Epigenetic mechanisms of osteoarthritis risk in human skeletal development

Euan McDonnell[1], Sarah E Orr[2], Matthew J Barter[2], Danielle Rux[3], Abby Brumwell[2], Nicola Wrobel[4], Lee Murphy[4], Lynne M Overmann[5], Antony K Sorial[2], David A Young[2], Jamie Soul[7]*, Sarah J Rice[2]*

1. Computational Biology Facility, University of Liverpool, MerseyBio, Crown Street, United Kingdom
2. Biosciences Institute, Newcastle University, Central Parkway, Newcastle upon Tyne, United Kingdom
3. Orthopaedic Surgery, UConn Health, Farmington, Connecticut, USA
4. Edinburgh Clinical Research Facility, University of Edinburgh, Edinburgh, United Kingdom.
5. Human Developmental Biology Resource, Newcastle University, International Centre for Life, Central Parkway, Newcastle upon Tyne, United Kingdom.
6. Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, United Kingdom.

*These authors made an equal contribution

Address correspondence to:

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

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

Word Count: 5199

Running Title: Articular cartilage methylome in human development

Abstract

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

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

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

Introduction

The development of human limbs and synovial joints commences between four and eight weeks post conception (pcw). During limb bud outgrowth, mesenchymal condensations form the cartilage anlagen (driven by the transcriptional master-regulator, SOX9) to establish the rudimentary embryonic skeletal elements. The limb synovial joints are established through coordinated dedifferentiation of the nascent cartilage anlagen at the presumptive joint locations to form a unique pool of progenitors at each location known as the interzone. From these progenitors, all structures of the mature synovial joint are established including articular cartilage, synovial lining, meniscus and intrajoint ligaments. For the knee joint, chondrogenesis within the femoral condyles is advanced by the start of seven pcw (Carnegie Stage (CS) 19), and a clear interzone has formed between the femur and the tibia, with full cavitation of the joint and formation of the synovial cavity commencing at eight pcw[1]. The first appearance of menisci and ligaments also occurs within a similar developmental window, whilst the formation of the infrapatellar fat pad does not appear until much later in development (18pcw)[1]. This dynamic process of limb and synovial joint development requires the orchestrated expression genes essential for the development of defined and differentiated tissues[2]. The transcriptome is primarily regulated through the spatiotemporal expression of transcription factors (TFs), yet epigenetic processes, including DNA methylation (DNAm) both underlie and reinforce transitional plasticity during development[3].

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

of repressive methyl-binding proteins. However, the relationship between DNAm and gene expression is far from straightforward, with gene body methylation often being associated with active transcription[5]. It is generally considered that DNAm within *cis*-regulatory elements (CREs) is repressive to the expression of the gene target[6,7].

The methylome of human articular cartilage has been extensively studied, primarily in the context of the chronic joint disease osteoarthritis (OA). OA is a global leading cause of disability amongst older adults, hallmarked by the degradation of articular cartilage in the joints, most commonly the hip or knee. Studies of the aged cartilage methylome have revealed distinct epigenomic signatures between disease states[8,9], and between joint sites themselves[10,11], increasing our understanding of the joint specificity of disease. OA is multifactorial, with genetic risk factors contributing to ~30% of the lifetime risk of developing knee OA. The integration of DNAm data into genetic studies, such as the statistical fine mapping of GWAS signals, has identified the co-localisation of methylation quantitative trait loci (mQTLs) with genetic risk signals[8,11–13]. This interplay between DNA sequence and CpG methylation status[14] has further been shown to underpin tissue-specific molecular mechanisms of gene expression within the joint[15,16].

Developmental factors also play a role in the risk of OA onset and progression in older age[17]. This includes both the shape of the joint[18–20] (which affects the biomechanical properties and weight-bearing capacity) and the biochemical composition of the articular cartilage (which affects the resilience of the tissue to withstand stresses throughout the life course). Our recent targeted study of 39 CpGs investigated the presence of OA mQTLs in human foetal limbs, at seven well-characterised OA genetic risk loci[21]. We identified that at 85% of the CpGs, the significant OA-mQTLs replicated in the foetal tissues, demonstrating for the first time that functional epigenetic mechanisms associated with a musculoskeletal disease of older age operate from the start of life[21].

Murine knockout or inactivation of the three known DNA methyltransferases (DNMTs) is lethal embryonically (*Dnmt1* and *Dnmt3b*) or postnatally (*Dnmt3a*), demonstrating the vital role of DNAm throughout development[22,23]. DNA methylating enzymes are expressed in proliferating chondrocytes during embryonic development and persist in articular chondrocytes during post-natal development[24]. Investigation of the methylome of human developmental cartilage has been limited to two studies to date, which utilised small sample numbers (n=8) falling within a narrow developmental window (14-19pcw)[25], or, in the case of our previous study, applying targeted approaches which captured only a handful of CpGs[21]. Here, we present a comprehensive study of the methylomic trajectory of human articular cartilage in a broader window of development to capture novel risk loci for joint disease.

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

Methods

Sample Collection

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

Generation of methylation data and genotyping

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

Transcript expression analysis measured by RT-qPCR.

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

analysed relative to the expression of three housekeeping genes: *18S, GAPDH* and *HPRT1* using the $2^{-\Delta ct}$ method as previously described[11].

mQTL analysis

mQTL analysis was performed using the matrixEQTL R package (v2.3), including Sentrix ID (batch), sex and gestation stage as covariates and testing for *cis*-associations based on a distance threshold of 500kb[29]. To account for ancestry, zero to four ancestry principal components (PC) from the 1000 genomes PC analysis (PCA) were trialled for inclusion in the mQTL analysis. The number of nominally significant mQTLs was compared between the different models and two PCs (79.86% of the total variance) were included in the final model. Resultant mQTLs were filtered based on an FDR < 0.05 or a conservative Bonferroni correction calculated using the matrixEQTL-reported total number of statistical tests. Significant mQTLs were associated with the protein-coding genes that they were closest to, based on EnsDb (v75) annotations[30].

Colocalisation analysis

The previously reported 100 independent genome-wide significant SNVs associated with 11 OA skeletal site phenotypes were used to assess colocalisation with the mQTL data[31]. Summary statistics for each phenotype were obtained from the musculoskeletal knowledge portal (https://msk.hugeamp.org/). MAF was calculated from the effect allele frequency, and case and control sizes for each phenotype were extracted from the respective publication. To assess mQTLs within a range of ±500kb from independent OA risk loci signal SNVs, we repeated the mQTL analysis while retaining all mQTL results (pvOutputThreshold=1). For methylation probes within ±500kb of the risk loci, we tested associated SNVs for colocalisation with genetic variants found in both the *cis* mQTL and the GWAS data. Colocalisation analyses were performed separately for each GWAS phenotype using the coloc R package[32] (v 5.1.0) 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

Results

The DNA methylome remains dynamic throughout cartilage development.

