current issues in molecular biology

MDPI

Article

Beyond the Case–Control Signal: MMP13 Is Embedded in a Reproducible Hypertrophic and Fibro-Stromal Co-Expression Programme Within Osteoarthritic Joint Tissues

Elif Defne Aydoğdu [1], Sajjad Eslamkhah [2,3], Hakan Gülmez [2], Elif Sibel Aslan [1] and Hilal Şahin [4,*]

[1] Department of Molecular Biology and Genetics, Biruni University, Istanbul 34015, Türkiye; 250401035@st.biruni.edu.tr (E.D.A.); easlan@biruni.edu.tr (E.S.A.)
[2] Biruni Advanced Technology Application and Research Center (BAMER), Biruni University, Istanbul 34015, Türkiye; mtaghizad@biruni.edu.tr (S.E.); hakan.glmz56@gmail.com (H.G.)
[3] Program of Pathology Laboratory Techniques, Vocational School, Biruni University, Istanbul 34015, Türkiye
[4] Department of Medical Biology and Genetics, Faculty of Medicine, Atlas University, Istanbul 34403, Türkiye
* Correspondence: hilal.sahin@atlas.edu.tr

Abstract

Matrix metalloproteinase-13 (MMP13) is the principal collagenase of osteoarthritic (OA) cartilage, but it is unclear whether its expression distinguishes OA from control tissue and which genes vary with it within diseased tissue. We re-analysed seven public microarray datasets from four joint tissues (123 OA, 66 control samples). MMP13 did not consistently separate OA from control tissue (pooled Hedges' g = 0.69, 95% confidence interval (CI): −0.39 to 1.77), no age-annotated discovery dataset retained a difference after age adjustment, and MMP13 was not higher in lesioned than in preserved cartilage. Within OA tissue, 1510 genes were positively and 407 negatively co-expressed with MMP13, which tracked a hypertrophy score ($\rho = 0.65$). This neighbourhood was no larger than those of expression-matched genes, and FAP and THY1 were more central, but MMP13's partners replicated better than those of 83–99% of random genes. Composition adjustment retained 124 of the top 150 positive but only 26 of the top 150 negative partners. In held-out cartilage and synovium, 80% and 93% of positive and 71% and 64% of negative partners were concordant, but after platform adjustment, MMP13 no longer tracked the negative module, and the hypertrophy correlation replicated only in synovium. In single-cell data, MMP13 showed its highest mean expression in a COL10A1^{+}IBSP^{+} cluster, with a confidence interval overlapping that of another cluster. In the discovery data, case–control differences in MMP13 could not be separated from donor age; within OA tissue, MMP13 is a consistent member of a reproducible fibro-stromal programme.

Keywords: osteoarthritis; MMP13; chondrocyte hypertrophy; co-expression meta-analysis; transcriptomics; single-cell RNA sequencing; synovium; donor age

Check for updates

Academic Editor: Li Liao

Received: 15 October 2026
Revised: 16 November 2026
Accepted: 20 November 2026
Published: 26 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Osteoarthritis (OA) affected an estimated 595 million people worldwide in 2020, and its prevalence is projected to keep rising with population ageing [1]. It is a leading cause of pain and disability in older adults, yet no disease-modifying treatment is available [2]. OA is now understood as a disease of the whole joint, in which articular cartilage, subchondral bone, synovium and menisci undergo coordinated structural and molecular change [3].

Matrix metalloproteinase-13 (MMP13, collagenase-3) has long been regarded as the main mediator of cartilage destruction in OA. Chondrocytes in OA cartilage produce it, and it

cleaves type II collagen at least tenfold faster than MMP1 [4]. Mice expressing constitutively active MMP13 in cartilage develop OA-like pathology [5], and MMP13-deficient mice are protected from cartilage erosion after surgical joint destabilisation [6]. The same knockout mice still form hypertrophic chondrocytes and osteophytes, so MMP13 is not required for this differentiation programme [6].

