Primary cartilage transcriptional signatures reflect cell-type-specific molecular pathways underpinning osteoarthritis

Authors

Georgia Katsoula, John E.G. Lawrence, Ana Luiza Arruda, ..., Sarah A. Teichmann, J. Mark Wilkinson, Eleftheria Zeggini

Correspondence

j.m.wilkinson@sheffield.ac.uk (J.M.W.), eleftheria.zeggini@helmholtz-muenchen.de (E.Z.)

This study reveals transcriptional changes in osteoarthritis cartilage linked to embryonic processes and cell types. Genetic risk enrichment in cartilage gene networks uncovers molecular drivers of disease progression. The findings identify effector genes, interaction partners, and potential therapeutic targets for drug repurposing.

Katsoula et al., 2024, The American Journal of Human Genetics *111*, 1–21
December 5, 2024 © 2024 The Authors. Published by Elsevier Inc. on behalf of American Society of Human Genetics.
https://doi.org/10.1016/j.ajhg.2024.10.019

Please cite this article in press as: Katsoula et al., Primary cartilage transcriptional signatures reflect cell-type-specific molecular pathways underpinning osteoarthritis, The American Journal of Human Genetics (2024), https://doi.org/10.1016/j.ajhg.2024.10.019

ARTICLE

Primary cartilage transcriptional signatures reflect cell-type-specific molecular pathways underpinning osteoarthritis

Georgia Katsoula,[1,2,3] John E.G. Lawrence,[4,5] Ana Luiza Arruda,[1,2] Mauro Tutino,[2] Petra Balogh,[6] Lorraine Southam,[2] Diane Swift,[7] Sam Behjati,[5,8] Sarah A. Teichmann,[5,9] J. Mark Wilkinson,[7,10,*] and Eleftheria Zeggini[2,3,10,*]

Summary

Translational efforts in osteoarthritis are hampered by a gap in our understanding of disease processes at the molecular level. Here, we present evidence of pronounced transcriptional changes in high- and low-disease-grade cartilage tissue, pointing to embryonic processes involved in disease progression. We identify shared transcriptional programs between osteoarthritis cartilage and cell populations in the human embryonic and fetal limb, pointing to increases in pre-hypertrophic chondrocytes' transcriptional programs in low-grade cartilage and increases in osteoblastic signatures in high-grade disease tissue. We find that osteoarthritis genetic risk signals are enriched in six gene co-expression modules and show that these transcriptional signatures reflect cell-type-specific expression along the endochondral ossification developmental trajectory. Using this network approach in combination with causal inference analysis, we present evidence of a causal effect on osteoarthritis risk for variants associated with the expression of ten genes that have not been previously reported as effector genes in genome-wide association studies in osteoarthritis. Our findings point to key molecular pathways as drivers of cartilage degeneration and identify high-value drug targets and repurposing opportunities.

Introduction

Osteoarthritis is the most prevalent degenerative joint disorder, affecting over 500 million individuals worldwide.[1] It is a complex disease, mainly characterized by loss of cartilage integrity and low-grade inflammation in tissues surrounding the joint.[2] Management strategies focus on pain relief and joint replacement surgery.[3] An in-depth understanding of the molecular processes leading to disease onset and progression is necessary for the development of disease-modifying therapies.

Genome-wide association studies (GWASs) have identified over 100 independent risk signals for osteoarthritis.[4] Although several high-confidence effector genes have been identified,[4–10] the translational and biological insight potential of resolving GWAS signals remains largely unfulfilled. A limiting factor is sample size and, hence, power of cartilage-specific functional genomics datasets. A further challenge resides in understanding how the biologically diverse effector genes contribute to osteoarthritis and which pathways are implicated in cartilage degeneration. Communities of co-expressed genes can be associated with disease processes, and genes with high network connectivity can be highly informative for disentangling disease architecture.[11]

The value of network approaches in understanding complex diseases lies in the fact that disease phenotypes result from the interplay of various pathobiological processes rather than solely from the dysfunction of individual effector genes.[12] For example, gene co-expression analysis has revealed that the genetic risk of type 2 diabetes (T2D) converges on an RFX6-mediated network that reduces insulin secretion by β cells.[13] Gene co-expression networks have also been widely utilized in studying complex neurological conditions including autism spectrum disorder (ASD) and schizophrenia (SCZ), pointing to networks of synaptic plasticity genes with a causal role for ASD[14] and networks of genes involved in brain development leading to atypical cortical connectivity in individuals at high risk for SCZ.[15] In the context of osteoarthritis, a network diffusion approach has been used to prioritize genes associated with joint damage in mice osteoarthritis models in the OATargets database.[16] The protein products of several known osteoarthritis risk genes (*COLGALT2*, *RUNX2*, *PLEC*, *MGP*, *TGFB1*, *GDF5*) as well as genes residing in osteoarthritis risk loci (*SMAD3*, *CDC5L*) have also been suggested to participate in protein-protein interaction networks related to extracellular matrix

[1]Technical University of Munich (TUM), School of Medicine and Health, Graduate School of Experimental Medicine, 81675 Munich, Germany; [2]Institute of Translational Genomics, Helmholtz Zentrum München – German Research Center for Environmental Health, 85764 Neuherberg, Germany; [3]Technical University of Munich (TUM) and Klinikum Rechts der Isar, TUM School of Medicine and Health, 81675 Munich, Germany; [4]Department of Trauma and Orthopaedics, Cambridge University Hospitals NHS Foundation Trust, Addenbrooke's Hospital, Box 37, Hills Road, Cambridge CB2 0QQ, UK; [5]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton CB10 1SA, UK; [6]Department of Cellular and Molecular Pathology, Royal National Orthopaedic Hospital, Brockley Hill, Stanmore HA7 4LP, UK; [7]School of Medicine and Population Health, University of Sheffield, Beech Hill Road, Sheffield S10 2RX, UK; [8]Department of Paediatrics, University of Cambridge, Cambridge CB2 0QQ, UK; [9]Department of Physics/Cavendish Laboratory, University of Cambridge, JJ Thomson, Cambridge CB3 0HE, UK
[10]The authors contributed equally
*Correspondence: j.m.wilkinson@sheffield.ac.uk (J.M.W.), eleftheria.zeggini@helmholtz-muenchen.de (E.Z.)
https://doi.org/10.1016/j.ajhg.2024.10.019.
© 2024 The Authors. Published by Elsevier Inc. on behalf of American Society of Human Genetics.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

(ECM) organization, tissue development, and cellular response to growth factor stimulus.[17] These findings stress the utility of network approaches for providing mechanistic insight in osteoarthritis. A well-powered study of gene co-expression networks in osteoarthritis is still missing.

Cartilage is an avascular tissue, in which the dominant cell type is the chondrocyte. Articular cartilage is a non-self-renewing tissue that forms the articulating element at the ends of adjacent bones within synovial joints. In contrast, growth plate cartilage, usually only present during skeletogenesis, facilitates skeletal growth and bone formation during endochondral ossification.[18] This process entails the differentiation of multipotent mesenchymal cells into lineage-committed chondroprogenitors, which themselves differentiate into immature, resting chondrocytes.[19] These subsequently undergo a rapid proliferation (proliferating chondrocytes) before entering a transitional phase, known as pre-hypertrophy.[20] Finally, chondrocytes rapidly increase in size to become hypertrophic, before becoming calcified and subsequently replaced by bone, resulting in skeletal growth.

Articular chondrocyte phenotype alterations, such as initiation of hypertrophic differentiation and matrix calcification, have been linked to the degradation of cartilage in the course of osteoarthritis in mice and *in vitro* chondrocyte models.[21–25] Osteoarthritis chondrocytes shift from the normally quiescent articular cartilage phenotype to a more proliferative growth-plate-like phenotype seen during limb development.[26,27] Single-cell RNA sequencing (scRNA-seq) has been recently employed in a small number of individuals with osteoarthritis to characterize chondrocyte heterogeneity,[28–32] with only partial agreement on the identified cell type populations. A consensus map of cartilage composition in osteoarthritis is still missing.

Here, we perform a comprehensive gene expression study of paired, macroscopically intact (low-grade) and degraded (high-grade) *ex vivo* knee osteoarthritis cartilage collected at the point of joint replacement surgery across 300 individuals. We combine our data with publicly available scRNA-seq data from osteoarthritis cartilage and explore cell type heterogeneity between disease grades. We further utilize scRNA-seq of the developing human embryonic and fetal hindlimb and shed light on the transcriptional changes underlying the key pathological process of hypertrophic differentiation in osteoarthritis. We identify deregulated transcriptional programs during disease progression, reflecting causal mechanisms in osteoarthritis phenotypes and specific cell types along the chondrocyte developmental trajectory. Lastly, we pinpoint putative causal effects on the risk of osteoarthritis phenotypes associated with variants linked to the expression of ten previously unreported genes in the largest osteoarthritis GWAS meta-analysis to date,[4] highlighting potential druggable targets.

Methods

Study samples

We recruited individuals with osteoarthritis undergoing total knee replacement for osteoarthritis with no history of significant knee surgery (apart from meniscectomy), knee infection, or fracture and no malignancy within the previous 5 years. It was further verified that no individual had been treated with corticosteroids (systemic or intra-articular) within the previous 6 months or any other drug associated with immune modulation. The matched cartilage samples were isolated from the weight-bearing parts of the joint to ensure that biomechanical loading did not influence within-pair differences in gene expression and were scored macroscopically using the International Cartilage Repair Society (ICRS) scoring system.[33] From each individual, we obtained one cartilage sample of ICRS grade 0 or 1 denoting low-grade osteoarthritis degeneration ("low-grade sample") and one sample of ICRS grade 3 or 4 denoting high-grade osteoarthritis degeneration ("high-grade sample"). Informed consent was given from all study participants, and samples were collected under Human Tissue Authority license 12182 and National Research Ethics Service approval 15/SC/0132 and 20/SC/0144, South Yorkshire and North Derbyshire Musculoskeletal Biobank, University of Sheffield, UK. The biobank is overseen by a steering committee, which includes two lay members. The lay members reviewed this project proposal prior to its initiation and had the opportunity to comment upon and make edits to the study design, as did the Sheffield Lay Advisory Panel for Bone Research. The conduct of the biobank and its outputs are also reviewed by the biobank lay committee members.

Description and justification of population descriptors

RNA-seq data were obtained from individuals of European ancestry based on available matching genotype data.[10] The matching genotype data were used to help determine whether individuals were most genetically similar to a European reference panel and, if so, were included in the study. The rationale of this descriptor was based on the fact that osteoarthritis GWAS data predominantly included individuals of European ancestry, and we aimed for consistency between the ancestry of our RNA-seq samples and the population from which the GWAS data were derived. To ensure accurate heritability estimates, we used the EUR linkage disequilibrium (LD) 1000 Genomes reference panel,[34] which is appropriate for European-ancestry populations. This choice minimizes potential biases that could arise from population stratification when analyzing genetic associations.

Fetal limb tissue collection

To explore links between osteoarthritis and fetal developmental programs, we utilized data from our atlas of endochondral ossification.[35] For this study, first-trimester human fetal tissue was collected from an elective termination of pregnancy procedure at Addenbrooke's Hospital, Cambridge, UK, through the ethically approved Wellcome-MRC Cambridge Stem Cell Institute and Department of Clinical Neurosciences tissue bank (REC-96/085). Written, informed consent was given for tissue collection by the individuals in accordance with The Declaration of Helsinki 2000. Fetal age (post-conception weeks [PCWs]) was estimated using the independent measurement of the crown rump length (CRL) using the formula PCW (days) $= 0.9022 \times$ CRL (mm) $+ 27.372$.

Fetal tissue dissection and dissociation

Detailed methods for tissue dissection can be found in our previous publication.[35] In brief, for long bone samples, femora, tibiae, and fibulae were dissected from the fetal hind limbs by a specialist bone and soft tissue pathologist (P.B.) under a microscope using sterile microsurgical instruments. Each sample was then processed into single-cell suspensions. Other limb samples were processed whole (first-trimester whole-limb samples and second-trimester bone samples). The tissue was digested in a 5 μg/mL Liberase TH working solution prepared from Liberase TH powder (Sigma 5401135001) and 1× phosphate-buffered saline (PBS) on a shaking platform (750 rpm) at 37°C for 30 min. The tissue was gently agitated using a P1000 pipette after 15 min. 5 mL 2% fetal bovine serum (FBS) in PBS was then added to stop the dissociation prior to second-stage digestion with 0.25% trypsin solution for a further 30 min at 37°C, with pipette agitation every 5 min. Cells were then spun down at 750*g* at 4°C for 5 min and resuspended in 50–200 μL of 2% FBS in PBS. Fetal cells were then loaded for scRNA-seq).

Bulk RNA-seq of osteoarthritis cartilage samples and preprocessing

RNA was extracted using Qiagen AllPrep RNA Mini Kit, following the manufacturer's instructions as previously described in Steinberg et al.[10] Samples were collected and sequenced in 5 sequencing batches and processed uniformly. Poly(A)-tailed RNA was isolated from total RNA using Illumina's TruSeq RNA Sample Prep v.2 kits, while for cohort 5, a SMART-Seq Ultra Low Input RNA Kit was used. RNA fragmentation and library preparation were performed according to standard Illumina protocols. The libraries were sequenced on the Illumina HiSeq 2000 and HiSeq 4000 (75 bp paired ends) as well as Novaseq6000 (cohort 5), yielding a median of 80.8 million reads per sample (interquartile range [IQR]: 70,110,708–91,380,240). The raw data of batches 1–3 have been deposited to the European Genome/Phenome Archive (EGA) (EGA: EGAD00001005215, EGA: EGAD00001003355, and EGA: EGAD00001001331). Summary statistics of data quality across all sequencing batches are available in Table S1 and Figure S1.

We aligned the reads to the reference genome Ensembl GRCh38 release 105 (cDNA and non-coding) using STAR version 2.7.9a.[36] Reads were summarized to genes using featureCounts (Subread 2.0.3) software,[37] and different gene biotypes were evaluated.[29] We further projected STAR alignment files to the reference transcriptome using RSEM software[38] and generated the count matrix by aggregating the results on the gene level as previously done (https://github.com/broadinstitute/gtex-pipeline/blob/master/rnaseq/src/aggregate_rsem_results.py). We additionally marked optical duplicates using Picard Tools (https://broadinstitute.github.io/picard/) and dupRadar software.[39] Extensive quality control (QC) was additionally performed using RseQC v.4.0.0 software.[40] All the above steps were performed in a snakemake pipeline.[41] We performed an initial sample-level QC excluding samples that had an RNA integrity number <5, a percentage of uniquely mapped reads <70%, a percentage of reads mapping to genomic features <40%, a percentage of reads mapping to intronic regions or rRNA >30%, a percentage of reads that failed RseQC strandness check >30%, and <10 million reads. We excluded 77 samples according to the above criteria. For individuals with bilateral knee replacement, we excluded one pair of matched samples each (3 samples), keeping only the sample pair with the best quality (Figure S1).