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

We profiled the DNA methylation (DNAm) of the 72 samples using Illumina HumanMethylation850 EPIC v1 microarrays. One sample was dropped due to mismatch between predicted and labelled sex. The remaining 71 samples (Table S2) had a minimum bisulphite conversion rate of 89.69% and following QC we retained data at 678,267 CpGs. Data were normalised via quantile-normalisation, followed by principal component analysis (PCA), where developmental stage was found to strongly correlate with PC1, which accounted for 17% of the total variance in global methylation (Fig. 1B). Across the captured developmental window, we identified 53,800 significantly differentially methylated probes (DMPs; FDR<0.001, log2fc>0.1; Table S3), with 48% becoming significantly hypomethylated through skeletogenesis. We analysed the trend in methylation changes across the developmental stages, identifying three clusters. Clusters 1 and 2 became significantly hypomethylated across the developmental window (Fig. 1C; red and blue, respectively). The primary difference separating the two groups was an observable lag in decreased methylation in Cluster 2, where DNAm remained stable until >10pcw. Conversely, the CpGs belonging to Cluster 3 (Fig. 1C; green) became hypermethylated throughout development. GO analysis revealed similar and shared significant terms (FDR<0.05) for Clusters 1 and 2, including biological processes such as “actin filament-based processes” and “cell morphogenesis”

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

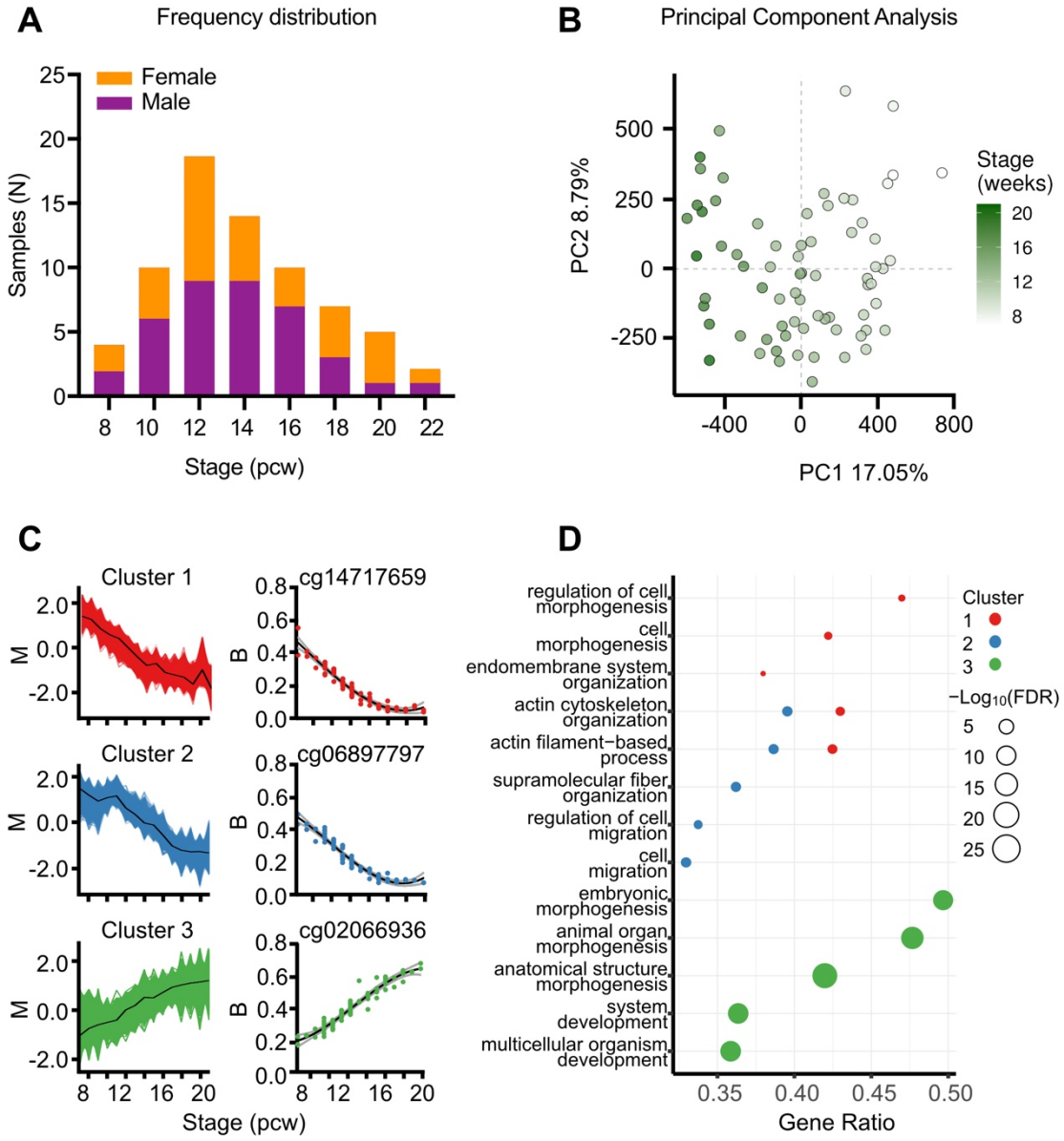


Figure 1 DNA methylation significantly changes at individual CpGs in foetal knee cartilage across the 7-21pcw developmental window. A, Histogram of samples (N=71) showing the distribution of developmental stage. Purple, Male; Orange, female. **B,** Principal component analysis of the samples revealed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