These findings made MMP13 an attractive drug target, but translation has been difficult. Broad-spectrum MMP inhibitors caused musculoskeletal toxicity: in a 12-month trial in knee OA, PG-116800 produced arthralgia and other musculoskeletal adverse events without slowing joint-space narrowing [7]. Selective MMP13 inhibitors that do not bind the catalytic zinc reduced cartilage damage in rat models without this toxicity [8,9], and intra-articular nanoparticles carrying Mmp13 small interfering RNA protected joint structure in mouse post-traumatic OA [10]; these approaches are reviewed in [11]. MMP13 also has physiological functions. It is expressed in the early proliferative zone of the growth plate [12] and takes part in cartilage and bone remodelling during skeletal repair [13], so blocking it may carry costs. A second obstacle is patient selection. OA is heterogeneous, and an MMP13-directed treatment can only help patients in whom MMP13-linked biology is active. Choosing those patients requires knowing whether MMP13 marks the disease itself or a programme that varies from one patient to the next.

That programme is closely linked to hypertrophy-like changes of articular chondrocytes, which resemble terminal differentiation in the growth plate and are accompanied by protease-mediated matrix degradation [14]. Transcriptional control of MMP13 integrates several signalling pathways and transcription factors [15], and hypoxia-inducible factor-2α directly induces MMP13 together with other catabolic genes in chondrocytes [16]. Single-cell transcriptomics of human OA cartilage has further resolved prehypertrophic, hypertrophic and fibrocartilage chondrocyte populations [17], suggesting that MMP13 expression may vary between patients and cell states rather than switching on uniformly with disease.

Public transcriptomes now cover several OA joint tissues, including paired lesioned and preserved cartilage [18], meniscus from knees with and without OA [19], cartilage from knees with meniscal tears and from knees with end-stage OA [20], synovium [21] and subchondral bone [22]. Most analyses of these data have compared OA with control tissue. Such comparisons are sensitive to differences in donor age, a major determinant of joint tissue biology [23], and they cannot show which programme MMP13 follows among patients who already have OA.

These data have been re-analysed many times. Bioinformatic studies have repeatedly ranked MMP13 or genes that we find in its co-expression neighbourhood such as POSTN and FAP, among "hub genes" derived from case–control contrasts in small public series. We cite three recent examples, chosen for illustration and not as a systematic sample, because all three relied on GSE169077, a cartilage series with six OA and five control arrays, either combined with GSE114007 [24] or GSE55235 [25], or alone [26]. Such rankings rarely report whether patients and controls are comparable in age. In parallel, unsupervised analyses of cartilage transcriptomes have identified molecularly distinct patient subgroups [27–29], so the variation among patients that case–control designs average out carries biological signal. The link between MMP13 and hypertrophy rests mainly on experiments: HIF-2α transactivates COL10A1, MMP13 and VEGFA [30], and Indian hedgehog expression in human OA cartilage rises with COL10A1, MMP13 and disease severity [31]. Whether these relationships hold across patients, tissues and cohorts once donor age is taken into account has not been tested.

Here, we re-analysed seven public datasets from four joint tissues with four aims: (i) to estimate the OA-versus-control difference in MMP13 across tissues and how far donor age accounts for it; (ii) to test, within patients, whether MMP13 is higher in lesioned than in preserved cartilage; (iii) to map the consensus co-expression neighbourhood of

MMP13 within OA tissue and the programmes and candidate regulators it reflects, to test its sensitivity to donor age, sex and cell-type composition, and to ask whether MMP13 is more central to it than comparable genes; and (iv) to test the pre-specified neighbourhood in independent bulk and single-cell datasets that were not used for its discovery.