For downstream exploratory analyses, the count matrix was normalized using trimmed mean of M (TMM) normalization.[42] We filtered lowly expressed genes using the *filterByExpr* function from edgeR.[43] We kept genes that were expressed with at least 10 counts for at least 158 samples (representing the 70% of the smallest group being high-grade osteoarthritis cartilage – 226 samples after initial sample-level QC), detecting expression for a total of 16,993 genes. Outlier samples were identified and excluded (18 samples) from further downstream analysis using robust principal-component analysis (PCA) and taking into consideration the first three principal components (PCs; PcaGrid method rrcov R package).[44] The final number of samples after the exclusions was 498, derived from a total of 300 individuals. We performed dimensionality reduction of the resulting count matrix using PCA and checked for the presence of batch effects. This analysis showed that PC1, explaining 28.91% of variation in the data, corresponds to sequencing batch, while PC2, explaining 11.01% of variation, corresponds to the cartilage degradation grade (Figures S2 and S3). Summary statistics of characteristics of the osteoarthritis cohort as well as aggregate data on individuals' demographic information and sample quality metrics included in the study after QC can be seen in Table S1. All the analyses after generation of the count matrix were performed using R v.4.1.3 and Bioconductor v.3.1.

scRNA-seq data from osteoarthritis cartilage

We obtained publicly available scRNA-seq data from Fan et al.,[32] accessed through NCBI's Gene Expression Omnibus (GEO) under accession ID GEO: GSE255460. Utilizing the integrated object provided by the authors, we re-annotated cell type clusters using broader labels to minimize the effects of highly correlated or rare cell types, which can compromise the accuracy of deconvolution methods.[45] We assessed data quality by examining key metrics, including total counts, gene numbers, and the percentage of mitochondrial gene counts (Figure S8). Cells selected for deconvolution analysis met criteria of having between 500 and 7,000 genes with a mitochondrial fraction below 0.25. After QC, a total of 62,611 cells were retained for analysis. Annotated cells, along with their representative gene markers and uniform manifold approximation and projection (UMAP) representations created using Scanpy,[46] are shown in Figures S9 and S10.

scRNA-seq of fetal tissues and preprocessing

Single-cell data were obtained from the fetal single-cell atlas of endochondral ossification.[35] This includes 18 long bone samples dissected from three fetuses aged 7–9 PCWs, embryonic limbs integrated with fetal limb data from another 14 embryos (5–9 PCWs),[47] and fetal bone marrow data from 9 embryos (12–19 PCWs).[48] These data are available via the European Nucleotide Archive (ENA) via the accession number ENA: PRJEB28278 (fetal long bone samples) and via ArrayExpress (first-trimester whole-limb samples: ArrayExpress: E-MTAB-8813; second-trimester bone marrow samples: ArrayExpress: E-MTAB-9389).

Briefly, single-cell suspensions were loaded onto a Chromium 10× Genomics single-cell 3′ v.2 library chip as per the manufacturer's protocol (10× Genomics; PN-120233), aiming for a cell capture recovery of 5,000 cells per channel. cDNA sequencing libraries were prepared according to the manufacturer's protocol and sequenced on an Illumina Hi-seq 4000 (2 × 50 bp paired-end reads).

Raw sequence reads in FASTQ format from fetal, pediatric, and organoid samples were processed and aligned to the

GRCh38-1.2.0 human reference transcriptome using the CellRanger v.3.0.2 pipeline (10× Genomics), available at http://cf.10xgenomics.com/supp/cell-exp/refdata-cellranger-GRCh38-3.0.0.tar.gz, with default parameters. The resulting expression matrices were processed with the SoupX package (v.1.3.0) for R to estimate and remove cell-free mRNA contamination prior to analysis with v.3 of the Seurat package.[49,50] In addition, each run was processed with the Scrublet pipeline,[51] with cells called as doublets subsequently removed from analysis.

Cells with fewer than 300 genes and greater than 7,500 genes expressed were filtered out, as well as those in which mitochondrial genes represented 10% or greater of total gene expression (Figure S12). To account for variations in cell cycle stage, Seurat's[50] *CellCycleScoring* function was applied on the remaining cells to produce a quantitative estimation of the cell cycle stage. Logarithmic normalization was then performed prior to data scaling, which used the cell cycle score, mitochondrial gene expression level, and total unique molecular identifiers (UMIs) per cell as regression variables.

The final developmental atlas included 249,151 cells spanning 5–17 PCWs.[35] For cell clustering, we used the top 2,000 highly variable genes, selected using Seurat's *FindVariableFeatures* function. Dimensionality reduction was performed using PCA on these highly variable genes. Cells were subsequently visualized using UMAP plots and clustering performed using the Seurat implementation of the Louvain algorithm (Figure S13). Differential gene expression testing between clusters was performed using the Wilcoxon rank-sum test within Seurat's *FindMarkers* function. Finally, cluster annotation was performed based on known cell-type-specific genes from the literature.

Power estimation of differential gene expression between osteoarthritis cartilage tissue grades

We estimated the statistical power to detect differential gene expression between low- and high-grade osteoarthritis cartilage along with sample size requirements using powsimR[52] (v.1.2.3). To this end, we simulated counts using a negative binomial distribution framework and estimated the mean-dispersion relationship from the data (count matrix containing 272 low-grade osteoarthritis cartilage samples). For parameter estimation, we used the total number of detected genes after QC (16,993), their respective lengths, and TMM value normalization[42] under the assumption of negative binomial distribution. We set up the simulations specifying the following parameters: (1) sample sizes ranging from 10 to 1,000 samples; (2) expected percentage of differentially expressed (DE) genes set to 30%, as determined by previous studies[10,53]; (3) effect size ($\log_2$ fold change [$\log_2$FC]) sampled from a narrow gamma distribution, also in agreement with previous observations[10,53]; and (4) mean, dispersion, and their relationship estimated from the low-grade cartilage count matrix. The library sizes for the simulations were sampled from our data. We performed 50 simulations of differential gene expression using the limma-voom[54] differential testing method and TMM normalization[42] under the assumption that 100% of replicates per group express the phenotypic differences in gene expression. To this end, we estimated the proportions and error rates at an alpha nominal level of 0.05 stratified by mean gene expression (mean $\log_2$ counts per million [$\log_2$CPMs]) and $\log_2$FC and evaluated the marginal and conditional ($\log_2$FC and mean expression) true positive rate (TPR) and false discovery rate (FDR) (Figure S4). We further evaluated the simulations using powsimR functions that report summary metrics such as the receiver operator characteristic (ROC) curve as well as the accuracy, F1 score, and Matthews correlation coefficient. This analysis showed that the FDR was controlled at the nominal level of 0.05 across all sample sizes (Figure S5).

Differential gene expression between osteoarthritis cartilage tissue grades

We performed differential expression analysis[55] between low- and high-grade cartilage using linear mixed models from limma blocking for the individual ID on filtered counts and voom[54] transformation (observational-level weights) as proposed for RNA-seq data.[54] The voom-limma pipeline was followed by applying the TREAT[56] (testing significance relative to a FC threshold) criteria (FC >1.2 in either direction) to calculate the t statistics, $\log_2$FC, and adjusted p values for all genes (Table S2). We further evaluated the distribution of DE genes among Ensembl[57] biotypes (Figure S6). We compared the results of differential expression analysis to those of 13 other RNA-seq and microarray studies that have performed similar analyses (for protein coding genes and long non-coding RNAs [lncRNAs]) with a smaller number of biological replicates after summarizing their respective results to the Ensembl[57] Gene stable IDs (see supplemental methods).

We conducted a differential gene expression analysis, adjusting for cell type proportions derived from the deconvolution analysis, to identify genes that are DE between conditions independent of variations in cell type composition (Table S4). The updated model for differential expression included all original covariates, with the addition of cell type proportions, except for the proportion of regulatory chondrocytes, which was excluded due to the introduction of collinearity. To validate our findings, we performed a correlation analysis comparing these results with those obtained from the original model, which adjusted for individual identity and sequencing batch. The effect sizes (logFC) and significance (adjusted p value) of genes found to be DE in both models were highly correlated, with Spearman ρ values of 0.93 for logFC and 0.8 for adjusted p values.

Enrichment analyses

Gene set enrichment was performed on the results of differential gene expression using fgsea (v.1.20.0) software,[58] ranking the genes by their effect size ($\log_2$FC). Overrepresentation analysis was performed separately for upregulated and downregulated DE genes using the *enrichr* and *enrichGO* (for Gene Ontology[59] biological processes) functions from ClusterProfiler[60] (v.4.2.2). Reactome[61] and Hallmark gene sets[62] were retrieved from the molecular signatures database (MSgDB v.7.5.1)[63] (Figure S7). For both enrichment approaches, we restricted our analyses to gene sets sized from 10 to 1,500 genes. For all overrepresentation analyses, the background was set to include all the genes detected after QC. Gene sets with an FDR-adjusted p value <0.05 were reported as significant.

Deconvolution analysis of bulk RNA-seq using osteoarthritis chondrocyte signatures

To identify chondrocyte cell types in low- and high-grade cartilage samples, we conducted a deconvolution analysis of bulk RNA-seq data using AutoGeneS (v.1.0.3)[64] and publicly available single-cell data from weight-bearing osteoarthritis cartilage provided by Fan et al.[32] After performing QC, we re-annotated the cell types using the integrated dataset to achieve a broader classification, informed

by cartilage literature and reported marker genes. Representative marker genes for each chondrocyte type are shown in Figure S6. Chondrocyte signatures were automatically identified from 4,000 highly variable genes using the function ag.optimize(ngen = 5,000, seed = 0, nfeatures = 400, mode = 'fixed', offspring_size = 100). The model was then applied to bulk RNA-seq data to estimate the proportions of specific cell types through regression analysis.

Analysis of shared transcriptional programs between osteoarthritis tissue and fetal osteochondral tissue

The *CellSignalAnalysis* python package was used to search for shared transcriptional programs between all fetal limb cell clusters (scRNA-seq) and bulk transcriptomes generated from osteoarthritis cartilage tissue as previously described[65] (Figure S14). In brief, this method models the bulk RNA-seq data as a weighted linear combination of the single-cell transcriptomes based on a constrained generalized linear model. In contrast to deconvolution, it does not quantify the cellular composition but rather estimates the major cellular signals (transcriptional programs) present in the bulk transcriptomes when the single-cell reference and bulk RNA-seq samples are different. This difference is quantified through the allocation of an "unexplained signal," which represents the model's intercept term and is used to limit spurious associations. The relative contribution of each single-cell reference subtype to the bulk transcriptomes is calculated per sample (optimized beta values of the generalized linear model), and library size normalization is then performed to make the estimations comparable across samples. Goodness of fit is calculated using the McFadden's pseudo R-squared value.[65] Results were visualized using the ggplot2 package for R.[66]

Construction of weighted gene co-expression network and module identification in osteoarthritis cartilage

We conducted weighted gene co-expression analysis[67] (WGCNA) so as to find genes that show coordinated up- or downregulation between low- and high-grade osteoarthritis cartilage. We constructed a signed co-expression network using gene expression data regressed for known batch effects (sequencing batch). Regressed gene expression was generated using the *removeBatchEffect* function from the limma package[68] (Figure S15). We selected a soft threshold power of 7 to ensure scale-free topology ($R^2 = 0.88$). We employed a robust version of WGCNA (rWGCNA) to minimize outlier influence on the network construction, as previously implemented[14] (see supplemental methods; Figure S16). Module eigengenes were defined for each module, representing the gene expression value for a module that is comparable to the module's PC1. We further calculated module membership (defined as kME) for each gene using the *signedKME* function from the WGCNA R package by providing the table of biweight midcorrelation between each gene and each module eigengene. Intramodular connectivity was further calculated using the *intramodularConnectivity* function from the WGCNA R package.

Functional characterization of co-expression modules

We evaluated the association of consensus module eigengenes (from regressed expression data with the removal of known technical covariates) with cartilage grade. To find modules that are significantly associated with osteoarthritis progression, we used linear mixed models blocking for the individual identifier as done for differential expression testing. A module was considered significantly associated with cartilage grade if it had an FDR-corrected p value <0.05. We further evaluated module correlation with other technical and biological covariates using Pearson correlation (Figure S17). We further performed enrichment analyses for Gene Ontology[59] biological processes using overrepresentation analysis as described for DE genes (Table S8), using the total number of genes used to construct the network as background. Modules were further tested for enrichment for knee-related osteoarthritis GWASs,[4] lists of genes with functional genomics evidence of involvement in osteoarthritis, microRNA, transcription factor (TF) targets[69] (Tables S6, S7, and S10; Figures S17 and S18), and protein-protein interaction networks (StringDB[70]) (see supplemental methods; Table S10). We further looked at protein-protein interactions mediated solely by physical association using the StringDB,[70] Intact,[71] and Bioplex[72] databases.

The enrichments for GWAS osteoarthritis traits, including knee osteoarthritis, osteoarthritis at any site, and total knee and total joint replacement, were assessed using MAGMA (v.1.10) software.[73] To this end, we calculated a gene-based association score based on the distance of the associated variant to the transcription start site of the genes present in the co-expression modules. MAGMA was run specifying a 35 kb upstream and 10 kb downstream window as previously done to capture both gene body and regulatory regions[74–76] and genome build 37 to match the GWAS summary statistics (see supplemental methods). p values underwent correction for multiple testing using across all modules and traits, and gene set-trait associations with an FDR-corrected p value $<5\%$ were defined as significant. The modules that showed significant enrichment for GWAS osteoarthritis traits were prioritized for further analyses (correlation between prioritized modules seen in Figure S19).

We employed LD score regression to partition heritability.[14,77] This approach allowed us to break down the heritability of the tested osteoarthritis phenotypes across distinct functional categories represented by the prioritized gene co-expression modules. Leveraging data from the GWAS summary statistics from the same meta-analysis[4] and ancestry-matched LD modeling from the 1000 Genomes reference panel,[34] we quantified the proportion of the overall genome-wide heritability that can be attributed to genetic variants within explicitly defined functional categories.[77] We calculated the partitioned heritability with stratified LD score regression using 10 kb windows around genes contained in each module as previously done.[14,77] Enrichment was also calculated as the ratio of SNP heritability attributed to each module compared to the proportion of total SNPs within that module. These categories were added to the full baseline model of 53 categories of broad regulatory functions described in Funicane et al.[77] to enhance model accuracy. The final model included all 53 categories and the six prioritized modules. Module enrichments with an FDR-corrected p value <0.05 (across traits and number of prioritized modules) were considered significant (Table S9).

To identify genes associated with mouse musculoskeletal (MSK) phenotypes upon perturbation in the prioritized modules, we utilized data from two curated databases, including the International Mouse Phenotyping Consortium (IMPC)[78] and the Rat Genome Database.[79] To define the MSK phenotypes, we incorporated the subsequent mouse knockout (KO) characteristics associated with MSK terms "bone," "muscle," "cartilage," "skeleton," "osteo," "arthritis," "limb," "muscular," "joint," "body size," "growth," "skeletal," "stature," and "height." To assess whether the observed number of genes associated with MSK phenotypes within each prioritized module was significantly higher than expected by chance,

we conducted a one-sided Fisher's exact test using the total number of genes in the co-expression network analysis as the background. Modules with an FDR-corrected p value <0.05 were reported to have significant enrichment.

Last, we checked the presence of genes targeted by drugs and further examined the ones in the "approved" or "approved investigational" phase among drugs in DrugBank[80] v.5.1.10 (Table S13).

Causal inference analysis

To assess the causality of the expression of the co-expressed genes within the six prioritized modules, we performed two-sample Mendelian randomization (MR) analyses[81] (Table S12). We used expression quantitative trait locus (eQTL) summary statistics from low-grade and high-grade cartilage[10] as the exposure and knee-related osteoarthritis GWAS summary statistics[4] as the outcome. We used the MR-Base-curated *TwoSampleMR* R package (v.0.5.7).[82] To define independent instrumental variables (IVs) for each gene in each of the studied tissues, we conducted LD-based clumping on the eQTL variants using a strict threshold of $R^2 = 0.001$ over a 10 Mb window on either side of the index variant with the European reference LD panel from 1000 Genomes.

To ensure that the selected IVs are strongly associated with the exposure, we selected only the ones with an F-statistic larger than ten, which was estimated as $(beta^2/se^2)$,[81] where *beta* is the effect size estimate and *se* the corresponding standard error. In the case of only one strong IV, we employed the Wald ratio method to infer causality. If more than one IV was available, then the inverse-variance-weighted (IVW) method was applied along with the weighted median and the MR-Egger methods. The latter was only applied if more than three independent IVs were available. To test for heterogeneity, we calculated the Q-statistics. For IVs not included in the outcome GWAS, we conducted an LD-based proxy search on the European ancestry 100 Genome reference panel using the R package *LDlinkR* (v.1.3.0). Finally, to define evidence of causality, we used the FDR-adjusted p value of the IVW or Wald ratio results.