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

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

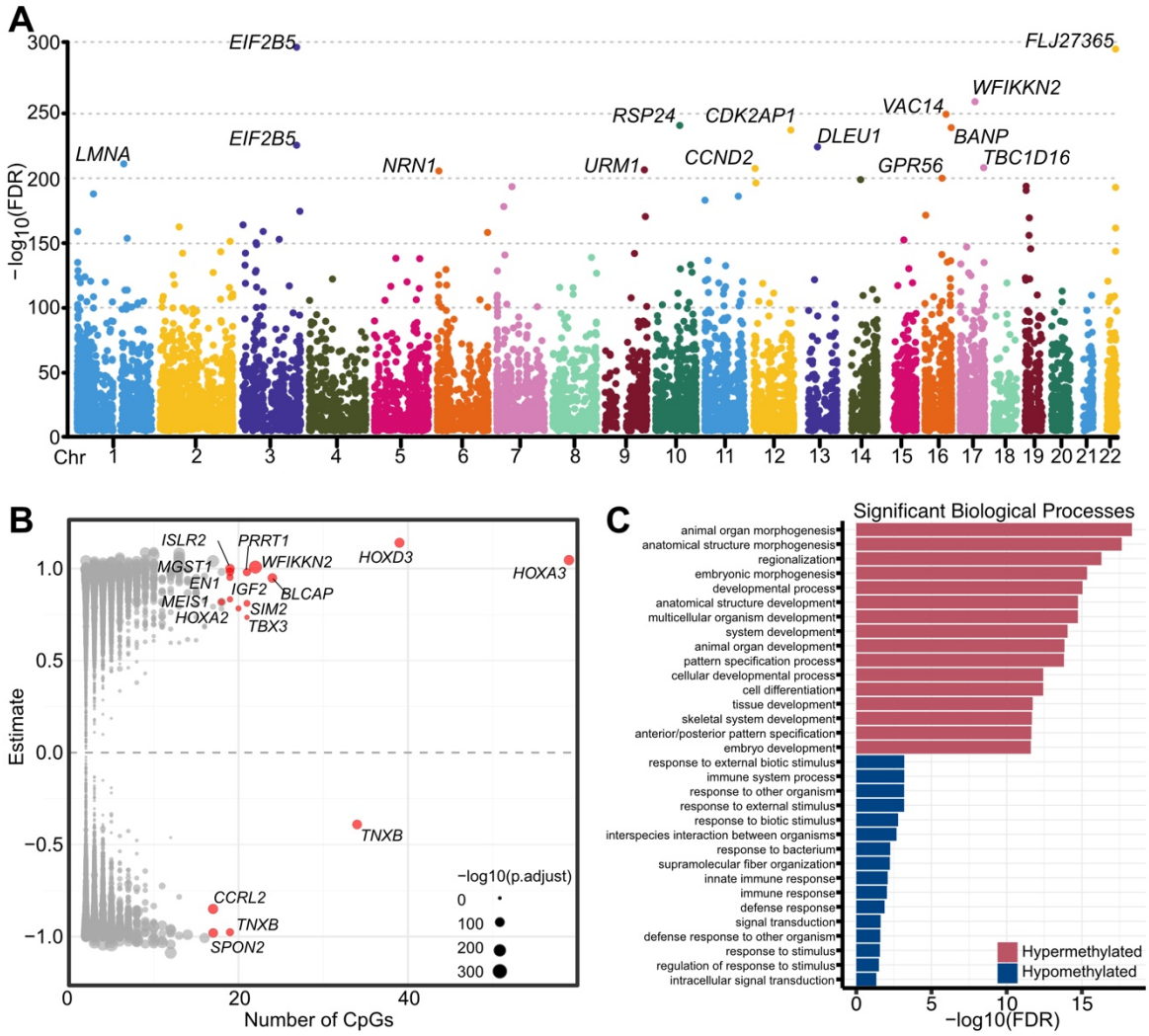


Figure 2. **Co-regulation of DNAm at individual CpGs results in developmental differentially methylated regions (dDMRs) in foetal knee cartilage. A,** Manhattan plot of dDMRs across the epigenome. The nearest protein-coding gene to the most significant dDMRs is labelled. **B,** Volcano plot of dDMRs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

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

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

To investigate the putative functional role of these DMRs in developmental gene regulation, we measured the expression eight of the genes mapping to the dDMRs presented in Table 1. The six transcripts mapping to hypermethylated regions (*MEIS1, EN1, HOXD3, IRX1, HOXA3, SIM2*) were expressed at low levels throughout the captured developmental window (Fig.S2, red). The expression of *HOXA3* negatively correlated with DNAm at 25/59 individual CpGs within the dDMR (P=0.048-0.008; Table S9, Fig.3A LHS). Interestingly, the *HOX* genes mapping to this dDMR do not have reported functions in limb development. We compared their expression to that of *HOXA10* and *HOXA11*, which have recently been reported to be expressed specifically in the zone between the proximal and distal human hindlimb at 5.6pcw, in the region where the knee cavity will form[33]. These genes are also located at the *HOXA* gene locus, a region of tightly controlled co-regulation (Fig.3A upper panel). However, their expression showed no correlation with DNAm within this region (P>0.05; Fig.3A middle and RHS; Table S9).

All three genes mapping to the largest hypomethylated regions, *CCRL2* (labelled dDMR 2153), *SPON2* (dDMR 2624) and *TNXB* (dDMRs 3804 and 3813), fell within putative chondrocyte enhancers and regions of open chromatin in foetal chondrocytes (Fig.3B-D). The expression of all three genes significantly increased in correlation with decreasing methylation within the mapped dDMRs at 3/17 (P=0.017 - 0.041), 6/17 (P=2.8x10^{-6} - 0.002) and 42/53 (P=1.82x10-^{7} – 0.017) CpGs, respectively (Fig.3B-D).

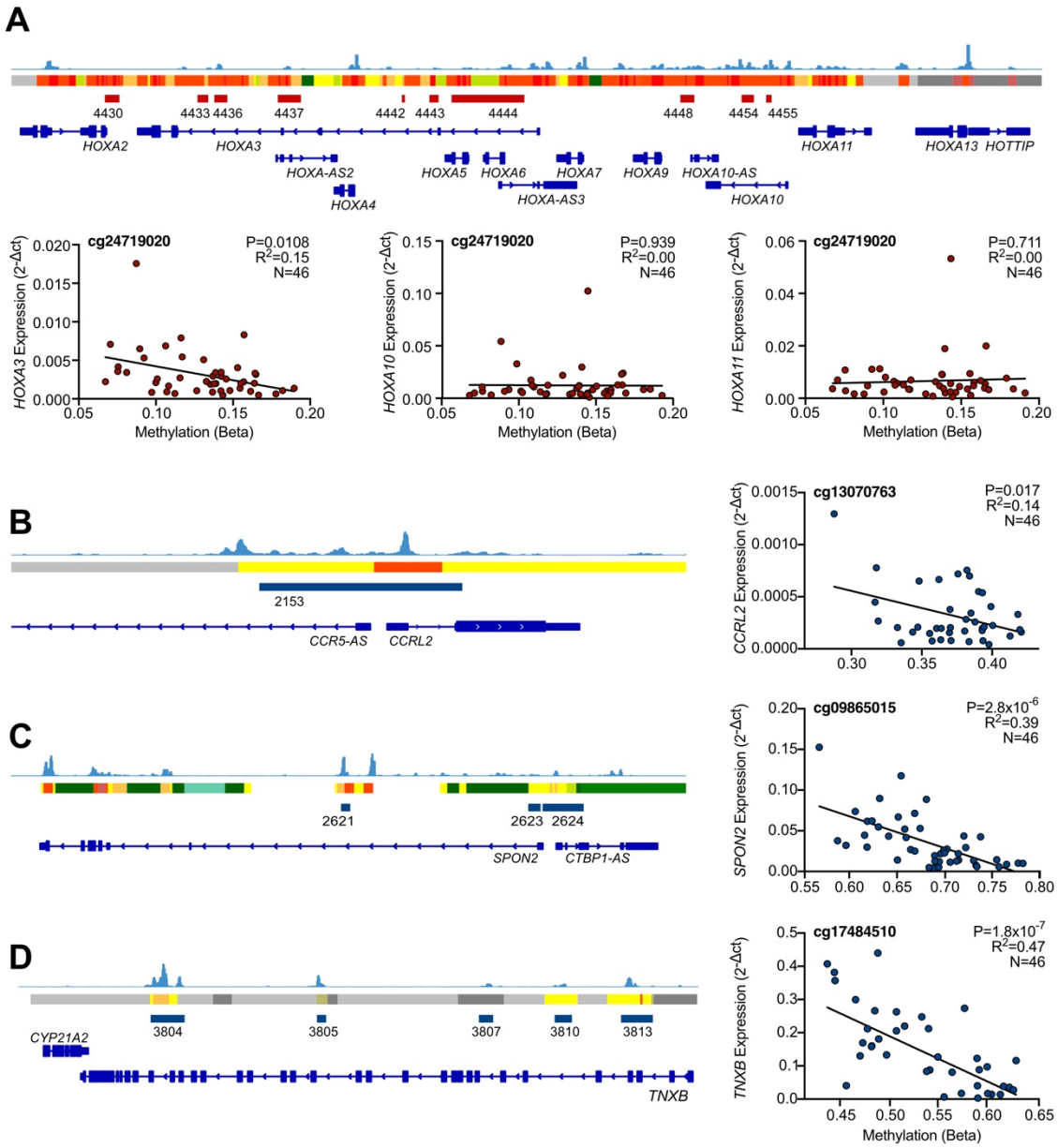


Figure 3. The expression of genes mapping to dDMRs significantly correlates with methylation within the regions. A, (Upper) Screenshot of the Integrated Genome Viewer (IGV) displaying the HOXA gene locus on chromosome 7. Foetal knee cartilage ATAC-seq peaks (light blue) show regions of open chromatin. 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.