2. Materials and Methods

2.1. Dataset Selection and Preprocessing

Processed microarray series were retrieved from the Gene Expression Omnibus (GEO) [32] through the GEOparse Python package (version 2.0.4; Tables 1 and S1). Probes were matched to gene symbols using the platform annotation and reduced to a single probe per gene, keeping the probe with the highest mean expression. Values were log2-transformed when their 99th percentile exceeded 50. Genes absent from more than 20% of samples were dropped, and any remaining gaps were filled with the gene mean. All analyses used publicly available, de-identified data; no new samples were collected.

The seven discovery datasets were selected from a list, compiled by the authors before analysis, of public GEO series that compared OA with control tissue of the same type in human cartilage, subchondral bone, synovium or meniscus. No systematic database search was performed, so a PRISMA 2020 flow diagram [33] could not be constructed; instead, every series that was assessed is listed in Table S5 with the reason for its inclusion or exclusion. Series were used if they (i) profiled human joint tissue from patients with primary OA; (ii) included control tissue of the same type from donors without end-stage OA; (iii) had at least five samples per group; (iv) used a genome-wide expression microarray with processed data deposited in GEO; and (v) annotated OA and control status at the sample level. Series were excluded if they profiled cultured or stimulated cells, repeated arrays of another included series, or lacked usable sample annotation; RNA-sequencing series were reserved for validation, and the microarray series GSE82107 was assigned to validation in advance to provide an independent synovium cohort. In total, 12 series were assessed: 8 microarray series for discovery (seven used), 3 pre-specified validation series (two used) and 1 single-cell series.

Table 1. Datasets analysed.

Dataset	Tissue	Platform	OA/Control (n)	Control Definition (GEO Annotation)	Median Age, Control vs. OA (Years)	Original Study
GSE57218	Cartilage	GPL6947	33/7	Healthy cartilage	18 vs. 67	Ramos et al. [18]
GSE117999	Cartilage	GPL20844	12/12	Knees undergoing arthroscopic partial meniscectomy	52 vs. 64	GEO series
GSE169077	Cartilage	GPL96	6/5	Normal	Not annotated	GEO series
GSE98918	Meniscus	GPL20844	12/12	Knees undergoing arthroscopic partial meniscectomy	52 vs. 64	Brophy et al. [19]
GSE55235	Synovium	GPL96	10/10	Healthy control	Not annotated	Woetzel et al. [21]
GSE55457	Synovium	GPL96	10/10	Normal control	57 vs. 73	Woetzel et al. [21]
GSE51588	Subchondral bone	GPL13497	40/10	Normal	35 vs. 68	Chou et al. [22]

GSE117999 and GSE98918 share patient identifiers, ages and body-mass indices, so they were merged into one cohort. Rheumatoid arthritis samples inside GSE55235 and GSE55457 were excluded. Joint of origin: GSE57218 OA cartilage came from 22 hip and 11 knee joints [18], and the original report does not state the joint of the seven healthy donors; GSE117999 and GSE98918, knee; GSE169077, knee; GSE51588, knee (tibial plateaus) [22]; GSE55235 and GSE55457, synovium obtained at joint replacement or synovectomy, without the joint being named in the original report [21]. Control synovium in GSE55235 and GSE55457 came from donors without known joint disease, obtained post mortem ($n = 15$) or after traumatic joint injury ($n = 5$) across the two series (the split per series is not reported [21]), and no histological grading of these controls is available in GEO. Control subchondral bone in GSE51588 came from donors with nonarthritic knees, obtained at autopsy or within 8 h of amputation [22]. The GEO record of GSE169077 does not state whether its arrays represent individual donors or pooled RNA; pooling is plausible, because a U133A study of normal and OA cartilage-hybridised RNA pooled from 5 to 10 donors per array [34].