Scoring the expression of osteoarthritis gene modules in fetal data and osteoarthritis adult cartilage

Module gene expression per single-cell population was calculated using the Seurat *AddModuleScore* function.[50,83] The genes of the prioritized modules were provided as input (six gene lists) along with the fetal single-cell Seurat object using default parameters. This function calculates the average expression of the module genes on a per-cell basis and then subtracts the aggregated expression of control gene sets. Briefly, first the average (aggregated) expression of all genes in the scRNA-seq dataset is calculated, and genes are ranked according to their average expression and binned into 25 bins (default). Then, 100 genes are extracted from the same bin for each gene that belongs to the gene sets of interest, here being the prioritized modules, to form the control set. Using this approach, two matrices are created: one containing the mean expression of control genes for each cell (control scores) and one containing the mean expression of all genes per module for each cell (feature scores). The control scores are then subtracted from the feature ones to get the final cell scores (close to zero difference corresponding to no enrichment of the genes in the modules in a cell). The average expression of gene modules in mesenchyme, Schwann cell precursors, and osteochondral tissue, together with classical marker genes, was visualized using the Seurat *DotPlot* function.

Results

Sample size and power

We collected paired (from within the same knee joint), macroscopically intact (low-grade), and degraded (high-grade) knee osteoarthritis cartilage from 300 individuals undergoing knee replacement surgery. We performed deep RNA-seq and compared the transcriptional profiles across 272 low-grade and 226 high-grade osteoarthritis cartilage samples (198 pairs isolated from the same knee joint) following QC (Figures 1 and S1–S3; Table S1; supplemental methods). Compared to earlier studies, this amounts to a more than 2-fold increase in sample size[10,53] and a concomitant jump in power. Here, we have the power to detect ~68% (compared to 55% of the largest study to date[53]) of all DE genes between low- and high-grade osteoarthritis cartilage at 5% FDR. We investigated the sample size required to attain 80% power to detect DE genes and found that more than 900 samples per cartilage group would be needed for an exhaustive characterization of transcriptomic differences in osteoarthritis (Figures S4 and S5). In addition to marginal power, we also estimated conditional power by stratifying genes based on their expected $\log_2$FCs, as effect size is a commonly used metric to evaluate biological relevance.[56] We find that, with the current sample size, we can confidently (>99% conditional power on effect size) detect differential expression for genes with effect sizes corresponding to a 1.2-fold change in either direction.

Transcriptional changes in primary osteoarthritis cartilage

We identify 3,304 genes to be upregulated and 3,485 genes to be downregulated in high-grade compared to low-grade osteoarthritis cartilage (FDR-adjusted $p < 0.05$, FC > 1.2) (Figures 2A, 2B, and S6; Table S2). Of these, 1,175 genes have not been previously reported to be DE in osteoarthritis cartilage (527 protein-coding genes and 479 lncRNAs)[10,53,84–95] (Table S2). We find enrichment of DE genes in biological pathways with links to osteoarthritis[10,53] (Figure S7; Table S3). The strongest enrichment among downregulated genes in high-grade cartilage was observed for embryonic morphogenesis (normalized enrichment score [NES] = −1.83, FDR-adjusted $p = 1.93 \times 10^{-6}$) and Gene Ontology terms related to cartilage and limb development (Figure S7). These findings suggest that chondrocytes could be undergoing a differentiation process similar to those taking place during limb development in the early stages of cartilage degradation.

Distinct chondrocyte dynamics, vascularization, and immune cell infiltration in osteoarthritis

To gain insights into how chondrocyte populations may differ between low- and high-grade cartilage, we performed a deconvolution analysis using scRNA-seq data from the cartilage of individuals with osteoarthritis[32]

Please cite this article in press as: Katsoula et al., Primary cartilage transcriptional signatures reflect cell-type-specific molecular pathways underpinning osteoarthritis, The American Journal of Human Genetics (2024), https://doi.org/10.1016/j.ajhg.2024.10.019

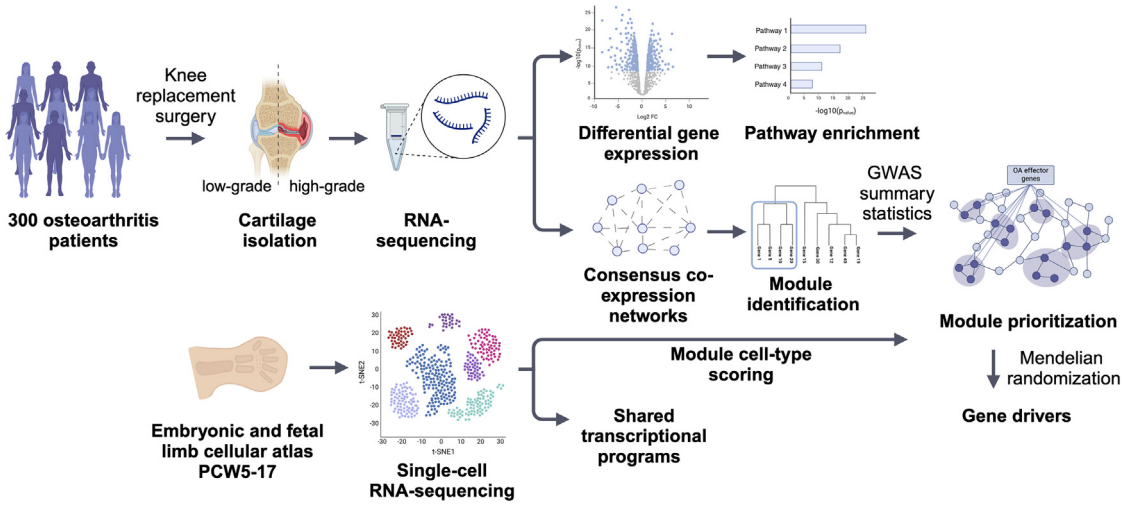

Figure 1. Overview of the study design

(Figures S8–S10). These samples were sourced from weight-bearing regions of articular cartilage, in line with the bulk RNA-seq data.

This analysis identified pre-fibrotic, pre-hypertrophic, and regulatory chondrocytes in both low- and high-grade cartilage, along with infiltrating macrophages and vascular endothelial cells (Figures 3A and 3B). These findings indicate that the differences observed between disease grades are likely due to variations in the proportions of these cell types. In particular, we find increased proportions of pre-fibrotic chondrocytes and macrophages in high-grade osteoarthritis cartilage compared to low-grade cartilage. This indicates that fibrocartilage formation (characterized by aberrant collagen expression) and inflammation are more prevalent in high-grade cartilage (Figure 3B). This finding is consistent with evidence pointing to the formation of cartilage of an inferior quality in an attempt to repair degenerated/injured articular cartilage.[96] Conversely, low-grade cartilage presents higher proportions of pre-hypertrophic chondrocytes, regulatory chondrocytes, and vascular endothelial cells. This suggests that less damaged cartilage is undergoing more active formation and maintenance processes, indicative of a healthier or less degenerated state. Regulatory chondrocytes play a role in maintaining cartilage homeostasis, responding to environmental cues, and potentially modulating inflammation and repair.[28] Meanwhile, the increased presence of vascular endothelial cells may point to neovascularization, a response to injury or inflammation.

We then used the cell type proportions estimated through deconvolution to identify DE genes and processes that are independent of cell type composition. We observed a reduction in the number of DE genes (3,747 vs. 6,789), which likely reflects the isolation of changes in gene regulation that occur consistently across cell types between low- and high-grade cartilage (Table S4). Upon exploring the biological processes enriched among these genes, we found nervous system processes, in addition to immune responses, to be upregulated in high-grade cartilage. This finding suggests that high-grade cartilage may exhibit increased neuroinflammatory activity, potentially contributing to disease progression and severity (Figure 3C). Downregulated genes in high-grade cartilage maintained enrichment for processes related to embryonic morphogenesis and limb development, along with regulation of protein synthesis (Figure 3C). This pattern may indicate that these developmental and maintenance pathways are active in the early stages of the disease across cell types and become downregulated as the disease progresses, possibly reflecting a shift in the tissue's developmental trajectory. Extracting the leading edge of the embryonic morphogenesis term, we find several genes with crucial functions in cartilage development and pattern formation, including HOX genes (*HOXA3*, *HOXA4*, *HOXA5*, *HOXA9*, *HOXA10*, *HOXA11*, *HOXA13*, *HOXD3*, *HOXD4*,*HOXA3*, *HOXD9*, *HOXD13*), SOX genes (*SOX6*, *SOX7*, *SOX8*, *SOX9*), and several growth factors (*GDF5*, *GDF7*, *TGFB3*, *BMP5*) (Figure S11). These genes are typically associated with developmental processes rather than healthy non-self-renewing adult cartilage, suggesting that their downregulation in high-grade cartilage could indicate a suppression of these developmental programs after being reactivated as a putative compensatory response at an earlier stage.

Shared transcriptional programs between limb development and osteoarthritis

There is some evidence to suggest that during osteoarthritis, chondrocytes undergo a phenotypic change toward states that resemble those found during endochondral ossification.[21,27,97] While the precise recapitulation of fetal cell states in osteoarthritis is not biologically likely, we postulated that common transcriptional programs (cellular signals) may exist between the two, driving common biological processes such as chondrocyte hypertrophy and osteoblastogenesis. The progression from healthy adult

Please cite this article in press as: Katsoula et al., Primary cartilage transcriptional signatures reflect cell-type-specific molecular pathways underpinning osteoarthritis, The American Journal of Human Genetics (2024), https://doi.org/10.1016/j.ajhg.2024.10.019

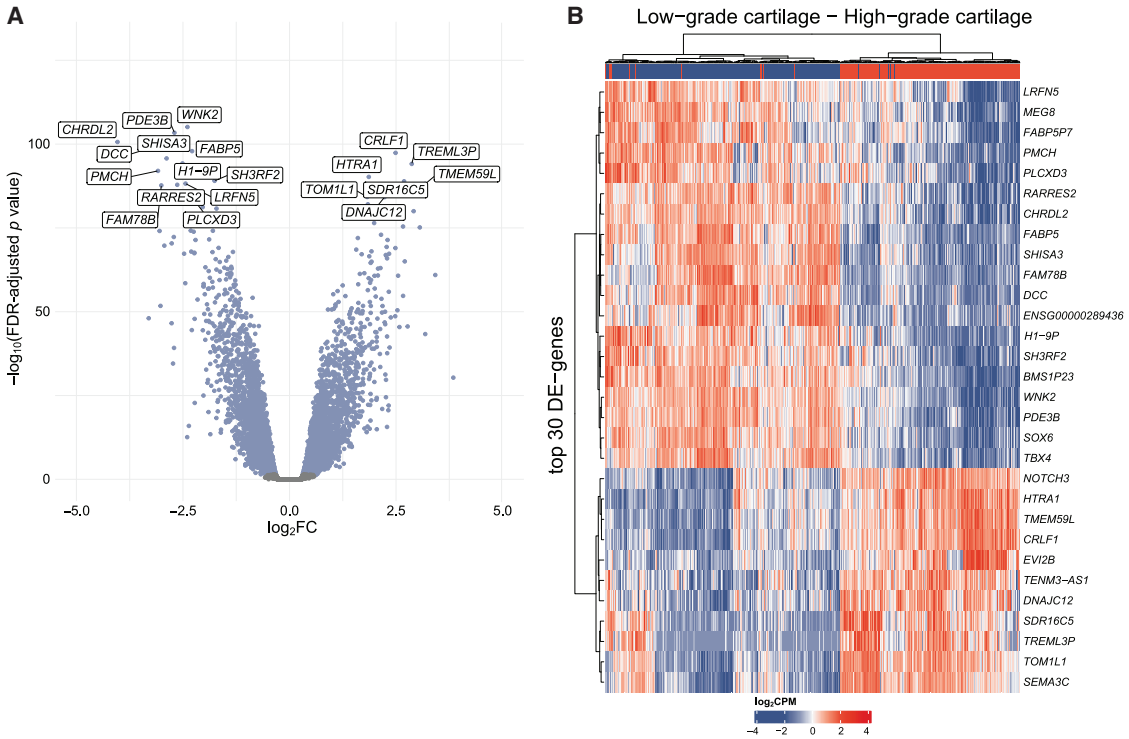

Figure 2. Differential gene expression between low- and high-grade osteoarthritis cartilage
(A) Volcano plot of differentially expressed (DE) genes between low- and high-grade osteoarthritis cartilage with respect to logarithmic fold change ($\log_2$FC) and significance (FDR-adjusted *p* value). Each dot represents a gene. Blue dots represent DE genes. Highlighted are the top 20 DE genes between low- and high-grade osteoarthritis cartilage.
(B) Heatmap of gene expression ($\log_2$CPM) of the top 30 DE genes between low- and high-grade osteoarthritis cartilage.

cartilage to early and then advanced osteoarthritis may be marked by an increasing presence of transcriptional programs associated with fetal endochondral ossification. We therefore applied the python package *CellSignalAnalysis*, trained on the fetal limb development scRNA-seq dataset derived from PCWs 5–17,[35] to the bulk RNA-seq data of osteoarthritis tissue. This fetal dataset encompasses cells from multiple lineages, such as lateral plate mesoderm (including osteochondral cells), glia, endothelial cells, striated and smooth muscle cells, adipose cells, and hematopoietic cells.[35]

We find that the largest fraction of the transcriptome in both low- and high-grade osteoarthritis cartilage is attributed to hypertrophic chondrocytes' transcriptional programs (Figure 4). Notably, in both tissues, the hypertrophic chondrocyte signatures are more pronounced compared to those of articular chondrocytes, which are the primary cell type in healthy articular cartilage.[98] This finding suggests that osteoarthritis chondrocytes have already started undergoing the process of hypertrophic differentiation before damage is macroscopically observed in early disease stages.

We next compared the cell type contribution between the bulk transcriptomes of low- and high-grade osteoarthritis cartilage. We find increased cellular signals from articular, hypertrophic, and pre-hypertrophic chondrocytes in low-grade compared to high-grade osteoarthritis cartilage. The increased articular chondrocyte cellular signals may reflect the lesser degree of degradation within the low-grade osteoarthritis cartilage. The elevated signals from hypertrophic and pre-hypertrophic chondrocytes indicate that chondrocytes in low-grade cartilage are undergoing a more active transition from pre-hypertrophic to hypertrophic states compared to those in high-grade cartilage. Conversely, early and mature osteoblast signals are increased in high-grade cartilage (Figure 4). Osteoblastic signatures have not been previously reported in osteoarthritis cartilage.[28–32] Considering that osteophyte formation and subchondral bone sclerosis are features of osteoarthritis progression, this suggests that osteoblasts have putative roles in these processes. Other fetal cell types contributed only minor signals to the overall bulk transcriptome, including fibroblasts, macrophages and adipogenic cells (Figure S14). There was no difference between the two categories in the fraction of the transcriptome that could not be explained by fetal cellular signals. Overall, these results support a key role of growth-plate transcriptional programs in the progression of osteoarthritis in cartilage.