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

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

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

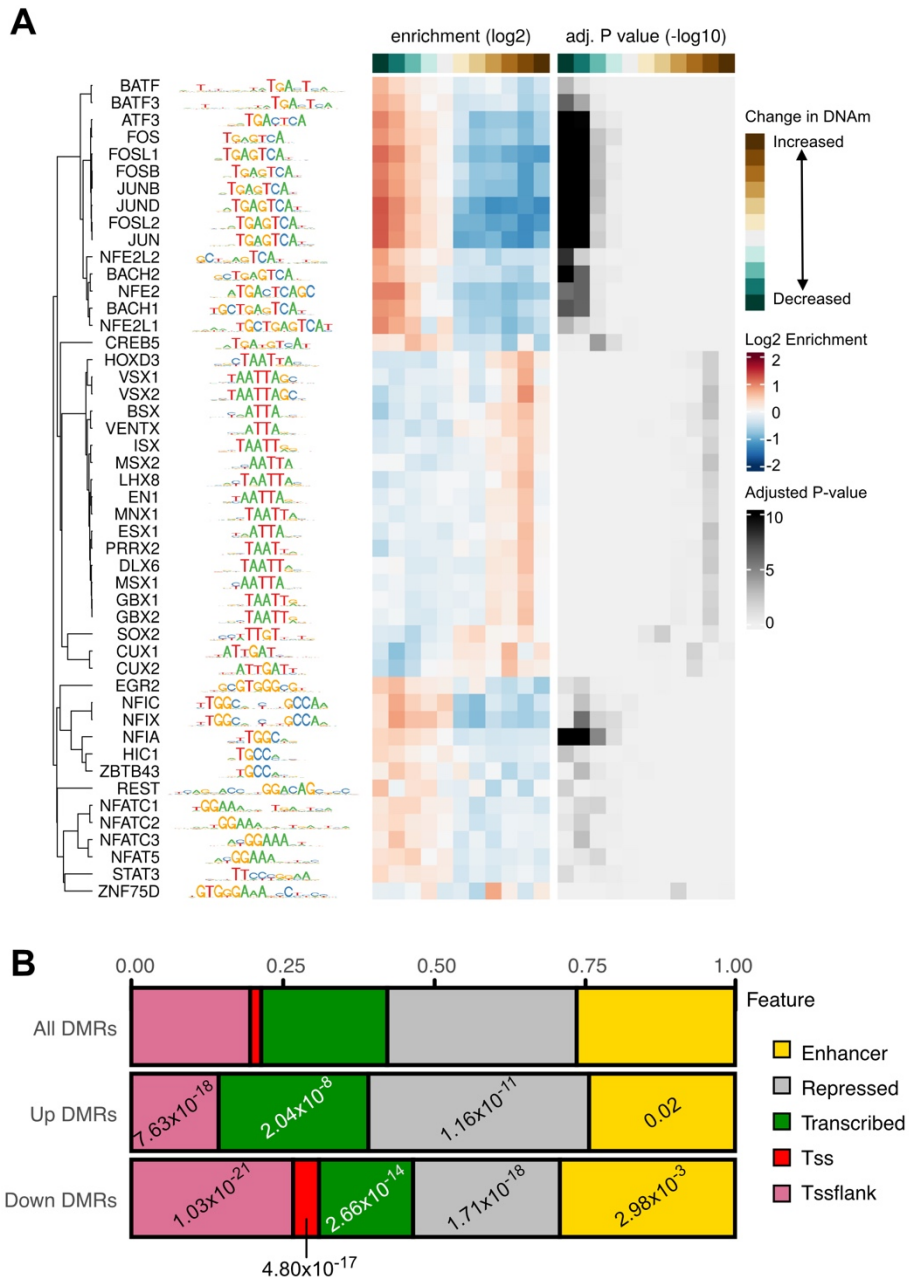


Figure 4 Developmental DMRs are enriched in regulatory elements. A, Transcription factor (TF) motifs taken from the Hocomoco database which are enriched in dDMRs **Left,** name and logo plot of enriched TF motif within the dDMRs, **Centre,** Heatmap of Log2 enrichment value of the motifs within hypermethylated (brown) or hypomethylated (green) dDMRs. **Right,** Heatmap of -Log10 adjusted P-value. Darker colour indicates higher statistical significance. **B,** Frequency of identified dDMRs overlapping with human cultured chondrocyte (Roadmap E049) chromatin states. Enhancer, yellow; repressed region, grey; transcribed region, green; transcription start site (TSS), red; region flanking TSS, pink. Statistical significance was calculated using Fisher's exact test with Bonferroni correction with all significant DMRs as background. Significant under- or overrepresentation is labelled with the adjusted P-value.

Chondrocyte DNA methylation in articular cartilage development exhibits sexual dimorphism.

Previously, methylome analysis performed in 358 blood samples from extremely premature neonates (at 23-27 weeks' gestation, which relates to ~21-25pcw) showed that a large proportion of CpGs exhibited sexually dimorphic levels of DNAm[37]. Using the EPIC array, Santos *et al.,* identified 5595 DMPs (FDR<0.00001), 95% of which were hypermethylated in females. We tested for significant sex-specific differences in DNAm amongst our developmental cartilage samples. Epigenome-wide analysis revealed 811 sex-specific DMPs (sDMPs; FDR<0.05), mapping to 534 genes (Table S11, Fig.S3A). Of these sites, 523 (64.5%) were hypermethylated in the female samples. Fifty-seven percent (461) of the sDMPs we identified in developing articular cartilage also showed significant sex effects in the neonatal blood dataset (Fig.S3B). There was a direct overlap between all of the top 20 reported probes, with methylation effects occurring in the same direction (Table 2, Fig.S3C). Amongst the genes mapping to these CpGs are *NAB1, RFTN1,* and *CEP170.*

We further identified 60 significant sDMRs (FDR<0.01), consisting of 178 CpGs (Table S12). At 46 identified regions (77%), DNAm was significantly higher in the female samples than in male. No significant GO terms (all FDR>0.05) were associated with either the sDMPs or sDMRs, consistent with the findings of Santos *et al.,* in their analysis of neonatal blood samples[37].

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

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

383 to the genes *ELMOD1* (P.adj=3.7x10^{-50}), *ATPIF1* (P=1.4x10^{-48}), and *PHGDH* (P.adj=4.0x10^{-}
384 ^{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 chondrocytes[21], regions of H3K27ac
392 enrichment in foetal limb tissues[38], and with designated chromatin states in mesenchymal
393 stem cells (MSCs) and chondrocytes[39]. 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.