2.2. Sample Groups and Duplicate Screening

Groups were assigned from GEO sample annotations. In GSE57218, the OA group comprised lesioned cartilage (one sample per patient); macroscopically preserved cartilage from the same 33 joints [18] was used only in the within-patient analysis. GSE12021 was left out because its OA arrays repeat those of GSE55457, with within-OA gene–MMP13 correlations identical across 95% of genes. Series on the same platform were also screened for duplicated arrays using residual sample–sample correlations (threshold $r > 0.98$); no further duplicates were found.

2.3. OA Versus Control Meta-Analysis

For each dataset, the difference in MMP13 between OA and control samples was expressed as Hedges' g and tested with Welch's *t*-test. Estimates were pooled by random-effects meta-analysis, using the DerSimonian–Laird estimator for between-study variance [35] and Hartung–Knapp–Sidik–Jonkman confidence intervals, which perform better when few studies are pooled [36]. Datasets from the same patients were averaged before pooling, under the conservative assumption of a correlation of 1 between them. Where donor age was annotated, MMP13 was modelled by ordinary least squares as MMP13~group and MMP13~group + age; the standardised effect is the group coefficient divided by the residual standard deviation. Sex was taken from GEO annotations where available and otherwise inferred from expression of XIST and Y-chromosome genes (RPS4Y1, DDX3Y, KDM5D and EIF1AY). Because the prevalence and severity of OA differ between women and men [37], the models were repeated with sex as an additional covariate. In sensitivity analyses, GSE169077 was left out of the pooled estimate, and the four GSE57218 control donors younger than 20 years were excluded.

2.4. Within-Patient Analysis

In GSE57218, lesioned and preserved samples were paired by patient identifier and compared with the Wilcoxon signed-rank test. Module scores were defined as the mean within-dataset z-score of the 50 strongest positively and negatively co-expressed MMP13 partners, obtained from a meta-analysis that left this cohort out.

2.5. Co-Expression Meta-Analysis

Within the OA samples of each dataset (all samples where fewer than eight OA samples were available), Spearman correlations between MMP13 and every gene were computed and Fisher-transformed. Transformed correlations were averaged inside cohorts and then combined across cohorts with the sample-size-weighted Z-method [38], with weights of $n - 3$. Only genes measured in at least half of the cohorts were retained; *p* values were corrected by the Benjamini–Hochberg procedure. Robust partners required a false discovery rate (FDR) < 0.05 together with the same sign in at least 75% of cohorts. In an age sensitivity analysis, partial Spearman correlations controlling for age (rank residuals after linear regression on age) replaced raw correlations in age-annotated datasets. The Fisher-transformed coefficients were given the variance $1/(n - 3)$, which applies exactly to Pearson coefficients and gives the $n - 3$ weights; the combined statistic was $Z = \Sigma(n_c - 3)z_c/\sqrt{\Sigma}(n_c - 3)$, where z_c is the Fisher-transformed correlation of cohort c, which equals the inverse-variance weighted mean divided by its standard error. For Spearman coefficients, the variance is closer to $1.06/(n - 3)$ [39]. Because this factor is the same in every cohort, it leaves the weights and the ranking of genes unchanged and shrinks every meta-Z by a factor of $\sqrt{1.06}$ (about 3%); we nevertheless repeated the analysis with the corrected variance (Table S11). In the merged cohort (GSE117999 and GSE98918, same patients), the two transformed coefficients were averaged and weighted by the number of patients ($n = 12$), so that no

patient contributed twice. As a further sensitivity analysis, the meta-analysis was repeated without GSE169077, the only cohort in which OA and control samples were combined for co-expression.

To test whether MMP13 is more central to this neighbourhood than other genes, the co-expression meta-analysis was repeated with other genes as seed in place of MMP13: OLFML2B and THY1, its two strongest partners; POSTN and FAP, two fibro-stromal partners; MMP3, a catabolic gene not co-expressed with MMP13; and 1000 random genes matched to MMP13 on mean expression and variance within each cohort. For each seed, we recorded the number of robust partners and, in the validation cohorts, the fraction of its top 150 positive partners that were positively correlated with it. We also built a genome-wide co-expression network from the 5000 most variable genes, which included MMP13, using the cross-cohort meta-correlations, and computed each gene's weighted degree (soft threshold $\beta = 6$) and its eigenvector centrality.