A consensus co-expression network of osteoarthritis cartilage

To look deeper into the perturbed transcriptional networks during osteoarthritis progression, we studied coordinated gene deregulation and constructed a cartilage consensus gene co-expression network (Figures S15 and S16). We identify 55 gene co-expression modules or communities

Please cite this article in press as: Katsoula et al., Primary cartilage transcriptional signatures reflect cell-type-specific molecular pathways underpinning osteoarthritis, The American Journal of Human Genetics (2024), https://doi.org/10.1016/j.ajhg.2024.10.019

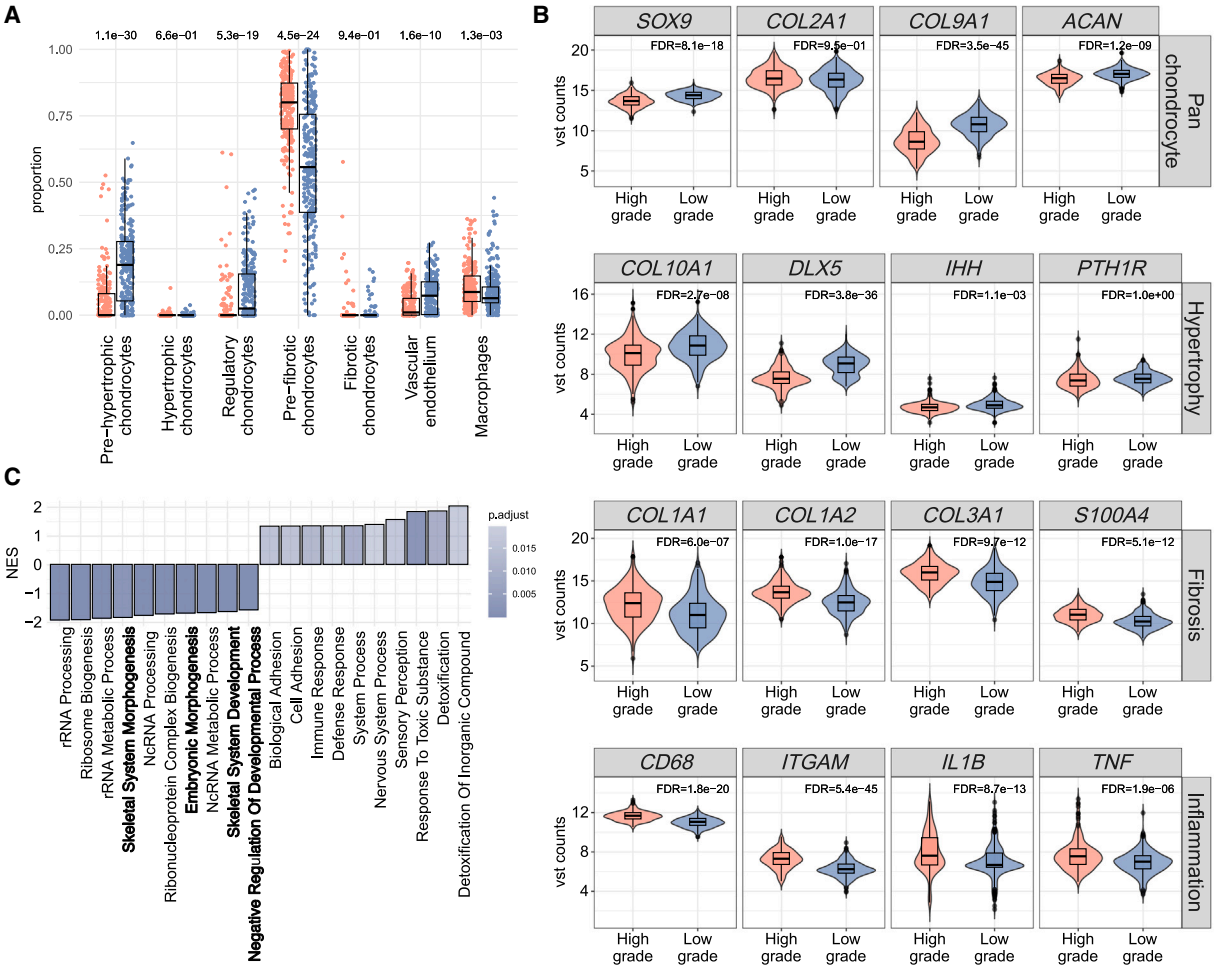

Figure 3. Differential cell type abundance between low- and high-grade cartilage

(A) Relative proportions of cell types present in bulk transcriptomes of high-grade and low-grade osteoarthritis cartilage samples obtained from deconvolution analysis (p values were generated using pairwise Wilcoxon rank-sum tests with FDR correction).

(B) Expression of representative marker genes used for cell type classification between low- and high-grade osteoarthritis cartilage (variance-stabilized-transformed [vst] gene expression counts after regression of the known batch effect).

(C) Bar plot of top 10 upregulated and top 10 downregulated Gene Ontology (GO) biological processes between low- and high-grade cartilage after adjustment for cell type proportions. Processes involved in skeletal system development and embryonic morphogenesis remain downregulated in high-grade cartilage after adjusting for cell type composition (top terms are highlighted in bold). p value is adjusted by FDR.

of genes with similar expression patterns (supplemental methods; Figures S16F and S16G; Table S6) (size ranging from 60 to 759 genes). Of these, 27 are upregulated and 24 are downregulated in high-grade osteoarthritis cartilage (linear mixed model, FDR-adjusted $p < 0.05$). We find that the association of these modules with cartilage grading (low-grade vs. high-grade osteoarthritis cartilage) is highly specific, with little to no association with biological and technical variables (Figure S17; Table S6). We examined potential regulatory relationships captured by the consensus co-expression by overlapping known curated TF-target interactions[69] with the identified modules. We found 488 matching edges reflecting captured TF-target interactions. These interactions spanned 34 modules (Figure S18).

To identify central transcriptional regulators of cartilage degeneration in knee osteoarthritis, we ranked module genes by their connectivity (both in the whole network and within each module) due to evidence pointing out that genes with high network connectivity are enriched for disease heritability.[11] We found that known osteoarthritis effector genes[4] had higher intramodular connectivity compared to all other genes in the network (Wilcoxon $p = 0.013$). In contrast, there was no significant difference in the inter-modular connectivity (Wilcoxon $p = 0.25$). This indicates that the modules and module-specific metrics capture the disease biological processes better than the whole network.

We next looked specifically at the most connected genes encoding for TFs within the identified modules. We find that several have important roles in chondrogenesis and embryonic limb morphogenesis (*DLX5*, *HOXA9*, *TBX4*, *HOXC10*) and cell differentiation/proliferation (*MYC*, *FOS*, *NFIA*, *ERG*, *FOXC1*, *FOXC2*, *FOXO1*, *FOXO3*) (Figure 5). *TBX4* and *ERG* have previously been identified

Please cite this article in press as: Katsoula et al., Primary cartilage transcriptional signatures reflect cell-type-specific molecular pathways underpinning osteoarthritis, The American Journal of Human Genetics (2024), https://doi.org/10.1016/j.ajhg.2024.10.019

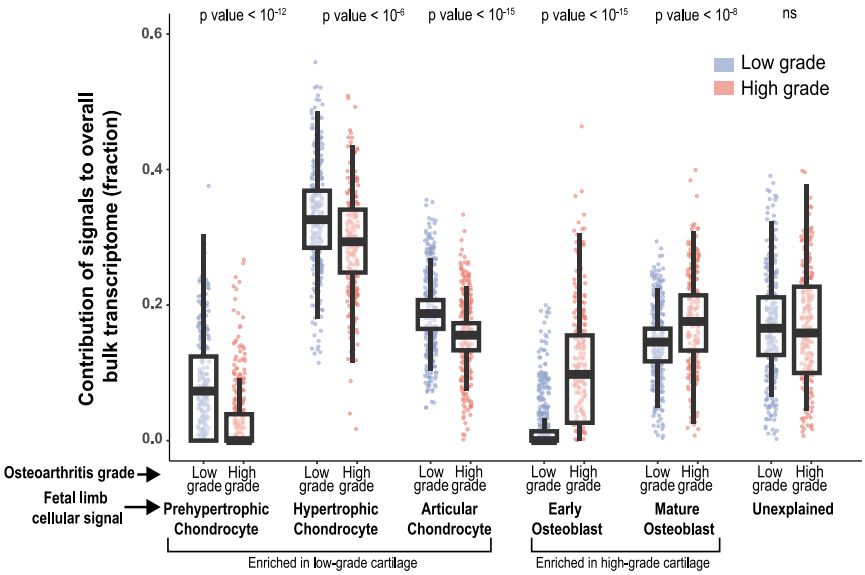

Figure 4. The relative contribution of single-cell-derived signals from fetal limb tissue in explaining the bulk transcriptomes of 226 high-grade and 272 low-grade osteoarthritis cartilage samples
The relative contribution of each signal to each bulk RNA-seq sample is shown on the y axis. Each signal/sample combination is represented by a single point and boxplots showing the distribution with median (middle line), first and third quartiles (box limits), and 1.5 times the interquartile range (whiskers). p values generated by the two-sided Wilcoxon rank-sum test of significant difference between the contribution of cellular signals to low- and high-grade osteorthritis cartilage transcriptomes are displayed above each pair of cell type plots. Color scale represents mean expression within a cluster. ns, non-significant.

as effector genes for osteoarthritis.[4] *TBX4* is known to be specific in hindlimb (posterior limb giving rise to legs) development in both mice and humans,[47,99] while *ERG* has been proposed to have key roles in the development and function of both transient and permanent cartilages.[100] *TBX4* is also among the hub genes of the module that showed the highest upregulation in low-grade cartilage (linear mixed model beta = −0.071, FDR-adjusted $p = 6.6 \times 10^{-69}$) and showed specific enrichment for processes related to embryonic morphogenesis and cartilage development. These results further support the role of developmental processes in osteoarthritis and provide evidence for the involvement of specific TFs and their regulatory networks.

Co-expression modules are enriched for osteoarthritis genetic risk signal

We prioritized modules capturing causal processes for osteoarthritis development by evaluating their enrichment for osteoarthritis risk variants within or in close proximity to module genes using MAGMA[73] (methods). To this end, we utilized summary statistics from the largest GWAS meta-analysis on osteoarthritis to date (826,690 individuals).[4] We identified significant enrichment for six modules (FDR-adjusted $p < 0.05$) among knee osteoarthritis, total knee, and total joint replacement phenotypes (Figures 6A and 6B; Table S7). This demonstrates that the average genetic association of genes with osteoarthritis in each of these modules is higher than the mean genetic association among all other genes in the captured gene space (gene drivers of this association within each module can be seen in Table S7). Out of the six prioritized modules, we find three to be upregulated and three to be downregulated in high-grade compared to low-grade osteoarthritis cartilage (Figures 6C and S17). Correlations between prioritized modules are shown in Figure S19.

We identify three modules enriched for knee osteoarthritis genetic risk (GWAS N = 396,054): the cartilage development module (top hub genes: *FAM78B*, *WNK2*, and *TBX4*; FDR-adjusted $p = 0.0006$), a module reflective of transcription regulation by RNA polymerase II (top hub genes: *SETD5*, *OTUD7B*, and *SRSF4*; FDR-adjusted $p = 0.011$), and a module representative of endoplasmic reticulum (ER) stress (top hub genes: *DNAJB11*, *SPCS2*, and *HM13*; FDR-adjusted $p = 0.012$) (Figure S17; Table S8). Upon closer inspection of the gene-gene interaction patterns within those modules, we find several captured TF-target relationships including regulons of known TFs, including *SMAD3*, *SIX4*, *YY1*, *MECP2*, and *NFE2L2* (transcription regulation module) and *NFIA* and *AR* (cartilage development module) (Figure S20).

Three modules show enrichment for the risk of total knee replacement surgery (GWAS N = 252,041): an ECM module (top hub genes: *MRC2*, *FMOD*, and *PLBD1*; FDR-adjusted $p = 0.046$), a module enriched for Wnt signaling genes (top hub genes: *DUSP2*, *TSC22D2*, and *MYC*; FDR-adjusted $p = 0.021$), and a module reflective of chromatin modification processes (top hub genes: *NIPBL*, *UBXN7*, and *PUM1*; FDR-adjusted $p = 0.021$) (Figure 6A). The ECM module is also enriched for genetic risk of total joint replacement (GWAS N = 368,576) (FDR-adjusted $p = 0.007$), with several fibrosis gene markers (*S100A4*, *THY1*, *FAP*, *COL1A2*, *COL3A1*, *FMOD*) in its co-expression network, suggesting that increased fibrosis is associated with knee/joint replacement risk.

In order to glean further insight into the genetic signals explained by the prioritized modules, we partitioned GWAS heritability[77] into the genes from each of the six modules (Figures 6D and 6E; Table S9). We find that the cartilage development module shows specific enrichment for knee osteoarthritis genetic risk (2.06-fold, FDR-adjusted $p = 2.29 \times 10^{-2}$), accounting for 1.70% of SNP-based heritability and containing 0.83% of the SNPs. The ECM

Please cite this article in press as: Katsoula et al., Primary cartilage transcriptional signatures reflect cell-type-specific molecular pathways underpinning osteoarthritis, The American Journal of Human Genetics (2024), https://doi.org/10.1016/j.ajhg.2024.10.019

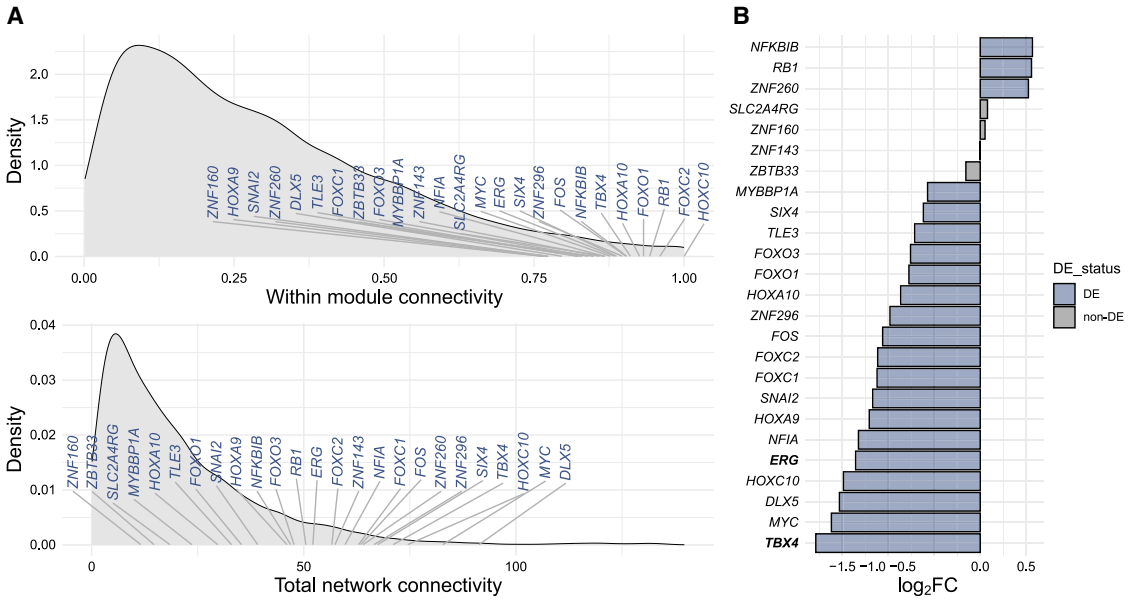

Figure 5. Network connectivity of genes encoding for transcription factors
(A) Within-module (intramodular) and between-module (inter-modular) connectivity based on consensus co-expression networks; top 25 genes encoding for transcription factors with the highest intramodular connectivity are labeled.
(B) Fold change of transcription factor genes shown in (A) between low- and high-grade cartilage. Highlighted in bold are known osteoarthritis effector genes.

module is enriched for knee replacement surgery risk (3.56-fold, FDR-adjusted $p = 8.41 \times 10^{-4}$), accounting for 3.12% of SNP-based heritability despite only containing 0.88% of the SNPs. We find that the transcription regulation module is enriched in common genetic variation for all three tested phenotypes (knee osteoarthritis: 1.94-fold, FDR-adjusted $p = 7.58 \times 10^{-3}$; total knee replacement: 2.14-fold, FDR-adjusted $p = 2.29 \times 10^{-2}$; total joint replacement: 1.90-fold, FDR-adjusted $p = 2.29 \times 10^{-2}$), accounting for >2.6% of SNP-based heritability across phenotypes (2.99% total knee replacement, 2.71% knee osteoarthritis, 2.66% total joint replacement) and containing 1.39% of SNPs. This module may be reflective of regulatory processes affecting the chondrocyte phenotype with a driving role in osteoarthritis. Our findings demonstrate the utility of gene co-expression networks in assigning common genetic variation patterns to particular biological functions.