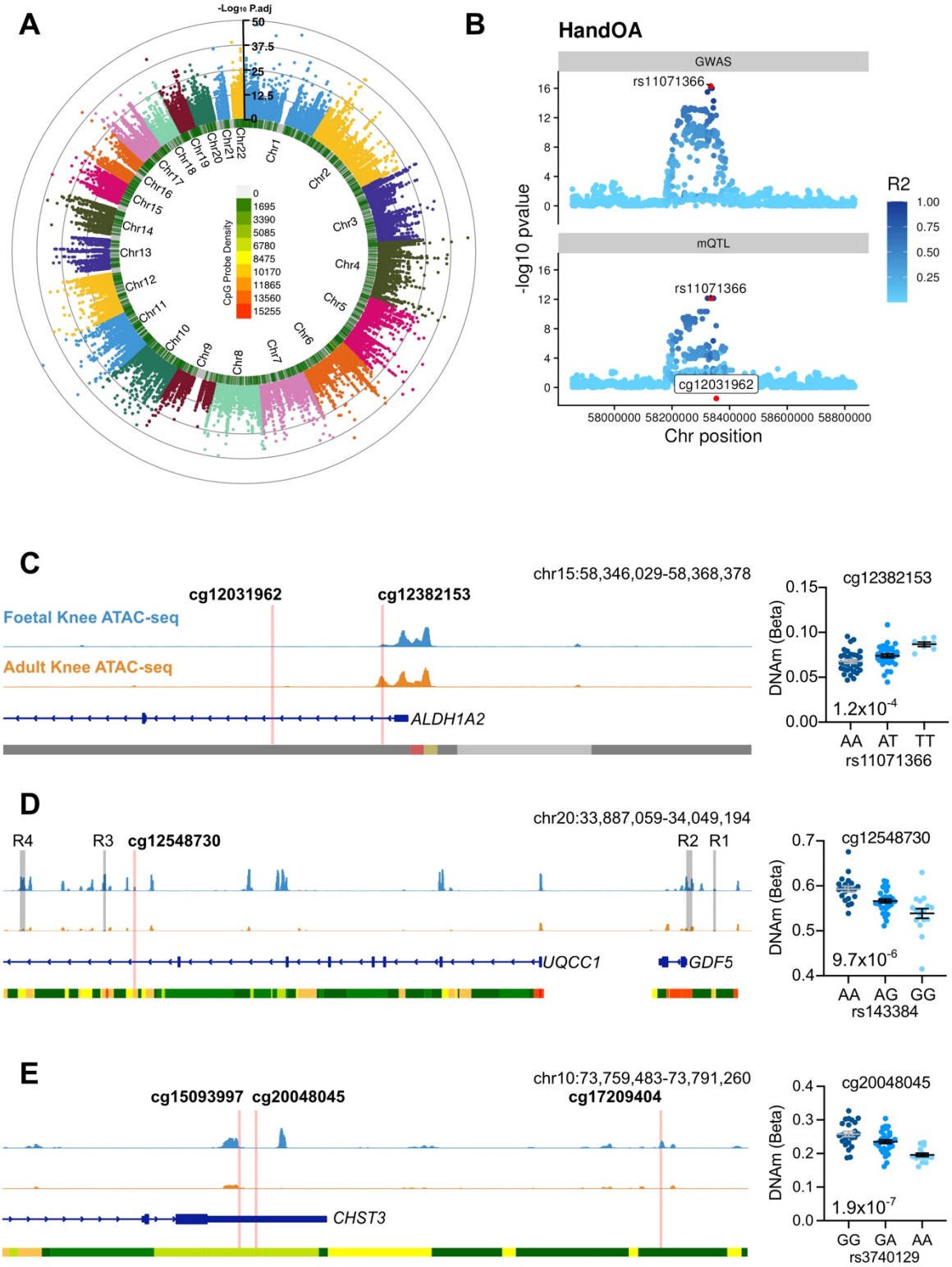
398 One example of OA-mQTLs operating in both foetal and aged human cartilage is on
399 chromosome 15, where mQTLs at cg12031962 and cg12382153 map to *ALDH1A2*
400 (P=7.4x10^{-13} and 1.3x10^{-4}, respectively). These mQTLs co-localise with the primary
401 phenotypes “Hand OA” (PP=0.99, Fig. 5B), “Thumb 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 cartilage[8].

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

phenotypes (Table S14). The CpG is located within an intron of *UQCC1*, downstream of previously reported *GDF5* cartilage enhancers R3 and R4[40,41]. The OA effect allele (G) significantly correlated with lower levels of DNAm ($P=9.7x10^{-6}$; Fig.5D). Analysis of our existing knee cartilage ATAC-seq dataset in human foetal (distal femur) and adult cartilage samples showed that the CpG is located adjacent to an open chromatin peak in foetal cartilage (blue), which is not an accessible region in OA cartilage (orange). This region has not previously been investigated as a regulatory element for *GDF5.* A previous mQTL at this locus has been reported in adult OA cartilage, located adjacent to the promoter of *UQCC1* (cg14752227)[11].

At the majority of the loci (19), no previous reports of OA-mQTLs have been reported through analyses of adult cartilage. On chromosome 10, three OA-mQTLs were identified that fell within putative regulatory elements of *CHST3,* encoding carbohydrate sulfotransferase 3 (Fig.5E). Two of the CpGs, cg15093997 and cg20048045, located within the 3'UTR of the gene, near to open chromatin regions (Fig.5E). At both CpGs, the effect allele, A, at rs3740129 (G>A) correlated with significantly decreased levels of DNAm ($P=1.6x10^{-12}$ and $1.9x10^{-7}$, respectively). Conversely, at the more distal CpG (cg17209404), located within an intergenic open chromatin region only in foetal cartilage, DNAm was significantly increased in the presence of the OA effect allele (P=3.9x10-4; Table S14).

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



444
445

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

Finally, we tested whether the 11,299 CpGs which comprise the identified significant mQTLs across the epigenome are amenable to external influences occurring between sexes (811 sDMPs) and through articular cartilage development (80,536 dDMPs). Between the mQTLs and dDMPs, an overlap of just 744 CpGs was identified, indicating that only 6.6% of CpGs that are significantly regulated by DNA sequence are impacted by developmental stage ($P=5.2x10^{-80}$). Similarly, between the mQTL CpGs and the 811 sDMPs, an overlap of only 6 CpGs (0.74%) was observed (P= 0.0184). This indicates that the CpGs at which DNAm is heavily influenced by the genetic architecture of the region are less amenable to environmental influences on the epigenome.

Discussion

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

In 2023, Zhang *et al* published the first single-cell and spatial transcriptomic atlas of the developing human hindlimb[33]. They provided evidence of nascent articular chondrocytes

that expressed high levels of *PRG4* in the distal femur by 8pcw. In our data, the overwhelming impact of the developmental stage upon the cartilage methylome partly reflects the changing nature of the cartilage cellular composition captured in our bulk analysis. In earlier stages (<12pcw) we have likely captured epiphyseal chondrocytes along with the nascent articular cartilage that will continue to develop postnatally, whereas after 12pcw, we were able to isolate the nascent articular cartilage more specifically. Our gene expression profiling data provide strong evidence that the primary population of cells analysed in this study are nascent epiphyseal and articular chondrocytes. Underscoring this, our data also align with studies of embryonic development of epiphyseal and articular cartilage in mouse models where time-course expression and lineage tracing studies have shown that both *Gdf5* and *Sox9* expression decreases, *Col2a1* remains relatively stable and *Prg4* increases as joint development proceeds through cavitation and beyond[42–44]. In 2015, Spiers *et al.* applied the Illumina HumanMethylation450 array to DNA extracted from 179 human foetal brain samples between 3 and 26pcw[45]. Comparable to our dataset, they also identified that DNAm at ~7% of captured CpGs significantly changed across the developmental window, despite the development within brain tissue of a more heterogenous cell population.