Because co-expression in bulk tissue can arise from differences in cell-type proportions between samples, each sample was scored for 11 cell types with marker genes (Table S7), following the marker-based approach of MCP-counter [40]: chondrocytes, fibroblasts, synovial lining fibroblasts, osteoblasts, osteoclasts, macrophages, T cells, B and plasma cells, endothelial cells, pericytes and smooth muscle cells, and adipocytes. Scores were means of within-dataset z-scores. Marker genes that were themselves locked MMP13 partners were removed from the scores, so that the adjustment did not remove the signal by construction. MMP13 partial correlations were then recomputed within each cohort, controlling for the first two principal components of the composition scores (one component in a further sensitivity analysis); the partial correlations were given the same variance, $1/(n - 3)$, without reduction for the covariates.

The co-expression meta-analysis was also repeated within the control samples of the discovery datasets, and the locked partners were also examined in the control samples of the two validation datasets. For each gene, the OA–control difference in Fisher-transformed correlation was tested with a Z statistic in each cohort and combined across cohorts with the same weighted Z-method.

2.6. Programme Scores

Catabolic (MMP13, MMP3, MMP1, ADAMTS4, ADAMTS5, MMP9), anabolic (COL2A1, ACAN, SOX9, PRG4, COMP, TIMP3) and hypertrophy (RUNX2, COL10A1, IHH, ALPL, VEGFA, SPP1, PTH1R) scores were mean within-dataset z-scores. Two balance indices were specified: the Catabolic Switch Index, $z(\text{MMP13}) - z(\text{TIMP3})$, and the Matrix Turnover Balance, catabolic minus anabolic score. Separation of OA from control samples was summarised by the area under the receiver operating characteristic curve (AUC) with 1000 bootstrap resamples. Several of the hypertrophy genes (SPP1, ALPL, RUNX2) also mark osteoblast-lineage cells, so outside cartilage we call this a hypertrophy/osteogenic score, and MMP13–score correlations are reported for each tissue separately.

2.7. Enrichment Analyses

Pre-ranked gene-set enrichment analysis (GSEA) [41] was carried out in GSEApy (version 1.3.1) [42] against the Molecular Signatures Database (MSigDB) Hallmark collection (2020 release, 1000 permutations, gene-set size: 10–500) [43], with genes ranked by consensus co-expression meta-Z. As a sensitivity analysis, top and bottom MMP13 tertiles within OA were contrasted (Welch's t ranking) in datasets with at least 20 OA samples. Over-representation of the top 150 positive and top 150 negative partners was tested with Enrichr (https://maayanlab.cloud/Enrichr/, accessed through GSEApy 1.3.1 on 12 October 2026) [44] against TRRUST [45], ChEA [46], Reactome [47] and Gene Ontology (GO)

Biological Process libraries (Enrichr releases TRRUST 2019, ChEA 2022, Reactome 2022 and GO Biological Process 2023), with all tested genes serving as the background. Pre-ranked GSEA permutes genes and therefore ignores correlation between genes. We therefore re-tested the Hallmark sets with two correlation-aware procedures: the pre-ranked form of CAMERA (correlation-adjusted mean-rank gene set test; implemented in Python, see Section 2.10) [48], with the inter-gene correlation of each set estimated from OA samples, and a sample-label permutation test in which MMP13 values were permuted within each cohort (1000 permutations; in the merged cohort, the same permutation was applied to both series, so that patients stayed aligned), the meta-analysis was rerun, and the mean meta-Z of each set was compared with its permutation distribution.

2.8. Drug–Gene Interactions

Approved-drug interaction records for the strongest partners were retrieved from the Drug–Gene Interaction Database (DGIdb) [49] on 12 October 2026. They are reported as database records only (Figure S4).

