁻²¹	<i>HDGFRP3</i>	0	M	M
cg03626220	16	71052200	1.47x10 ⁻²⁴	<i>HYDIN</i>	0	M	M
cg03218192	17	33914403	4.68x10 ⁻²²	<i>AP2B1</i>	0	F	F
cg12607525	17	42286849	1.94x10 ⁻²²	<i>UBTF</i>	0	M	M
cg17612569	21	27107221	2.38x10 ⁻²⁵	<i>ATP5J</i>	0	M	M

894

895 **Table 2. Top 20 sDMPs in developing articular cartilage.** The most significant
896 DMPs between male and female cartilage samples are listed. Chr, chromosome;
897 FDR, false discovery rate; B, beta value; M, male; F, female.

Chr	SNV	Gene (SNP)	mQTL CpG	mQTL P-value	H4 PP	Effect on DNAm	EA	OA mQTL Locus	Same CpG?	H3K27ac Foetal Limb	ATAC peak Foetal?	ATAC peak OA?	Chon	Evidence of function?
2	rs74676797	<i>TMEM18</i>	cg14382829	6.7x10 ⁻⁹	0.84	+	A				N	N		D
3	rs3774354	<i>ITIH1</i>	cg06068388	0.04	0.98	+								L
			cg15147215	2.5x10 ⁻⁷	0.89	-		Y ¹¹	N					
			cg27294008	4.7x10 ⁻⁸	0.87	+		Y ⁵⁸	N					
			cg21742876	5.1x10 ⁻⁷	0.87	+	A	Y	N	N	N			
			cg25917518	4.7x10 ⁻¹¹	0.93	+			N	N				
			cg06706579	8.2x10 ⁻¹¹	0.94	+			N	N	N			
			cg18404041	6.0x10 ⁻⁸	0.92	-		Y ⁵⁸	N	N	N			
4	rs1530586	<i>TACC3</i>	cg18815463	1.7x10 ⁻⁶	0.89	+	T	Y ^{8,13}		N	N	N		
4	rs11729628	<i>PRDM5</i>	cg07870920	1.2x10 ⁻³⁵	0.86	-	T			N	N	N		
5	rs10062749	<i>FGF1</i>	cg15672022	6.8x10 ⁻⁵	0.81	-	T							L
6	rs9396861	<i>RNF144B</i>	cg04025088	1.6x10 ⁻⁴	0.95	-	A					N		D
6	rs79220007	<i>HFE</i>	cg00854024	1.8x10 ⁻¹⁴	1.00	+	T				N	N		D
6	rs2856821	<i>HLA-DPA1</i>	cg17362900	2.4x10 ⁻⁸	0.97	+				N				L
			cg19053046	4.0x10 ⁻⁶	0.93	+								
			cg25045942	1.3x10 ⁻⁶	0.98	+	T							
			cg13921245	9.0x10 ⁻⁷	0.84	+		Y ^{8,13,58}	N	N	N			
			cg14310225	0.004	0.87	+			N	N	N			
7	rs116934101	<i>CUX1</i>	cg23686983	1.3x10 ⁻⁵	0.94	-	A			N	N	N		
9	rs1330349	<i>TNC</i>	cg12835026	0.005	0.88	-	A			N	N	N		
10	rs3740129	<i>CHST3</i>	cg15093997	1.6x10 ⁻¹²	0.97	-					N			L
			cg20048045	1.9x10 ⁻⁷	0.98	-	A				N	N		
			cg17209404	3.9x10 ⁻⁴	0.90	+						N		
11	rs3993110	<i>TEAD1</i>	cg25843174	1.3x10 ⁻¹¹	0.98	+	A				N	N		D
11	rs1631174	<i>PTPRJ</i>	cg11545521	1.5x10 ⁻⁵	0.80	-	A			N	N	N		
11	rs67924081	<i>EHBP1L1</i>	cg01366692	2.1x10 ⁻⁷	0.98	-	A			N	N	N		L
			cg17120908	6.3x10 ⁻¹⁸	0.99	-								

11	rs1149620	<i>TSKU</i>	cg17107561	0.002	0.81	-	A		N	N	N			
15	rs11071366	<i>ALDH1A2</i>	cg12031962	7.4×10^{-13}	0.99	-	A	Y	Y ^{11,58}	N	N	N		L
			cg12382153	1.3×10^{-4}	0.90	+			Y ⁸		N			
15	rs12908498	<i>SMAD3</i>	cg15164144	8.4×10^{-6}	0.90	+				N		N		D
			cg21106065	3.2×10^{-17}	0.99	-	C			N		N		
16	rs34195470	<i>WWP2</i>	cg00854324	0.015	0.84	+	A	Y ¹³		N	N	N		
17	rs216175	<i>SMG6</i>	cg04375931	0.046	0.99	+	A	Y ⁸						L
17	rs9908159	<i>C17orf67</i>	cg24142633	7.5×10^{-5}	0.88	+	T			N				L
17	rs7212908	<i>NACA2</i>	cg03721816	0.006	0.90	+	A			N		N		D
19	rs10405617	<i>SLC44A2</i>	cg01654627	8.5×10^{-5}	0.83	-	A	Y ⁸		N		N		D
			cg16876852	4.5×10^{-8}	0.95	+				N	N	N		
19	rs8112559	<i>IRF2BP1</i>	cg16791036	2.7×10^{-6}	0.95	+	C				N	N		D
20	rs143384	<i>GDF5</i>	cg12548730	9.7×10^{-6}	0.92	-	A	Y ¹¹		N		N		D
21	rs9981884	<i>BRWD1</i>	cg11644478	1.2×10^{-5}	0.84	+								L
			cg05821552	4.3×10^{-8}	0.96	-	A			N	N	N		
			cg14227474	1.7×10^{-5}	0.91	-				N	N	N		
22	rs12160491	<i>H1FO</i>	cg17109681	1.2×10^{-6}	0.92	+						N		L
			cg06521852	1.1×10^{-8}	0.97	+					N	N		
			cg11795195	2.8×10^{-6}	0.97	+	A				N	N		
			cg25404088	2.8×10^{-7}	0.95	+								
			cg13303067	2.9×10^{-4}	0.86	+								
			cg26028571	8.0×10^{-5}	0.91	+				N	N			

898

899 **Table 3. OA genetic risk loci co-localizing with significant mQTLs in human foetal knee cartilage.** The nearest gene to the SNV is listed.
900 MA, minor allele; EA, effect allele H4 PP, poster probability that both traits are associated and share a single causal variant. We investigated
901 whether the CpGs fell within or adjacent to (within 150bp of) narrow peaks in our cartilage ATAC-seq data sets (GSE214394) or within regions
902 enriched for H3K27ac (a marker of enhancer activity) in human limb tissues at embryonic day 47 (CS19; Gene Expression Omnibus accession
903 GSE42413). In human cultured chondrocytes (E049) the 15 ROADMAP states were collapsed into 5 functional classifications: Yellow: enhancer
904 (Enh, EnhG, EnhBiv), White: repressed (ZNF/Rpts, Quies), Grey: heterochromatic (ReprPC, ReprPCWk), Red: promoters (TssBiv, TssA,
905 TssAFlnk, BivFlnk) and Green: transcribed (TxWk, Tx, TxFlnk). Regions were overlapped with the CpG locations and are indicated by the denoted
906 colours. L, lifecourse; D, development