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

gene transcriptional start sites. These observations also align with functional studies in murine models of limb development and the recent spatial maps of gene expression in human limb development[33]. *Meis1* is a regulator of proximodistal identity[46,47] and was further shown by Zhang *et al* to be expressed specifically in the proximal hindlimb (developing femur) at 5.6pcw[33]. *En1* and *Irx1,* which orchestrate early murine limb bud patterning and digit formation, respectively are most highly expressed earlier and in the most distal locations in the limb[48].

Conversely, several large hypomethylated regions mapped to genes encoding ECM proteins including Spondin-2 and Tenascin-X, highlighting the role of the DNA methylome in underlying transcriptional changes. We investigated the expression of the three genes mapping to the largest significant hypomethylated dDMRs, identifying that all significantly increased with decreasing DNAm. Across the developmental window, hypomethylated dDMRs were enriched in CREs, particularly annotated chondrocyte enhancers (29% overlapped with enhancers, compared to 25% of all dDMRs). This is consistent with earlier reports using an MSC *in vitro* model that, throughout differentiation, hypomethylation of DNA is enriched in enhancer regions[49].

Dimorphism in DNAm levels between the biological sexes has been identified in a multitude of tissues across the human lifespan and is believed to contribute to the discrepancy in disease trajectories between males and females[50]. A recent study demonstrated that developmental sexual dimorphism in DNAm is tissue-specific, with a trend for hypomethylation in the placental tissues of female premature neonates, and hypermethylation in blood[37]. Due to the known discrepancies in the prevalence of knee OA in men and women, we investigated sDMRs at autosomal loci to look for sex-specific differences within the developing knee joint. The most significant sDMR mapped to the gene *NAB1*, encoding nerve growth factor induced (NGFI)-A-binding protein 1, a transcriptional repressor[51] which is also the most significant DMP identified in neonatal blood. We found no significant interactions in our samples between sex and developmental stage upon DNAm. Sex differences in methylation profiles could potentially be reflective of different cell type populations in the same tissues, or subsets of 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 blood[37].

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 life[21]. 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 morphogenesis[52]. Mutations in *CHST3* lead to skeletal dysplasia, highlighting the important
550 function of this protein in skeletogenesis[53]. 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,*

encoding RALDH2, integral for all-*trans* retinoic acid (atRA) synthesis. *G*enetic studies of hand OA have been conducted on *ALDH1A2* based on the lead variant rs11071366, first identified in the Icelandic population[54]. Allelic expression imbalance at the locus showed a decrease in gene expression driven by the OA risk allele, which was present in osteochondral tissue from the trapezium, along with cartilage from the hip and knee[55,56]. This effect was also replicated in our targeted analysis of human foetal cartilage[21]. Furthermore, an OA-mQTL was identified at cg12031962, where the minor and effect allele, C, correlated with a decrease in DNAm[11]. In this study, we colocalised the OA risk SNV rs11071366 (A>T)[31], which is in moderate pairwise linkage disequilibrium in European populations (r^2=0.56). In our analysis of the foetal tissues, we also identified an OA-mQTL at the same site, where the effect allele, A, was also associated with decreased DNAm. Recent studies have shown that atRA has an anti-inflammatory effect within the joint, and that pharmacologically blocking its cellular metabolism exhibited a chondroprotective effect[56]. As an example of an epigenetic OA risk mechanism occurring throughout the life course, this knowledge could influence the key therapeutic window for future targeted interventions.

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

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

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

Acknowledgements

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

Author Contributions

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

Declaration of Interests

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

Data Availability

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

Supplemental information

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

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

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

Szpak, M., Thormann, A., Tricomi, F. F., Urbina-Gómez, D., Veidenberg, A., Walsh, T. A., Walts, B., Willhoft, N., Winterbottom, A., Wass, E., Chakiachvili, M., Flint, B., Frankish, A., Giorgetti, S., Haggerty, L., Hunt, S. E., Iisley, G. R., Loveland, J. E., Martin, F. J., Moore, B., Mudge, J. M., Muffato, M., Perry, E., Ruffier, M., Tate, J., Thybert, D., Trevanion, S. J., Dyer, S., Harrison, P. W., Howe, K. L., Yates, A. D., Zerbino, D. R. & Flicek, P. Ensembl 2022. *Nucleic Acids Res* **50,** D988–D995 (2022).

31. Boer, C. G., Hatzikotoulas, K., Southam, L., Stefánsdóttir, L., Zhang, Y., Coutinho de Almeida, R., Wu, T. T., Zheng, J., Hartley, A., Teder-Laving, M., Skogholt, A. H., Terao, C., Zengini, E., Alexiadis, G., Barysenka, A., Bjornsdottir, G., Gabrielsen, M. E., Gilly, A., Ingvarsson, T., Johnsen, M. B., Jonsson, H., Kloppenburg, M., Luetge, A., Lund, S. H., Mägi, R., Mangino, M., Nelissen, R. R. G. H. H., Shivakumar, M., Steinberg, J., Takuwa, H., Thomas, L. F., Tuerlings, M., Loughlin, J., Arden, N., Birrell, F., Carr, A., Deloukas, P., Doherty, M., McCaskie, A. W., Ollier, W. E. R., Rai, A., Ralston, S. H., Spector, T. D., Wallis, G. A., Martinsen, A. E., Willer, C., Fors, E. A., Mundal, I., Hagen, K., Nilsen, K. B., Lie, M. U., Børte, S., Brumpton, B., Nielsen, J. B., Fritsche, L. G., Zhou, W., Heuch, I., Storheim, K., Tyrpenou, E., Koukakis, A., Chytas, D., Evangelopoulos, D. S., Efstathios, C., Pneumaticos, S., Nikolaou, V. S., Malizos, K., Anastasopoulou, L., Abecasis, G., Baras, A., Cantor, M., Coppola, G., Deubler, A., Economides, A., Lotta, L. A., Overton, J. D., Reid, J. G., Shuldiner, A., Karalis, K., Siminovitch, K., Beechert, C., Forsythe, C., Fuller, E. D., Gu, Z., Lattari, M., Lopez, A., Schleicher, T. D., Padilla, M. S., Widom, L., Wolf, S. E., Pradhan, M., Manoochehri, K., Bai, X., Balasubramanian, S., Boutkov, B., Eom, G., Habegger, L., Hawes, A., Krasheninina, O., Lanche, R., Mansfield, A. J., Maxwell, E. K., Nafde, M., O'Keeffe, S., Orelus, M., Panea, R., Polanco, T., Rasool, A., Salerno, W., Staples, J. C., Li, D., Sharma, D., Banerjee, I., Bovijn, J., Locke, A., Verweij, N., Haas, M., Hindy, G., De, T., Akbari, P., Sosina, O., Ferreira, M. A. R., Jones, M. B., Mighty, J., LeBlanc, M. G., Mitnaul, L. J., Babis, G. C., Cheung, J. P. Y., Kang, J. H., Kraft, P., Lietman, S. A., Samartzis, D., Slagboom, P. E., Stefansson, K., Thorsteinsdottir, U., Tobias, J. H., Uitterlinden, A. G., Winsvold, B., Zwart, J. A., Davey Smith, G., Sham, P. C., Thorleifsson, G., Gaunt, T. R., Morris, A. P., Valdes, A. M., Tsezou, A., Cheah, K. S. E., Ikegawa, S., Hveem, K., Esko, T., Wilkinson, J. M., Meulenbelt, I., Lee, M. T. M., van Meurs, J. B. J., Styrkársdóttir, U. & Zeggini, E. Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations. *Cell* **184,** 4784-4818.e17 (2021).