We evaluated module enrichment for known osteoarthritis risk genes. These include curated lists of high-confidence effector genes (with three or more lines of evidence in support of their involvement in disease), putative effector genes (with at least one line of evidence) as defined by Boer et al.,[4] and genes with at least one eQTL in high- or low-grade cartilage[53] (Figure 6F; Table S10). The ECM and the cartilage development modules show the strongest enrichment for putative effector genes in osteoarthritis (FDR-adjusted $p = 0.002$ and 0.021, respectively). The ECM module is also the only one enriched for high-confidence osteoarthritis effector genes (FDR-adjusted $p = 1.9 \times 10^{-5}$), reflecting the important role of ECM remodeling in knee osteoarthritis pathogenesis.[101] Effector genes within the enriched modules include several TFs (*NFATC1*, *FOSB*, *CREBBP*, *TWIST1*, *SMAD3*) and TF targets (*CIC*, *LAMC1*, *KMT2D*, *COL27A1*, *TEAD1*, *HDAC7*) (Figure S20).

We investigated known protein-protein interactions from StringDB[70] and found significant enrichment within all six prioritized modules (FDR-adjusted $p < 0.05$). This indicates that the prioritized modules contain more interactions (edges) than one would expect from a random collection of proteins and is a further line of evidence that their gene products participate in common biological processes, providing a foundation for disentangling regulatory relationships.

Lastly, we investigated the presence of genes that show a MSK phenotype upon perturbation in the IMPC[78] and Rat Genome[79] databases (Table S11), and find that 41% of genes in all prioritized modules show a mouse MSK phenotype (transcription regulation: 168, chromatin modification: 154, ECM: 93, cartilage development: 70, Wnt signaling: 63, ER stress: 62). For five out of six modules (transcription regulation, chromatin modification, ECM, cartilage development, Wnt signaling) the observed number of genes associated with a mouse MSK phenotype was significantly higher than expected by chance (one-sided Fisher FDR adjusted $p < 0.05$, Figure 6F). This finding further supports the functional relevance of the identified modules for knee osteoarthritis.

To better understand the role of the prioritized transcriptional networks in osteoarthritis, we performed two-sample MR analysis in all eGenes (genes with at least one significant eQTL in either low- or high-grade cartilage tissue) within the six prioritized modules (methods). We identify 14 unique eGenes that have a causal association

Please cite this article in press as: Katsoula et al., Primary cartilage transcriptional signatures reflect cell-type-specific molecular pathways underpinning osteoarthritis, The American Journal of Human Genetics (2024), https://doi.org/10.1016/j.ajhg.2024.10.019

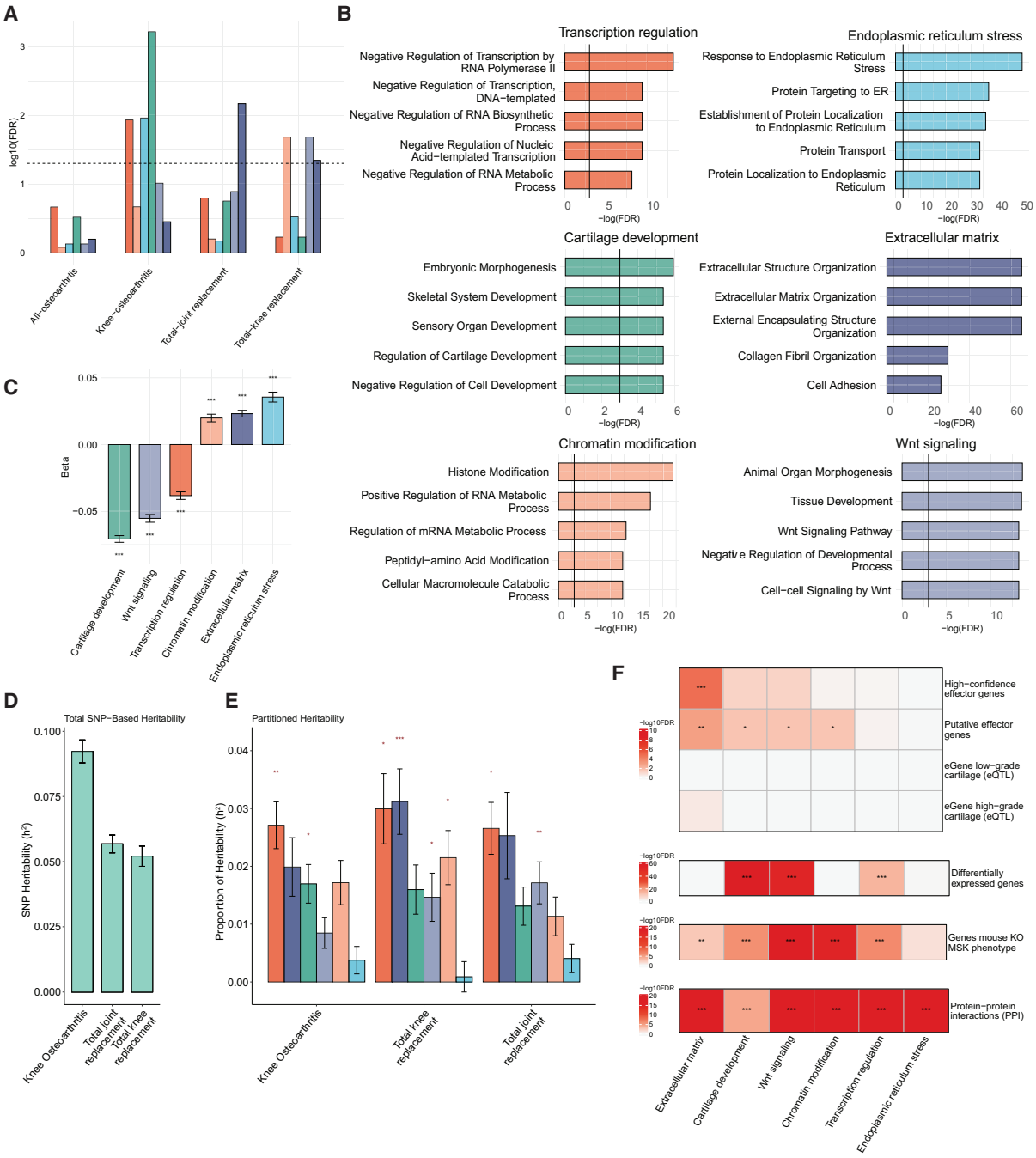

Figure 6. Co-expression modules that show enrichment for osteoarthritis GWAS signals and their functional annotation

(A) Significant enrichment is observed for knee osteoarthritis risk and total joint and total knee replacement surgery among 6 prioritized modules.

(B) Top 5 Gene Ontology biological processes for the prioritized modules depicted in (A).

(C) Module-level differential expression between low- and high-grade osteoarthritis cartilage. Bar plot of β values from linear mixed effect model of module eigengene association with disease status (FDR-corrected $p < 0.05$).

(D) Total SNP-based heritability (h^2: liability scale for osteoarthritis-relevant phenotypes) calculated from GWAS using LD score regression. Error bars represent the standard error of the mean (SEM) for total heritability.

(E) Partitioned heritability for each osteoarthritis trait that can be assigned to prioritized co-expression modules FDR-corrected $p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$. Error bars represent jackknife standard errors around the enrichment estimates.

(F) Enrichment $-\log_{10}$(FDR) (one-tailed Fisher test) for osteoarthritis risk genes (high-confidence effector genes, putatively causal genes, gene with an eQTL in low- or high-grade cartilage), differentially expressed genes, protein-protein interactions, and genes with a mouse musculoskeletal phenotype for the prioritized modules depicted in (A).

with osteoarthritis (Figure S21). Out of those, 10 are newly reported, as their eQTL regulation has not previously been causally linked to osteoarthritis in GWASs.[4] (Table S12). These are involved in immune response and cell signaling, enzymatic processes and metabolism, cellular processes and development, DNA repair, and protein trafficking. Their respective mouse model KO phenotypes[78,79,102] included skeletal/limb changes (*MIA3*, *P4HA2*), growth/body size (*MIA3*, *REV1*, *AGA*, *GABBR1*), aging/mortality (*RHBDL3*, *PAM*), neurological (*RHBDL3*, *GABBR1*, *AGA*, *CD4*), and the immune system (*CD4*). These results point to the relevance of the identified causal associations for these genes and their transcriptional networks in osteoarthritis biology.

Osteoarthritis co-expression modules are enriched for distinct cell types in the developing skeleton

In order to understand if the prioritized co-expression networks reflect the expression of specific cell types, we evaluated their enrichment among cell types found in both osteoarthritis cartilage and the developing fetal limb.[35] Through scoring module expression at the single-cell level (methods), we first examined their expression in cell populations identified in weight-bearing osteoarthritis cartilage.[32]

We found that that pre-hypertrophic and hypertrophic chondrocyte signatures are specifically expressed by the cartilage development module, indicating its specificity to these chondrocyte subtypes (Figure S22). Additionally, pre-hypertrophic chondrocytes' genes are expressed in the Wnt signaling, ECM, and ER stress modules, underscoring the importance of these pathways in regulating pre-hypertrophic chondrocytes during osteoarthritis. The ECM module is mainly expressed in pre-fibrotic and pre-hypertrophic chondrocyte subtypes, with little to no expression in regulatory chondrocytes and other non-chondrocyte populations (Figure S22). This pattern suggests that the ECM module is particularly relevant to cells actively engaged in cartilage matrix production and remodeling, emphasizing its specificity to matrix-producing chondrocytes in osteoarthritis.

By evaluating module expression in single-cell fetal data, we find that the six prioritized gene co-expression modules exhibit distinct expression patterns during the endochondral ossification developmental trajectory (Figure 7). The multipotent mesenchyme of the limb highly expresses both transcription regulation (low-grade osteoarthritis) and chromatin modification (high-grade osteoarthritis) gene modules, possibly reflecting the highly proliferative nature of these tissues (Figure 7). Two of the modules upregulated in low-grade osteoarthritis (cartilage development and Wnt signaling) are strongly and specifically expressed by developing chondrocytes (Figure 7). In particular, the cartilage development module, showing the largest upregulation in low-grade vs. high-grade osteoarthritis cartilage, presents increased activity across the chondroprogenitor → early resting → resting → proliferating → pre-hypertrophic chondrocyte trajectory. By contrast, two high-grade modules (ECM and ER stress) were weakly expressed by chondrocytes of the fetal limb tissue, with higher expression in fetal articular chondrocytes and perichondrium (ECM module only) and in osteoblasts (both modules). These findings suggest that changes in chondrocyte homeostasis during the early stages of cartilage degradation in osteoarthritis may share some biological similarities with the process of chondrocyte development during endochondral ossification,[27,103] with high-grade disease sharing similarities with bone remodeling, and with the replacement of cartilage with bone.[27]

Drug targets and repurposing opportunities

We find that 112 genes among the prioritized modules are targets of known drugs in the approved or approved investigational phase (targeted by 263 unique compounds from DrugBank[80]), presenting a clinical opportunity for drug repurposing in knee osteoarthritis (Table S13). Six out of 112 genes were also among the module genes driving the genetic association with osteoarthritis (MAGMA gene drivers; Table S7; Figure S23) (including among others the potassium calcium-activated channel subfamily M alpha 1 [*KCNMA1*], the phosphoinositide-3-kinase regulatory subunit 3 [*PIK3R3*], the complement factor H [*CFH*], and the apolipoprotein D [*APOD*]), and two were among the genes identified to be causally associated with osteoarthritis through MR (*CD4*, *GABBR1*) (Tables S13 and S14).

GABBR1 (transcription regulation module), which codes for a receptor of the gamma-aminobutyric acid (GABA) inhibitory neurotransmitter, is targeted by two approved drugs (baclofen and vigabatrin). Baclofen, a drug used for muscle spasticity, has a potential role in the treatment of collagen-induced arthritis (CIA) based on mouse experiments.[104] Vigabatrin has been found to attenuate chondrocyte hypertrophy and protect against injury-induced osteoarthritis in mice through the inhibition of 4-aminobutyrate aminotransferase.[105]

KCNMA1 (cartilage development module), coding for calcium-activated potassium channel subunit alpha-1, is abundantly expressed in articular cartilage and suggested to be involved in osteoarthritis progression. *KCNMA1* is targeted by five approved drugs (chlorzoxazone, bendroflumethiazide, cromoglicic acid, diazoxide, Trimebutine). Several calcium channels are also among the approved drug targets in the transcription regulation module. Specifically, *CACNB1*, *CACNB2*, and *CACNB4* are all targeted by Levomenthol, an organic compound used to treat minor pain. Levomenthol is currently in phase 2/3 for the treatment of pain in knee osteoarthritis.[80]

Isoprenaline, a β-adrenergic agonist targeting *PIK3R3* (transcription regulation module), is in an approved investigational stage for bradycardia and heart block.[80] Isoprenaline has also been found to inhibit interleukin (IL)-1β-mediated genotoxicity and ECM degradation by downregulating *MMP1*, *MMP3*, *MMP9*, and *ADAMTS5* in *in vitro* and *in vivo* osteoarthritis models.[106]

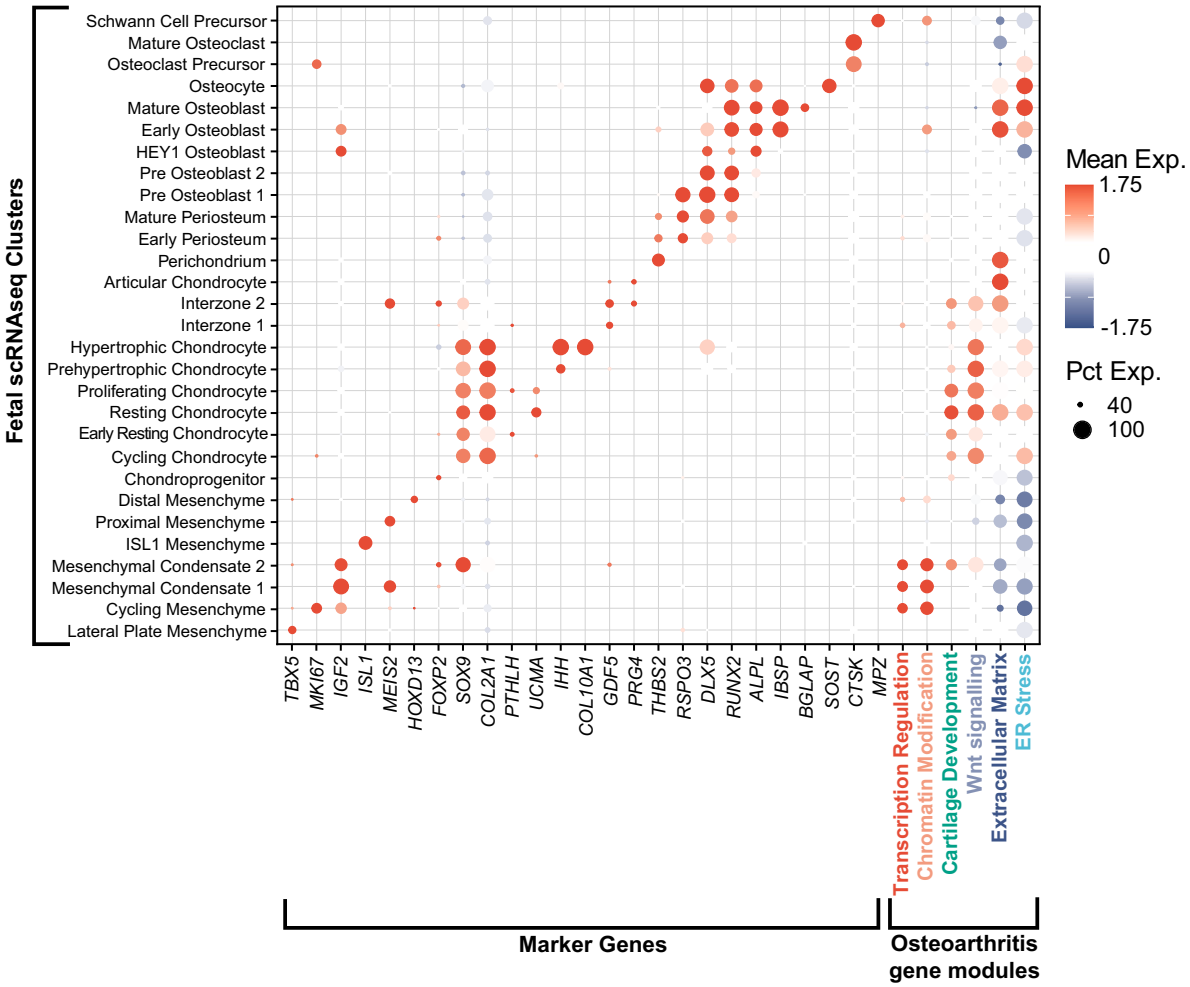

Figure 7. Dot plot of mean expression values of marker genes and osteoarthritis gene modules among cells of the developing fetal skeleton

Dot size corresponds to the percentage of cells with non-zero expression. The color scale represents the mean gene expression within a cluster. Exp, expression.