32. Wang, G., Sarkar, A., Carbonetto, P. & Stephens, M. A Simple New Approach to Variable Selection in Regression, with Application to Genetic Fine Mapping. *J R Stat Soc Series B Stat Methodol* **82,** 1273–1300 (2020).

33. Zhang, B., He, P., Lawrence, J. E. G., Wang, S., Tuck, E., Williams, B. A., Roberts, K., Kleshchevnikov, V., Mamanova, L., Bolt, L., Polanski, K., Li, T., Elmentaite, R., Fasouli, E. S., Prete, M., He, X., Yayon, N., Fu, Y., Yang, H., Liang, C., Zhang, H., Blain, R., Chedotal, A., FitzPatrick, D. R., Firth, H., Dean, A., Bayraktar, O. A., Marioni, J. C., Barker, R. A., Storer, M. A., Wold, B. J., Zhang, H. & Teichmann, S. A. A human embryonic limb cell atlas resolved in space and time. *Nature 2023* 1–11 (2023). doi:10.1038/s41586-023-06806-x

34. Zhang, F., Wang, Y., Zhao, Y., Wang, M., Zhou, B., Zhou, B. & Ge, X. NFATc1 marks articular cartilage progenitors and negatively determines articular chondrocyte differentiation. *Elife* **12,** (2023).

35. Zhang, C. H., Gao, Y., Hung, H. H., Zhuo, Z., Grodzinsky, A. J. & Lassar, A. B. Creb5 coordinates synovial joint formation with the genesis of articular cartilage. *Nature Communications 2022 13:1* **13,** 1–18 (2022).

36. Tavares, A. T., Tsukui, T. & Belmonte, J. C. I. Evidence that members of the Cut/Cux/CDP family may be involved in AER positioning and polarizing activity during chick limb development. *Development* **127,** 5133–5144 (2000).

37. Santos, H. P., Enggasser, A. E., Clark, J., Roell, K., Zhabotynsky, V., Gower, W. A., Yanni, D., Yang, N. G., Washburn, L., Gogcu, S., Marsit, C. J., Kuban, K., O'Shea, T.

M. & Fry, R. C. Sexually dimorphic methylation patterns characterize the placenta and blood from extremely preterm newborns. *BMC Biol* **21,** 1–12 (2023).

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

dimorphism in the human placenta: implications for susceptibility during the prenatal period. *http://dx.doi.org/10.2217/epi-2016-0132* **9,** 267–278 (2017).

51. Thiel, G., Kaufmann, K., Magin, A., Lietz, M., Bach, K. & Cramer, M. The human transcriptional repressor protein NAB1: expression and biological activity. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1493,** 289–301 (2000).
52. Kitagawa, H., Tsutsumi, K., Tone, Y. & Sugahara, K. Developmental regulation of the sulfation profile of chondroitin sulfate chains in the chicken embryo brain. *J Biol Chem* **272,** 31377–31381 (1997).
53. Thiele, H., Sakano, M., Kitagawa, H., Sugahara, K., Rajab, A., Höhne, W., Ritter, H., Leschik, G., Nürnberg, P. & Mundlos, S. Loss of chondroitin 6-O-sulfotransferase-1 function results in severe human chondrodysplasia with progressive spinal involvement. *Proc Natl Acad Sci U S A* **101,** 10155–10160 (2004).
54. Styrkarsdottir, U., Thorleifsson, G., Helgadottir, H. T., Bomer, N., Metrustry, S., Bierma-Zeinstra, S., Strijbosch, A. M., Evangelou, E., Hart, D., Beekman, M., Jonasdottir, A., Sigurdsson, A., Eiriksson, F. F., Thorsteinsdottir, M., Frigge, M. L., Kong, A., Gudjonsson, S. A., Magnusson, O. T., Masson, G., Hofman, A., Arden, N. K., Ingvarsson, T., Lohmander, S., Kloppenburg, M., Rivadeneira, F., Nelissen, R. G. H. H., Spector, T., Uitterlinden, A., Slagboom, P. E., Thorsteinsdottir, U., Jonsdottir, I., Valdes, A. M., Meulenbelt, I., Van Meurs, J., Jonsson, H. & Stefansson, K. Severe osteoarthritis of the hand associates with common variants within the ALDH1A2 gene and with rare variants at 1p31. *Nat Genet* **46,** 498–502 (2014).
55. Shepherd, C., Zhu, D., Skelton, A. J., Combe, J., Threadgold, H., Zhu, L., Vincent, T. L., Stuart, P., Reynard, L. N. & Loughlin, J. Functional Characterization of the Osteoarthritis Genetic Risk Residing at ALDH1A2 Identifies rs12915901 as a Key Target Variant. *Arthritis and Rheumatology* **70,** 1577–1587 (2018).
56. Zhu, L., Kamalathevan, P., Koneva, L. A., Zarebska, J. M., Chanalaris, A., Ismail, H., Wiberg, A., Ng, M., Muhammad, H., Walsby-Tickle, J., McCullagh, J. S. O., Watt, F. E., Sansom, S. N., Furniss, D., Gardiner, M. D. & Vincent, T. L. Variants in ALDH1A2 reveal an anti-inflammatory role for retinoic acid and a new class of disease-modifying drugs in osteoarthritis. *Sci Transl Med* **14,** (2022).
57. Kreitmaier, P., Park, Y.-C., Swift, D., Gilly, A., Wilkinson, J. M., Zeggini, E. & Wilkinson, † J Mark. Epigenomic profiling of the infrapatellar fat pad in osteoarthritis. *Hum Mol Genet* (2023). doi:10.1093/HMG/DDAD198
58. Aubourg, G., Rice, S. J., Bruce-Wootton, P. & Loughlin, J. Genetics of osteoarthritis. *Osteoarthritis Cartilage* (2021). doi:10.1016/j.joca.2021.03.002

885

Top hypermethylated DMRs							
DMR ID (Cluster)	Chr	Start	End	Gene	Classification/Function	FDR	CpGs (N)
1413 (3)	2	66665027	66668012	*MEIS1*	Transcriptional regulator of proximodistal limb identity	$2.03x10^{-36}$	18
1576 (3)	2	119607192	119611880	*EN1*	Engrailed Homeobox Transcription Factor	$2.26x10^{-44}$	19
1755 (3)	2	177021702	177030228	*HOXD3*	Homeobox Transcription Factor	$6.31x10^{-127}$	39
3054 (3)	5	3597311	3602413	*IRX1*	Iroquois Homeobox Transcription Factor	$1.98x10^{-18}$	18
3759 (3)	6	30848726	30851086	*DDR1*	Discoidin Domain Receptor	$2.01x10^{-97}$	18
3825 (3)	6	32115979	32118204	*PRRT1*	Transmembrane Protein	$1.43x10^{-69}$	21
4430 (3)	7	27142100	27143788	*HOXA2*	Homeobox Transcription Factor	$2.76x10^{-22}$	20
4444 (3)	7	27181067	27189610	*HOXA3*	Homeobox Transcription Factor	$1.48x10^{-140}$	59
6992 (3)	12	16760056	16763480	*MGST1*	Glutathione S-Transferase	$1.49x10^{-104}$	19
8319 (3)	15	74423308	74427499	*ISLR2*	Immunoglobulin Superfamily	$8.15x10^{-76}$	19
9322 (3)	17	48911141	48913317	*WFIKKN2*	Serine and metalloprotease inhibitor	$8.44x10^{-260}$	22
10141 (3)	20	36145902	36148775	*BLCAP*	Regulates cell proliferation and apoptosis	$1.86x10^{-112}$	24
10351 (3)	21	38076709	38082613	*SIM2*	Basic helix-loop-helix transcription factor	$7.42x10^{-37}$	21
Top hypomethylated DMRs							
DMR ID (Cluster)	Chr	Start	End	Gene	Classification/Function	FDR	CpGs (N)
2153 (2)	3	46447254	46449636	*CCRL2*	Chemokine receptor	$3.11x10^{-127}$	17
2624 (2)	4	1202780	1206150	*SPON2*	ECM Glycoprotein	$3.11x10^{-105}$	17
3804 (2)	6	32014059	32016535	*TNXB*	ECM Glycoprotein	$3.80x10^{-117}$	34
3813 (2)	6	32048632	32050947	*TNXB*	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-vlaue	Nearest Gene Symbol	Distance (bp)	Cartilage highest mean B	Blood highest mean B
cg03618918	1	160865097	$1.59x10^{-21}$	*ITLN1*	10136	M	M
cg26919182	1	202522232	$2.41x10^{-22}$	*PPP1R12B*	0	F	F
cg15817705	1	209406063	$2.86x10^{-23}$	*CAMK1G*	350998	M	M
cg25742246	1	214153460	$2.22x10^{-19}$	*PROX1*	3063	F	F
cg12691488	1	243053673	$1.27x10^{-39}$	*CEP170*	234056	M	M
cg15228509	1	243074717	$2.12x10^{-42}$	*CEP170*	213012	F	F
cg02989351	2	9770584	$1.16x10^{-23}$	*YWHAQ*	0	F	F
cg03226871	2	128128958	$2.88x10^{-20}$	*MAP3K2*	0	M	M
cg19765154	2	191524409	$7.70x10^{-36}$	*NAB1*	0	M	M
cg20262915	2	191524489	$2.02x10^{-22}$	*NAB1*	0	M	M
cg00148935	3	16398839	$1.19x10^{-27}$	*RFTN1*	0	F	F
cg07850329	9	33265029	$3.04x10^{-21}$	*CHMP5*	0	F	F
cg07852945	9	84303915	$9.64x10^{-20}$	*TLE1*	0	F	F
cg02716779	9	131016143	$3.27x10^{-22}$	*DNM1*	0	F	F
cg26355737	13	114292172	$2.84x10^{-21}$	*TFDP1*	0	M	M
cg11284736	15	83826708	$3.77x10^{-21}$	*HDGFRP3*	0	M	M
cg03626220	16	71052200	$1.47x10^{-24}$	*HYDIN*	0	M	M
cg03218192	17	33914403	$4.68x10^{-22}$	*AP2B1*	0	F	F
cg12607525	17	42286849	$1.94x10^{-22}$	*UBTF*	0	M	M
cg17612569	21	27107221	$2.38x10^{-25}$	*ATP5J*	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	*TMEM18*	cg14382829	$6.7x10^{-9}$	0.84	+	A				N	N		D
3	rs3774354	*ITIH1*	cg06068388	0.04	0.98	+	A	Y						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	+				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	*TACC3*	cg18815463	$1.7x10^{-6}$	0.89	+	T	Y[8,13]		N	N	N		
4	rs11729628	*PRDM5*	cg07870920	$1.2x10^{-35}$	0.86	-	T			N	N	N		
5	rs10062749	*FGF1*	cg15672022	$6.8x10^{-5}$	0.81	-	T							L
6	rs9396861	*RNF144B*	cg04025088	$1.6x10^{-4}$	0.95	-	A					N		D
6	rs79220007	*HFE*	cg00854024	$1.8x10^{-14}$	1.00	+	T				N	N		D
6	rs2856821	*HLA-DPA1*	cg17362900	$2.4x10^{-8}$	0.97	+	T			N				L
			cg19053046	$4.0x10^{-6}$	0.93	+								
			cg25045942	$1.3x10^{-6}$	0.98	+								
			cg13921245	$9.0x10^{-7}$	0.84	+			Y[8,13,58]	N	N	N		
			cg14310225	0.004	0.87	+				N	N	N		
7	rs116934101	*CUX1*	cg23686983	$1.3x10^{-5}$	0.94	-	A			N	N	N		
9	rs1330349	*TNC*	cg12835026	0.005	0.88	-	A			N	N	N		
10	rs3740129	*CHST3*	cg15093997	$1.6x10^{-12}$	0.97	-	A				N			L
			cg20048045	$1.9x10^{-7}$	0.98	-					N	N		
			cg17209404	$3.9x10^{-4}$	0.90	+						N		
11	rs3993110	*TEAD1*	cg25843174	$1.3x10^{-11}$	0.98	+	A				N	N		D
11	rs1631174	*PTPRJ*	cg11545521	$1.5x10^{-5}$	0.80	-	A			N	N	N		
11	rs67924081	*EHBP1L1*	cg01366692	$2.1x10^{-7}$	0.98	-	A			N	N	N		L
			cg17120908	$6.3x10^{-18}$	0.99	-								

11	rs1149620	*TSKU*	cg17107561	0.002	0.81	-	A			N	N	N			
15	rs11071366	*ALDH1A2*	cg12031962	$7.4x10^{-13}$	0.99	-	A	Y	$Y^{11,58}$	N	N	N			L
			cg12382153	$1.3x10^{-4}$	0.90	+			Y^{8}		N				
15	rs12908498	*SMAD3*	cg15164144	$8.4x10^{-6}$	0.90	+	C			N		N			D
			cg21106065	$3.2x10^{-17}$	0.99	-				N		N			
16	rs34195470	*WWP2*	cg00854324	0.015	0.84	+	A	Y^{13}		N	N	N			
17	rs216175	*SMG6*	cg04375931	0.046	0.99	+	A	Y^{8}							L
17	rs9908159	*C17orf67*	cg24142633	$7.5x10^{-5}$	0.88	+	T			N					L
17	rs7212908	*NACA2*	cg03721816	0.006	0.90	+	A			N		N			D
19	rs10405617	*SLC44A2*	cg01654627	$8.5x10^{-5}$	0.83	-	A	Y^{8}		N		N			D
			cg16876852	$4.5x10^{-8}$	0.95	+				N	N	N			
19	rs8112559	*IRF2BP1*	cg16791036	$2.7x10^{-6}$	0.95	+	C				N	N			D
20	rs143384	*GDF5*	cg12548730	$9.7x10^{-6}$	0.92	-	A	Y^{11}		N		N			D
21	rs9981884	*BRWD1*	cg11644478	$1.2x10^{-5}$	0.84	+	A								L
			cg05821552	$4.3x10^{-8}$	0.96	-				N	N	N			
			cg14227474	$1.7x10^{-5}$	0.91	-				N	N	N			
22	rs12160491	*H1F0*	cg17109681	$1.2x10^{-6}$	0.92	+	A					N			L
			cg06521852	$1.1x10^{-8}$	0.97	+					N	N			
			cg11795195	$2.8x10^{-6}$	0.97	+					N	N			
			cg25404088	$2.8x10^{-7}$	0.95	+									
			cg13303067	$2.9x10^{-4}$	0.86	+									
			cg26028571	$8.0x10^{-5}$	0.91	+				N	N				

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