CFH and *APOD* (Wnt signaling module) are both targeted by copper (approved investigational phase), which is mainly used for supplementation. *CFH* is also targeted by zinc-containing inorganic compounds (zinc, zinc acetate, zinc chloride; approved investigational phase). Of note, genetically predicted high levels of both copper and zinc are positively associated with osteoarthritis, with the regulation of their levels having been proposed as a promising prevention strategy for osteoarthritis progression.[107] Therefore, each of these medications offers a potential avenue for repurposing studies in knee osteoarthritis.

Discussion

Large-scale functional genomics studies in osteoarthritis-relevant tissues, and the insights they can offer to improve our understanding of disease processes, have been lacking. Here, we analyze transcriptional alterations between paired low- and high-grade disease cartilage tissue samples from 300 individuals with knee osteoarthritis at unprecedented power. We identify transcriptional programs that are deregulated in disease and refine our understanding of its molecular pathophysiology by identifying coordinated gene expression changes. We find enrichment of osteoarthritis genetic risk in distinct co-expression modules, reflective of both specific biological processes and cell types in osteoarthritis cartilage as well as transcriptional similarities to the developing fetal limb. We find genetic risk enrichment of cartilage co-expression networks for knee osteoarthritis, total joint replacement, and knee replacement surgery and identify putative causal effects for variants linked to the expression of ten previously unreported genes in the largest osteoarthritis GWAS to date.[4]

One of the key findings of our study is the identification of fetal endochondral ossification transcriptional programs in adult osteoarthritis cartilage. This finding highlights the possibility of similar biological changes during the processes of endochondral ossification and osteoarthritis of

the knee, a concept that has recently gained attention. For example, Richard et al.[108] provide a model of human knee evolution, suggesting that genetic variation in functionally constrained regulatory elements plays a pivotal role in joint disease. These regulatory elements, which are essential for knee development, structure, and maintenance, have likely evolved under both positive and purifying selection pressures. However, modern genetic drift and pleiotropy may disrupt these constraints, contributing to an increased risk of osteoarthritis. Our findings align with this model, as we observe the presence of fetal transcriptional programs within prioritized modules (i.e., those enriched for osteoarthritis genetic risk), particularly noting that the module reflective of cartilage development exhibits the greatest genetic enrichment for knee osteoarthritis risk. This suggests that the proposed aberrant reactivation of knee developmental programs in the non-self-renewing adult cartilage may not only be a pathological hallmark but could also be a potential consequence of evolutionary pressures that have shaped these regulatory elements over time. Complementing these insights, Rice et al.[109] further demonstrated that osteoarthritis-associated methylation QTLs (mQTLs) are active during skeletogenesis, underscoring the developmental origins of osteoarthritis genetic risk and emphasizing the need to consider early molecular events in translational strategies.

Our findings may have broader implications for complex diseases. For example, research into neuropsychiatric disorders has shown that significant genetic risk is linked to prenatal developmental stages. A study by Walker et al.[110] identified numerous eQTLs and splicing QTLs (sQTLs) in mid-gestational human brains, revealing that genetic liabilities for SCZ and ASD are associated with prenatal-specific regulatory regions. Similarly, research into Alzheimer disease (AD) has demonstrated that patient-derived induced neurons exhibit a downregulation of mature neuronal properties and an upregulation of progenitor-like pathways, reflecting a hypo-mature neuronal identity.[111] Additionally, a study on systemic sclerosis (SSc) found that the loss of adipose tissue and its transition to myofibroblasts plays a crucial role in the development of dermal fibrosis, suggesting that similar cellular transitions might underlie other degenerative processes.[112] These examples underscore how developmental gene networks and cellular transitions impact disease susceptibility and progression.

Previous studies in mice have demonstrated that chondrocyte hypertrophic differentiation takes place in response to joint instability and that hypertrophic chondrocytes (marked by *COL10A1* and *MMP13* expression) may be involved in degradation of the cartilage matrix.[113–116] Building on these insights from mice, we provide large-scale molecular evidence that similar processes occurring during joint development are recapitulated during osteoarthritis in humans. Our study refines this understanding by pinpointing specific cell populations involved in cartilage degeneration. We find that increased pre-hypertrophic and hypertrophic transcriptional programs are already present at lower cartilage degradation (more pronounced than those of articular chondrocytes). This suggests that low-grade cartilage may undergo hypertrophic differentiation earlier than previously recognized. We further find that osteoblasts' transcriptional programs are enriched in high-grade cartilage, pointing to a putative transition to bone. This indicates that the change in cellular phenotype that occurs through osteoarthritis progression may be driven by biological mechanisms closely linked to those at play during endochondral ossification.

We find that gene co-expression modules that exhibit increased expression in low-grade osteoarthritis (cartilage development and Wnt signaling modules) are representative of transcriptional programs expressed in fetal resting, proliferating and pre-hypertrophic chondrocytes (both modules), and hypertrophic chondrocytes (Wnt signaling only). This is in agreement with ectopic initiation of chondrocyte hypertrophy playing a key role in osteoarthritis pathogenesis.[27] In particular, a module reflective of cartilage developmental processes is strongly enriched for knee osteoarthritis genetic risk and represents a transition from chondroprogenitors to the formation of pre-hypertrophic chondrocytes. Given the capacity of this population to initiate hypertrophic differentiation,[117] a process driving osteoarthritis cartilage remodeling,[97] this co-expression network likely captures critical transcriptional changes from the articular to the growth-plate transcriptional program. It is important to note that the presence of endochondral ossification programs does not fully elucidate the specific trajectories occurring during osteoarthritis progression. Evidence suggests that chondrocytes might also transdifferentiate into osteoblasts, potentially driving both cartilage degradation and ectopic bone formation.[118] This adds complexity to our understanding of chondrocyte behavior in osteoarthritis, indicating that multiple pathways are likely to contribute to disease progression and stressing the need for well-powered multi-omics studies.

In high-grade osteoarthritis cartilage, which represents more advanced degeneration, we find upregulation of the ECM module. The latter is mainly reflective of fibrosis and osteoblasts' transcriptional programs and shows enrichment for total knee and joint replacement risk. This indicates an association between changes in the ECM and more severe disease progression or pain, which is associated with surgical intervention. This finding is in line with the role of ECM degradation in knee osteoarthritis pathogenesis[2] and the replacement of cartilage with bone in advanced osteoarthritis stages.[27]

We identify drug repurposing opportunities for genes in the prioritized modules, including ion channels and adrenergic signaling targets. Ion channels have been proposed as attractive targets for pain management in individuals with osteoarthritis.[119] Adrenergic receptors (α2a and β2) and the inhibition of their catabolic effects in cartilage have been suggested to be an appealing approach for osteoarthritis treatment.[120]

We have compared low- and high-grade cartilage in individuals with end-stage osteoarthritis. As a result, the reported transcriptional variations may be distinct from the molecular alterations related to early changes associated with osteoarthritis susceptibility. In addition, we used samples from both the tibia and femur, which prevents us from determining if the identified fetal-like programs are more prevalent in one location or equally significant in both. Although the single-cell data analyzed here cover a wide range of gestational ages, they do not capture the full extent of limb development, mainly due to the ethical and logistical considerations of working with human embryonic and fetal tissue. Additionally, in the context of tissues, such as cartilage and bone, the breadth of single-cell capture can be limited, with a relatively low number of these cells sequenced following dissociation. This is due to the challenge of liberating single cells from matrix-rich tissue, which requires harsh enzymatic digestion and a delicate balance between liberation and over-treatment, resulting in cell stress and/or death. Nevertheless, the atlas used here does cover a significant portion of cellular differentiation in these lineages, successfully capturing cells ranging from mesenchymal progenitors to *COL10A1*-expressing hypertrophic chondrocytes and from perichondrial cells through to *BGLAP*-expressing osteoblasts. All samples in the current study were derived from individuals of European ancestry. Functional genomics data from diverse global populations will be necessary going forward.[121]

In this work, we generate a comprehensive transcriptional characterization of osteoarthritis primary cartilage tissue. The data enhance our understanding of gene networks enriched for the genetic risk of osteoarthritis, generate new insights into biological processes underpinning disease development and progression, and identify high-value drug targets and repurposing opportunities. Our findings demonstrate the value of integrating gene network analysis, GWASs, and single-cell data to disentangle key transcriptional and genetic regulators in osteoarthritis as an exemplar complex disease.

Data and code availability

Summary statistics of all analyses are shared through the Musculoskeletal Knowledge Portal (mskkp.org). The code accompanying the manuscript can be obtained online from https://hmgubox.helmholtz-muenchen.de/index.php/s/Manuscript_code.

Acknowledgments

The authors wish to thank Peter Kreitmaier and Iris Fischer for their helpful contributions and all study participants for their willingness to donate study samples. The work on osteoarthritis cartilage samples was funded by the Wellcome Trust (206194). The work on the embryonic and fetal scRNA-seq atlas was supported by the Wellcome Trust through institutional/program grants and personal fellowships (222902/Z/21/Z, 206194, 220540/Z/20/A, and 211276/Z/18/Z).

Declaration of interests

In the last three years, S.A.T. has been a remunerated consultant for Sanofi, Foresite Labs. and Qiagen and is a consultant and equity holder of TransitionBio and EnsoCell Therapeutics.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2024.10.019.

Received: May 4, 2024
Accepted: October 25, 2024
Published: November 22, 2024

References

1. Vos, T., Lim, S.S., Abbafati, C., Abbas, K.M., Abbasi, M., Abbasifard, M., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., Abdelalim, A., et al. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet *396*, 1204–1222. https://doi.org/10.1016/S0140-6736(20)30925-9.
2. Martel-Pelletier, J., Barr, A.J., Cicuttini, F.M., Conaghan, P.G., Cooper, C., Goldring, M.B., Goldring, S.R., Jones, G., Teichtahl, A.J., and Pelletier, J.-P. (2016). Osteoarthritis. Nat. Rev. Dis. Primers *2*, 16072. https://doi.org/10.1038/nrdp.2016.72.
3. Arden, N.K., Perry, T.A., Bannuru, R.R., Bruyère, O., Cooper, C., Haugen, I.K., Hochberg, M.C., McAlindon, T.E., Mobasheri, A., and Reginster, J.-Y. (2021). Non-surgical management of knee osteoarthritis: comparison of ESCEO and OARSI 2019 guidelines. Nat. Rev. Rheumatol. *17*, 59–66. https://doi.org/10.1038/s41584-020-00523-9.
4. 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., et al. (2021). Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations. Cell *184*, 4784–4818.e17. https://doi.org/10.1016/j.cell.2021.07.038.
5. Rice, S.J., Tselepi, M., Sorial, A.K., Aubourg, G., Shepherd, C., Almarza, D., Skelton, A.J., Pangou, I., Deehan, D., Reynard, L.N., and Loughlin, J. (2019). Prioritization of PLEC and GRINA as Osteoarthritis Risk Genes Through the Identification and Characterization of Novel Methylation Quantitative Trait Loci. Arthritis Rheumatol. *71*, 1285–1296. https://doi.org/10.1002/art.40849.
6. Coutinho de Almeida, R., Tuerlings, M., Ramos, Y., Den Hollander, W., Suchiman, E., Lakenberg, N., Nelissen, R.G.H.H., Mei, H., and Meulenbelt, I. (2023). Allelic expression imbalance in articular cartilage and subchondral bone refined genome-wide association signals in osteoarthritis. Rheumatology *62*, 1669–1676. https://doi.org/10.1093/rheumatology/keac498.
7. Rice, S.J., Aubourg, G., Sorial, A.K., Almarza, D., Tselepi, M., Deehan, D.J., Reynard, L.N., and Loughlin, J. (2018). Identification of a novel, methylation-dependent, RUNX2 regulatory region associated with osteoarthritis risk. Hum. Mol. Genet. *27*, 3464–3474. https://doi.org/10.1093/hmg/ddy257.
8. Kreitmaier, P., Suderman, M., Southam, L., Coutinho de Almeida, R., Hatzikotoulas, K., Meulenbelt, I., Steinberg, J.,

Relton, C.L., Wilkinson, J.M., and Zeggini, E. (2022). An epigenome-wide view of osteoarthritis in primary tissues. Am. J. Hum. Genet. *109*, 1255–1271. https://doi.org/10.1016/j.ajhg.2022.05.010.

9. Shepherd, C., Reese, A.E., Reynard, L.N., and Loughlin, J. (2019). Expression analysis of the osteoarthritis genetic susceptibility mapping to the matrix Gla protein gene MGP. Arthritis Res. Ther. *21*, 149. https://doi.org/10.1186/s13075-019-1934-7.
10. Steinberg, J., Southam, L., Roumeliotis, T.I., Clark, M.J., Jayasuriya, R.L., Swift, D., Shah, K.M., Butterfield, N.C., Brooks, R.A., McCaskie, A.W., et al. (2021). A molecular quantitative trait locus map for osteoarthritis. Nat. Commun. *12*, 1309. https://doi.org/10.1038/s41467-021-21593-7.
11. Kim, S.S., Dai, C., Hormozdiari, F., van de Geijn, B., Gazal, S., Park, Y., O'Connor, L., Amariuta, T., Loh, P.-R., Finucane, H., et al. (2019). Genes with High Network Connectivity Are Enriched for Disease Heritability. Am. J. Hum. Genet. *104*, 896–913. https://doi.org/10.1016/j.ajhg.2019.03.020.
12. Barabási, A.-L., Gulbahce, N., and Loscalzo, J. (2011). Network medicine: a network-based approach to human disease. Nat. Rev. Genet. *12*, 56–68. https://doi.org/10.1038/nrg2918.
13. Walker, J.T., Saunders, D.C., Rai, V., Chen, H.-H., Orchard, P., Dai, C., Pettway, Y.D., Hopkirk, A.L., Reihsmann, C.V., Tao, Y., et al. (2023). Genetic risk converges on regulatory networks mediating early type 2 diabetes. Nature *624*, 621–629. https://doi.org/10.1038/s41586-023-06693-2.
14. Gandal, M.J., Haney, J.R., Wamsley, B., Yap, C.X., Parhami, S., Emani, P.S., Chang, N., Chen, G.T., Hoftman, G.D., de Alba, D., et al. (2022). Broad transcriptomic dysregulation occurs across the cerebral cortex in ASD. Nature *611*, 532–539. https://doi.org/10.1038/s41586-022-05377-7.
15. Dear, R., Wagstyl, K., Seidlitz, J., Markello, R.D., Arnatkevičiūtė, A., Anderson, K.M., Bethlehem, R.A.I., Lifespan Brain Chart Consortium, Raznahan, A., Bullmore, E.T., and Vértes, P.E. (2024). Cortical gene expression architecture links healthy neurodevelopment to the imaging, transcriptomics and genetics of autism and schizophrenia. Nat. Neurosci. *27*, 1075–1086. https://doi.org/10.1038/s41593-024-01624-4.
16. Soul, J., Barter, M.J., Little, C.B., and Young, D.A. (2021). OA-Targets: a knowledge base of genes associated with osteoarthritis joint damage in animals. Ann. Rheum. Dis. *80*, 376–383. https://doi.org/10.1136/annrheumdis-2020-218344.
17. Aubourg, G., Rice, S.J., Bruce-Wootton, P., and Loughlin, J. (2022). Genetics of osteoarthritis. Osteoarthritis Cartilage *30*, 636–649. https://doi.org/10.1016/j.joca.2021.03.002.
18. Yang, L., Tsang, K.Y., Tang, H.C., Chan, D., and Cheah, K.S.E. (2014). Hypertrophic chondrocytes can become osteoblasts and osteocytes in endochondral bone formation. Proc. Natl. Acad. Sci. USA *111*, 12097–12102. https://doi.org/10.1073/pnas.1302703111.
19. Hallett, S.A., Matsushita, Y., Ono, W., Sakagami, N., Mizuhashi, K., Tokavanich, N., Nagata, M., Zhou, A., Hirai, T., Kronenberg, H.M., and Ono, N. (2021). Chondrocytes in the resting zone of the growth plate are maintained in a Wnt-inhibitory environment. Elife *10*, e64513. https://doi.org/10.7554/eLife.64513.
20. Haseeb, A., Kc, R., Angelozzi, M., de Charleroy, C., Rux, D., Tower, R.J., Yao, L., Pellegrino da Silva, R., Pacifici, M., Qin, L., and Lefebvre, V. (2021). SOX9 keeps growth plates and articular cartilage healthy by inhibiting chondrocyte dedifferentiation/osteoblastic redifferentiation. Proc. Natl. Acad. Sci. USA *118*, e2019152118. https://doi.org/10.1073/pnas.2019152118.
21. Singh, P., Marcu, K.B., Goldring, M.B., and Otero, M. (2019). Phenotypic instability of chondrocytes in osteoarthritis: on a path to hypertrophy. Ann. N. Y. Acad. Sci. *1442*, 17–34. https://doi.org/10.1111/nyas.13930.
22. van der Kraan, P.M., and van den Berg, W.B. (2012). Chondrocyte hypertrophy and osteoarthritis: role in initiation and progression of cartilage degeneration? Osteoarthritis Cartilage *20*, 223–232. https://doi.org/10.1016/j.joca.2011.12.003.
23. Hosaka, Y., Saito, T., Sugita, S., Hikata, T., Kobayashi, H., Fukai, A., Taniguchi, Y., Hirata, M., Akiyama, H., Chung, U.i., and Kawaguchi, H. (2013). Notch signaling in chondrocytes modulates endochondral ossification and osteoarthritis development. Proc. Natl. Acad. Sci. USA *110*, 1875–1880. https://doi.org/10.1073/pnas.1207458110.
24. Saito, T., Fukai, A., Mabuchi, A., Ikeda, T., Yano, F., Ohba, S., Nishida, N., Akune, T., Yoshimura, N., Nakagawa, T., et al. (2010). Transcriptional regulation of endochondral ossification by HIF-2alpha during skeletal growth and osteoarthritis development. Nat. Med. *16*, 678–686. https://doi.org/10.1038/nm.2146.
25. Fuerst, M., Bertrand, J., Lammers, L., Dreier, R., Echtermeyer, F., Nitschke, Y., Rutsch, F., Schäfer, F.K.W., Niggemeyer, O., Steinhagen, J., et al. (2009). Calcification of articular cartilage in human osteoarthritis. Arthritis Rheum. *60*, 2694–2703. https://doi.org/10.1002/art.24774.
26. Sun, M.M.-G., and Beier, F. (2014). Chondrocyte hypertrophy in skeletal development, growth, and disease. Birth Defects Res. C Embryo Today. *102*, 74–82. https://doi.org/10.1002/bdrc.21062.
27. Pitsillides, A.A., and Beier, F. (2011). Cartilage biology in osteoarthritis–lessons from developmental biology. Nat. Rev. Rheumatol. *7*, 654–663. https://doi.org/10.1038/nrrheum.2011.129.
28. Ji, Q., Zheng, Y., Zhang, G., Hu, Y., Fan, X., Hou, Y., Wen, L., Li, L., Xu, Y., Wang, Y., and Tang, F. (2019). Single-cell RNA-seq analysis reveals the progression of human osteoarthritis. Ann. Rheum. Dis. *78*, 100–110. https://doi.org/10.1136/annrheumdis-2017-212863.
29. Lv, Z., Han, J., Li, J., Guo, H., Fei, Y., Sun, Z., Dong, J., Wang, M., Fan, C., Li, W., et al. (2022). Single cell RNA-seq analysis identifies ferroptotic chondrocyte cluster and reveals TRPV1 as an anti-ferroptotic target in osteoarthritis. EBioMedicine *84*, 104258. https://doi.org/10.1016/j.ebiom.2022.104258.
30. Chou, C.-H., Jain, V., Gibson, J., Attarian, D.E., Haraden, C.A., Yohn, C.B., Laberge, R.-M., Gregory, S., and Kraus, V.B. (2020). Synovial cell cross-talk with cartilage plays a major role in the pathogenesis of osteoarthritis. Sci. Rep. *10*, 10868. https://doi.org/10.1038/s41598-020-67730-y.
31. Swahn, H., Li, K., Duffy, T., Olmer, M., D'Lima, D.D., Mondala, T.S., Natarajan, P., Head, S.R., and Lotz, M.K. (2023). Senescent cell population with ZEB1 transcription factor as its main regulator promotes osteoarthritis in cartilage and meniscus. Ann. Rheum. Dis. *82*, 403–415. https://doi.org/10.1136/ard-2022-223227.
32. Fan, Y., Bian, X., Meng, X., Li, L., Fu, L., Zhang, Y., Wang, L., Zhang, Y., Gao, D., Guo, X., et al. (2024). Unveiling inflammatory and prehypertrophic cell populations as key

contributors to knee cartilage degeneration in osteoarthritis using multi-omics data integration. Ann. Rheum. Dis. *83*, 926–944. https://doi.org/10.1136/ard-2023-224420.

33. van den Borne, M.P.J., Raijmakers, N.J.H., Vanlauwe, J., Victor, J., de Jong, S.N., Bellemans, J., Saris, D.B.F.; and International Cartilage Repair Society (2007). International Cartilage Repair Society (ICRS) and Oswestry macroscopic cartilage evaluation scores validated for use in Autologous Chondrocyte Implantation (ACI) and microfracture. Osteoarthritis Cartilage *15*, 1397–1402. https://doi.org/10.1016/j.joca.2007.05.005.
34. 1000 Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for human genetic variation. Nature *526*, 68–74. https://doi.org/10.1038/nature15393.
35. Lawrence, J.E.G., Woods, S., Roberts, K., Sumanaweera, D., Balogh, P., Predeus, A.V., He, P., Li, T., Polanski, K., Prigmore, E., et al. (2023). Single cell transcriptomics reveals chondrocyte differentiation dynamics in vivo and in vitro. Preprint at bioRxiv. https://doi.org/10.1101/2023.12.20.572425.
36. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.
37. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.
38. Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinf. *12*, 323. https://doi.org/10.1186/1471-2105-12-323.
39. Sayols, S., Scherzinger, D., and Klein, H. (2016). dupRadar: a Bioconductor package for the assessment of PCR artifacts in RNA-Seq data. BMC Bioinf. *17*, 428. https://doi.org/10.1186/s12859-016-1276-2.
40. Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. Bioinformatics *28*, 2184–2185. https://doi.org/10.1093/bioinformatics/bts356.
41. Mölder, F., Jablonski, K.P., Letcher, B., Hall, M.B., Tomkins-Tinch, C.H., Sochat, V., Forster, J., Lee, S., Twardziok, S.O., Kanitz, A., et al. (2021). Sustainable data analysis with Snakemake. F1000Res. *10*, 33. https://doi.org/10.12688/f1000research.29032.2.
42. Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. *11*, R25. https://doi.org/10.1186/gb-2010-11-3-r25.
43. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140. https://doi.org/10.1093/bioinformatics/btp616.
44. Chen, X., Zhang, B., Wang, T., Bonni, A., and Zhao, G. (2020). Robust principal component analysis for accurate outlier sample detection in RNA-Seq data. BMC Bioinf. *21*, 269. https://doi.org/10.1186/s12859-020-03608-0.
45. Heumos, L., Schaar, A.C., Lance, C., Litinetskaya, A., Drost, F., Zappia, L., Lücken, M.D., Strobl, D.C., Henao, J., Curion, F., et al. (2023). Best practices for single-cell analysis across modalities. Nat. Rev. Genet. *24*, 550–572. https://doi.org/10.1038/s41576-023-00586-w.
46. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15. https://doi.org/10.1186/s13059-017-1382-0.
47. Zhang, B., He, P., Lawrence, J.E.G., Wang, S., Tuck, E., Williams, B.A., Roberts, K., Kleshchevnikov, V., Mamanova, L., Bolt, L., et al. (2023). A human embryonic limb cell atlas resolved in space and time. Nature, 1–11. https://doi.org/10.1038/s41586-023-06806-x.
48. Jardine, L., Webb, S., Goh, I., Quiroga Londoño, M., Reynolds, G., Mather, M., Olabi, B., Stephenson, E., Botting, R.A., Horsfall, D., et al. (2021). Blood and immune development in human fetal bone marrow and Down syndrome. Nature *598*, 327–331. https://doi.org/10.1038/s41586-021-03929-x.
49. Young, M.D., and Behjati, S. (2020). SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. GigaScience *9*, giaa151. https://doi.org/10.1093/gigascience/giaa151.
50. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.
51. Wolock, S.L., Lopez, R., and Klein, A.M. (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. Cell Syst. *8*, 281–291.e9. https://doi.org/10.1016/j.cels.2018.11.005.
52. Vieth, B., Ziegenhain, C., Parekh, S., Enard, W., and Hellmann, I. (2017). powsimR: power analysis for bulk and single cell RNA-seq experiments. Bioinformatics *33*, 3486–3488. https://doi.org/10.1093/bioinformatics/btx435.
53. Katsoula, G., Steinberg, J., Tuerlings, M., Coutinho de Almeida, R., Southam, L., Swift, D., Meulenbelt, I., Wilkinson, J.M., and Zeggini, E. (2022). A molecular map of long non-coding RNA expression, isoform switching and alternative splicing in osteoarthritis. Hum. Mol. Genet. *31*, 2090–2105. https://doi.org/10.1093/hmg/ddac017.
54. Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. *15*, R29. https://doi.org/10.1186/gb-2014-15-2-r29.
55. Law, C.W., Alhamdoosh, M., Su, S., Dong, X., Tian, L., Smyth, G.K., and Ritchie, M.E. (2016). RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. F1000Res *5*, ISCBCommJ-1408. https://doi.org/10.12688/f1000research.9005.3.
56. McCarthy, D.J., and Smyth, G.K. (2009). Testing significance relative to a fold-change threshold is a TREAT. Bioinformatics *25*, 765–771. https://doi.org/10.1093/bioinformatics/btp053.
57. 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., et al. (2022). Ensembl 2022. Nucleic Acids Res. *50*, D988–D995. https://doi.org/10.1093/nar/gkab1049.
58. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2021). Fast gene set enrichment analysis. Preprint at bioRxiv. https://doi.org/10.1101/060012.

59. The Gene Ontology Consortium (2019). The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res. *47*, D330–D338. https://doi.org/10.1093/nar/gky1055.
60. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation *2*, 100141. https://doi.org/10.1016/j.xinn.2021.100141.
61. Gillespie, M., Jassal, B., Stephan, R., Milacic, M., Rothfels, K., Senff-Ribeiro, A., Griss, J., Sevilla, C., Matthews, L., Gong, C., et al. (2022). The reactome pathway knowledgebase 2022. Nucleic Acids Res. *50*, D687–D692. https://doi.org/10.1093/nar/gkab1028.
62. Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. *1*, 417–425. https://doi.org/10.1016/j.cels.2015.12.004.
63. Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics *27*, 1739–1740. https://doi.org/10.1093/bioinformatics/btr260.
64. Aliee, H., and Theis, F.J. (2021). AutoGeneS: Automatic gene selection using multi-objective optimization for RNA-seq deconvolution. Cell Syst. *12*, 706–715.e4. https://doi.org/10.1016/j.cels.2021.05.006.
65. Young, M.D., Mitchell, T.J., Custers, L., Margaritis, T., Morales-Rodriguez, F., Kwakwa, K., Khabirova, E., Kildisiute, G., Oliver, T.R.W., de Krijger, R.R., et al. (2021). Single cell derived mRNA signals across human kidney tumors. Nat. Commun. *12*, 3896. https://doi.org/10.1038/s41467-021-23949-5.
66. Wickham, H. (2009). ggplot2, Voice in Settings: Elegant Graphics for Data Analysis (Springer-Verlag). https://doi.org/10.1007/978-0-387-98141-3.
67. Zhang, B., and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. Stat. Appl. Genet. Mol. Biol. *4*, Article17. https://doi.org/10.2202/1544-6115.1128.
68. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47. https://doi.org/10.1093/nar/gkv007.
69. Müller-Dott, S., Tsirvouli, E., Vazquez, M., Ramirez Flores, R.O., Badia-I-Mompel, P., Fallegger, R., Türei, D., Lægreid, A., and Saez-Rodriguez, J. (2023). Expanding the coverage of regulons from high-confidence prior knowledge for accurate estimation of transcription factor activities. Nucleic Acids Res. *51*, 10934–10949. https://doi.org/10.1093/nar/gkad841.
70. Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N.T., Legeay, M., Fang, T., Bork, P., et al. (2021). The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. *49*, D605–D612. https://doi.org/10.1093/nar/gkaa1074.
71. del Toro, N., Shrivastava, A., Ragueneau, E., Meldal, B., Combe, C., Barrera, E., Perfetto, L., How, K., Ratan, P., Shirodkar, G., et al. (2022). The IntAct database: efficient access to fine-grained molecular interaction data. Nucleic Acids Res. *50*, D648–D653. https://doi.org/10.1093/nar/gkab1006.
72. Geistlinger, L., Vargas, R., Jr., Lee, T., Pan, J., Huttlin, E.L., and Gentleman, R. (2023). BioPlexR and BioPlexPy: integrated data products for the analysis of human protein interactions. Bioinformatics *39*, btad091. https://doi.org/10.1093/bioinformatics/btad091.
73. de Leeuw, C.A., Mooij, J.M., Heskes, T., and Posthuma, D. (2015). MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLoS Comput. Biol. *11*, e1004219. https://doi.org/10.1371/journal.pcbi.1004219.
74. Pergola, G., Parihar, M., Sportelli, L., Bharadwaj, R., Borcuk, C., Radulescu, E., Bellantuono, L., Blasi, G., Chen, Q., Kleinman, J.E., et al. (2023). Consensus molecular environment of schizophrenia risk genes in coexpression networks shifting across age and brain regions. Sci. Adv. *9*, eade2812. https://doi.org/10.1126/sciadv.ade2812.
75. Coleman, J.R.I., Bryois, J., Gaspar, H.A., Jansen, P.R., Savage, J.E., Skene, N., Plomin, R., Muñoz-Manchado, A.B., Linnarsson, S., Crawford, G., et al. (2019). Biological annotation of genetic loci associated with intelligence in a meta-analysis of 87,740 individuals. Mol. Psychiatry *24*, 182–197. https://doi.org/10.1038/s41380-018-0040-6.
76. Bryois, J., Skene, N.G., Hansen, T.F., Kogelman, L.J.A., Watson, H.J., Liu, Z., Eating Disorders Working Group of the Psychiatric Genomics Consortium; International Headache Genetics Consortium; and 23andMe Research Team, and Brueggeman, L., et al. (2020). Genetic identification of cell types underlying brain complex traits yields insights into the etiology of Parkinson's disease. Nat. Genet. *52*, 482–493. https://doi.org/10.1038/s41588-020-0610-9.
77. Finucane, H.K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.-R., Anttila, V., Xu, H., Zang, C., Farh, K., et al. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat. Genet. *47*, 1228–1235. https://doi.org/10.1038/ng.3404.
78. Groza, T., Gomez, F.L., Mashhadi, H.H., Muñoz-Fuentes, V., Gunes, O., Wilson, R., Cacheiro, P., Frost, A., Keskivali-Bond, P., Vardal, B., et al. (2023). The International Mouse Phenotyping Consortium: comprehensive knockout phenotyping underpinning the study of human disease. Nucleic Acids Res. *51*, D1038–D1045. https://doi.org/10.1093/nar/gkac972.
79. Smith, J.R., Hayman, G.T., Wang, S.-J., Laulederkind, S.J.F., Hoffman, M.J., Kaldunski, M.L., Tutaj, M., Thota, J., Nalabolu, H.S., Ellanki, S.L.R., et al. (2020). The Year of the Rat: The Rat Genome Database at 20: a multi-species knowledgebase and analysis platform. Nucleic Acids Res. *48*, D731–D742. https://doi.org/10.1093/nar/gkz1041.
80. Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., et al. (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. *46*, D1074–D1082. https://doi.org/10.1093/nar/gkx1037.
81. Sanderson, E., Glymour, M.M., Holmes, M.V., Kang, H., Morrison, J., Munafò, M.R., Palmer, T., Schooling, C.M., Wallace, C., Zhao, Q., and Smith, G.D. (2022). Mendelian randomization. Nat. Rev. Methods Primers *2*, 6. https://doi.org/10.1038/s43586-021-00092-5.
82. Hemani, G., Zheng, J., Elsworth, B., Wade, K.H., Haberland, V., Baird, D., Laurin, C., Burgess, S., Bowden, J., Langdon, R., et al. (2018). The MR-Base platform supports systematic causal inference across the human phenome. Elife *7*, e34408. https://doi.org/10.7554/eLife.34408.
83. Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy,

G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189–196. https://doi.org/10.1126/science.aad0501.

84. van Hoolwerff, M., Metselaar, P.I., Tuerlings, M., Suchiman, H.E.D., Lakenberg, N., Ramos, Y.F.M., Cats, D., Nelissen, R.G.H.H., Broekhuis, D., Mei, H., et al. (2020). Elucidating Epigenetic Regulation by Identifying Functional cis-Acting Long Noncoding RNAs and Their Targets in Osteoarthritic Articular Cartilage. Arthritis Rheumatol. *72*, 1845–1854. https://doi.org/10.1002/art.41396.
85. Li, H., Yang, H.H., Sun, Z.G., Tang, H.B., and Min, J.K. (2019). Whole-transcriptome sequencing of knee joint cartilage from osteoarthritis patients. Bone Joint Res. *8*, 290–303. https://doi.org/10.1302/2046-3758.87.BJR-2018-0297.R1.
86. Coutinho de Almeida, R., Ramos, Y.F.M., Mahfouz, A., den Hollander, W., Lakenberg, N., Houtman, E., van Hoolwerff, M., Suchiman, H.E.D., Rodríguez Ruiz, A., Slagboom, P.E., et al. (2019). RNA sequencing data integration reveals an miRNA interactome of osteoarthritis cartilage. Ann. Rheum. Dis. *78*, 270–277. https://doi.org/10.1136/annrheumdis-2018-213882.
87. Ajekigbe, B., Cheung, K., Xu, Y., Skelton, A.J., Panagiotopoulos, A., Soul, J., Hardingham, T.E., Deehan, D.J., Barter, M.J., and Young, D.A. (2019). Identification of long non-coding RNAs expressed in knee and hip osteoarthritic cartilage. Osteoarthritis Cartilage *27*, 694–702. https://doi.org/10.1016/j.joca.2018.12.015.
88. Xiao, K., Yang, Y., Bian, Y., Feng, B., Li, Z., Wu, Z., Qiu, G., and Weng, X. (2019). Identification of differentially expressed long noncoding RNAs in human knee osteoarthritis. J. Cell. Biochem. *120*, 4620–4633. https://doi.org/10.1002/jcb.27750.
89. Soul, J., Dunn, S.L., Anand, S., Serracino-Inglott, F., Schwartz, J.-M., Boot-Handford, R.P., and Hardingham, T.E. (2018). Stratification of knee osteoarthritis: two major patient subgroups identified by genome-wide expression analysis of articular cartilage. Ann. Rheum. Dis. *77*, 423. https://doi.org/10.1136/annrheumdis-2017-212603.
90. Fisch, K.M., Gamini, R., Alvarez-Garcia, O., Akagi, R., Saito, M., Muramatsu, Y., Sasho, T., Koziol, J.A., Su, A.I., and Lotz, M.K. (2018). Identification of transcription factors responsible for dysregulated networks in human osteoarthritis cartilage by global gene expression analysis. Osteoarthritis Cartilage *26*, 1531–1538. https://doi.org/10.1016/j.joca.2018.07.012.
91. Akagi, R., Akatsu, Y., Fisch, K.M., Alvarez-Garcia, O., Teramura, T., Muramatsu, Y., Saito, M., Sasho, T., Su, A.I., and Lotz, M.K. (2017). Dysregulated circadian rhythm pathway in human osteoarthritis: NR1D1 and BMAL1 suppression alters TGF-β signaling in chondrocytes. Osteoarthritis Cartilage *25*, 943–951. https://doi.org/10.1016/j.joca.2016.11.007.
92. Steinberg, J., Ritchie, G.R.S., Roumeliotis, T.I., Jayasuriya, R.L., Clark, M.J., Brooks, R.A., Binch, A.L.A., Shah, K.M., Coyle, R., Pardo, M., et al. (2017). Integrative epigenomics, transcriptomics and proteomics of patient chondrocytes reveal genes and pathways involved in osteoarthritis. Sci. Rep. *7*, 8935. https://doi.org/10.1038/s41598-017-09335-6.
93. Dunn, S.L., Soul, J., Anand, S., Schwartz, J.-M., Boot-Handford, R.P., and Hardingham, T.E. (2016). Gene expression changes in damaged osteoarthritic cartilage identify a signature of non-chondrogenic and mechanical responses. Osteoarthritis Cartilage *24*, 1431–1440. https://doi.org/10.1016/j.joca.2016.03.007.
94. Snelling, S., Rout, R., Davidson, R., Clark, I., Carr, A., Hulley, P.A., and Price, A.J. (2014). A gene expression study of normal and damaged cartilage in anteromedial gonarthrosis, a phenotype of osteoarthritis. Osteoarthritis Cartilage *22*, 334–343. https://doi.org/10.1016/j.joca.2013.12.009.
95. Ramos, Y.F.M., den Hollander, W., Bovée, J.V.M.G., Bomer, N., van der Breggen, R., Lakenberg, N., Keurentjes, J.C., Goeman, J.J., Slagboom, P.E., Nelissen, R.G.H.H., et al. (2014). Genes Involved in the Osteoarthritis Process Identified through Genome Wide Expression Analysis in Articular Cartilage; the RAAK Study. PLoS One *9*, e103056. https://doi.org/10.1371/journal.pone.0103056.
96. Armiento, A.R., Alini, M., and Stoddart, M.J. (2019). Articular fibrocartilage - Why does hyaline cartilage fail to repair? Adv. Drug Deliv. Rev. *146*, 289–305. https://doi.org/10.1016/j.addr.2018.12.015.
97. van Eegher, S., Perez-Lozano, M.-L., Toillon, I., Valour, D., Pigenet, A., Citadelle, D., Bourrier, C., Courtade-Gaïani, S., Grégoire, L., Cléret, D., et al. (2021). The differentiation of prehypertrophic into hypertrophic chondrocytes drives an OA-remodeling program and IL-34 expression. Osteoarthritis Cartilage *29*, 257–268. https://doi.org/10.1016/j.joca.2020.10.013.
98. Sophia Fox, A.J., Bedi, A., and Rodeo, S.A. (2009). The Basic Science of Articular Cartilage. Sports Health *1*, 461–468. https://doi.org/10.1177/1941738109350438.
99. Duboc, V., Sulaiman, F.A., Feneck, E., Kucharska, A., Bell, D., Holder-Espinasse, M., and Logan, M.P.O. (2021). Tbx4 function during hindlimb development reveals a mechanism that explains the origins of proximal limb defects. Development *148*, dev199580. https://doi.org/10.1242/dev.199580.
100. Iwamoto, M., Higuchi, Y., Enomoto-Iwamoto, M., Kurisu, K., Koyama, E., Yeh, H., Rosenbloom, J., and Pacifici, M. (2001). The role of ERG (ets related gene) in cartilage development. Osteoarthritis Cartilage *9*, S41–S47. https://doi.org/10.1053/joca.2001.0443.
101. Hardingham, T. (2008). Extracellular matrix and pathogenic mechanisms in osteoarthritis. Curr. Rheumatol. Rep. *10*, 30–36. https://doi.org/10.1007/s11926-008-0006-9.
102. Baldarelli, R.M., Smith, C.L., Ringwald, M., Richardson, J.E., Bult, C.J.; and Mouse Genome Informatics Group (2024). Mouse Genome Informatics: an integrated knowledgebase system for the laboratory mouse. Genetics *227*, iyae031. https://doi.org/10.1093/genetics/iyae031.
103. Bernabei, I., So, A., Busso, N., and Nasi, S. (2023). Cartilage calcification in osteoarthritis: mechanisms and clinical relevance. Nat. Rev. Rheumatol. *19*, 10–27. https://doi.org/10.1038/s41584-022-00875-4.
104. Huang, S., Mao, J., Wei, B., and Pei, G. (2015). The anti-spasticity drug baclofen alleviates collagen-induced arthritis and regulates dendritic cells. J. Cell. Physiol. *230*, 1438–1447. https://doi.org/10.1002/jcp.24884.
105. Shen, J., Wang, C., Ying, J., Xu, T., McAlinden, A., and O'Keefe, R.J. (2019). Inhibition of 4-aminobutyrate aminotransferase protects against injury-induced osteoarthritis in mice. JCI Insight *4*, e128568. https://doi.org/10.1172/jci.insight.128568.
106. Ajmal, I., Farooq, M.A., Abbas, S.Q., Shah, J., Majid, M., and Jiang, W. (2022). Isoprenaline and salbutamol inhibit pyroptosis and promote mitochondrial biogenesis in arthritic chondrocytes by downregulating β-arrestin and GRK2.

Front. Pharmacol. *13*, 996321. https://doi.org/10.3389/fphar.2022.996321.

107. Zhou, J., Liu, C., Sun, Y., Francis, M., Ryu, M.S., Grider, A., and Ye, K. (2021). Genetically predicted circulating levels of copper and zinc are associated with osteoarthritis but not with rheumatoid arthritis. Osteoarthritis Cartilage *29*, 1029–1035. https://doi.org/10.1016/j.joca.2021.02.564.
108. Richard, D., Liu, Z., Cao, J., Kiapour, A.M., Willen, J., Yarlagadda, S., Jagoda, E., Kolachalama, V.B., Sieker, J.T., Chang, G.H., et al. (2020). Evolutionary Selection and Constraint on Human Knee Chondrocyte Regulation Impacts Osteoarthritis Risk. Cell *181*, 362–381.e28. https://doi.org/10.1016/j.cell.2020.02.057.
109. Rice, S.J., Brumwell, A., Falk, J., Kehayova, Y.S., Casement, J., Parker, E., Hofer, I.M.J., Shepherd, C., and Loughlin, J. (2023). Genetic risk of osteoarthritis operates during human skeletogenesis. Hum. Mol. Genet. *32*, 2124–2138. https://doi.org/10.1093/hmg/ddac251.
110. Walker, R.L., Ramaswami, G., Hartl, C., Mancuso, N., Gandal, M.J., de la Torre-Ubieta, L., Pasaniuc, B., Stein, J.L., and Geschwind, D.H. (2019). Genetic Control of Expression and Splicing in Developing Human Brain Informs Disease Mechanisms. Cell *179*, 750–771.e22. https://doi.org/10.1016/j.cell.2019.09.021.
111. Mertens, J., Herdy, J.R., Traxler, L., Schafer, S.T., Schlachetzki, J.C.M., Böhnke, L., Reid, D.A., Lee, H., Zangwill, D., Fernandes, D.P., et al. (2021). Age-dependent instability of mature neuronal fate in induced neurons from Alzheimer's patients. Cell Stem Cell *28*, 1533–1548.e6. https://doi.org/10.1016/j.stem.2021.04.004.
112. Marangoni, R.G., Korman, B.D., Wei, J., Wood, T.A., Graham, L.V., Whitfield, M.L., Scherer, P.E., Tourtellotte, W.G., and Varga, J. (2015). Myofibroblasts in murine cutaneous fibrosis originate from adiponectin-positive intradermal progenitors. Arthritis Rheumatol. *67*, 1062–1073. https://doi.org/10.1002/art.38990.
113. Kawaguchi, H. (2008). Endochondral ossification signals in cartilage degradation during osteoarthritis progression in experimental mouse models. Mol. Cells *25*, 1–6.
114. Kamekura, S., Kawasaki, Y., Hoshi, K., Shimoaka, T., Chikuda, H., Maruyama, Z., Komori, T., Sato, S., Takeda, S., Karsenty, G., et al. (2006). Contribution of runt-related transcription factor 2 to the pathogenesis of osteoarthritis in mice after induction of knee joint instability. Arthritis Rheum. *54*, 2462–2470. https://doi.org/10.1002/art.22041.
115. Kamekura, S., Hoshi, K., Shimoaka, T., Chung, U., Chikuda, H., Yamada, T., Uchida, M., Ogata, N., Seichi, A., Nakamura, K., and Kawaguchi, H. (2005). Osteoarthritis development in novel experimental mouse models induced by knee joint instability. Osteoarthritis Cartilage *13*, 632–641. https://doi.org/10.1016/j.joca.2005.03.004.
116. Samvelyan, H.J., Madi, K., Törnqvist, A.E., Javaheri, B., and Staines, K.A. (2021). Characterisation of Growth Plate Dynamics in Murine Models of Osteoarthritis. Front. Endocrinol. *12*, 734988. https://doi.org/10.3389/fendo.2021.734988.
117. Dreier, R. (2010). Hypertrophic differentiation of chondrocytes in osteoarthritis: the developmental aspect of degenerative joint disorders. Arthritis Res. Ther. *12*, 216. https://doi.org/10.1186/ar3117.
118. Aghajanian, P., and Mohan, S. (2018). The art of building bone: emerging role of chondrocyte-to-osteoblast transdifferentiation in endochondral ossification. Bone Res. *6*, 19. https://doi.org/10.1038/s41413-018-0021-z.
119. Matta, C., Takács, R., Ducza, L., Ebeid, R.A., Choi, H., and Mobasheri, A. (2023). Ion channels involved in inflammation and pain in osteoarthritis and related musculoskeletal disorders. Am. J. Physiol. Cell Physiol. *325*, C257–C271. https://doi.org/10.1152/ajpcell.00040.2023.
120. Sohn, R., Rösch, G., Junker, M., Meurer, A., Zaucke, F., and Jenei-Lanzl, Z. (2021). Adrenergic signalling in osteoarthritis. Cell. Signal. *82*, 109948. https://doi.org/10.1016/j.cellsig.2021.109948.
121. Arruda, A.L., Morris, A.P., and Zeggini, E. (2024). Advancing equity in human genomics through tissue-specific multiancestry molecular data. Cell Genom. *4*, 100485. https://doi.org/10.1016/j.xgen.2023.100485.