2.9. Independent Validation (Pre-Specified)

Before any validation data were examined, the discovery signature was frozen: the top 150 positive and top 150 negative robust partners together with the genome-wide meta-correlations, saved with a time stamp. Validation used datasets that were not used in discovery: knee cartilage RNA-sequencing from 18 normal and 20 OA donors (GSE114007) [50] and synovium microarrays from 7 healthy and 10 OA donors (GSE82107; the joint is not stated in its GEO record). A third pre-specified series (GSE89408, synovium RNA-sequencing) could not be used because no supplementary expression table could be matched unambiguously to its GEO sample annotations. Expression tables came from GEO series matrices or Supplementary Files; Ensembl identifiers were mapped to gene symbols with MyGene.info (https://mygene.info/, accessed on 16 November 2026; queried with the mygene Python client, version 3.2.2) [51], and raw counts were converted to log2 counts per million. Groups followed GEO annotations once severity fields (for example, 'OA grade') had been removed, and a dataset was used only when the resulting group sizes matched the published design. Within OA samples, three pre-specified tests were applied: (i) genome-wide Spearman concordance between discovery meta-correlations and validation correlations with MMP13; (ii) one-sided Mann–Whitney tests of whether locked positive and negative partners had shifted relative to all other genes; and (iii) Spearman correlation of MMP13 with the 50-gene positive and negative module scores and with the hypertrophy score. Because GSE114007 was sequenced on two platforms, tests (i)–(iii) were repeated after centring each gene within the platform. MMP13 in OA versus control samples was compared by Welch's *t*-test and, where age was annotated, by linear regression on group and age, with the platform added as a covariate when samples spanned more than one platform.

For single-cell localisation, OA knee cartilage single-cell RNA-sequencing data (GSE104782) [17] were processed in Scanpy (version 1.12.4) [52]: cells with fewer than 200 detected genes were discarded (1597 cells retained; the original study analysed 1464 cells after its own quality control), counts were normalised and log-transformed according to library size, and 2000 highly variable genes were selected with the 'seurat' method of Scanpy. Data were scaled and reduced to 50 principal components, a neighbourhood graph was built from 15 nearest neighbours on all 50 components, and cells were clustered with the Leiden algorithm [53] at a resolution of 0.6 (random seed 7). In the primary analysis, no filter on mitochondrial read fraction, no doublet removal and no batch correction across donors were applied. In a sensitivity analysis, cells with a mitochondrial fraction above 20%

and doublets predicted by Scrublet (as implemented in Scanpy 1.12.4) [54] were removed, Harmony (harmonypy Python implementation) [55] was run on the sample label of the deposited matrix, cells were re-clustered, and the two clusterings were compared by the adjusted Rand index. The deposited matrix is a single cell-by-gene table with one sample label, so Harmony could not correct donor effects. The mean MMP13 expression and the fraction of cells with detectable MMP13 were computed per cluster, with 95% bootstrap confidence intervals over cells, which ignore donor structure. Locked module and hypertrophy scores were computed per cell, and clusters were described by their score profiles and their top marker genes from Wilcoxon tests, with no cell-type labels assigned in advance.

2.10. Software and Reproducibility

Analyses were run in Python on Google Colab (Google Colaboratory; Google LLC, Mountain View, CA, USA; https://colab.research.google.com/, accessed on 16 November 2026, a hosted service without version numbers) with pandas, NumPy, SciPy, statsmodels, scikit-learn, GSEApy, GEOparse, Scanpy, anndata, mygene, NetworkX and matplotlib (Python 3.13.16; pandas 3.0.6, NumPy 2.1.3, SciPy 1.16.3, statsmodels 0.15.0, scikit-learn 1.6.1, GSEApy 1.3.1, GEOparse 2.0.4, Scanpy 1.12.4, anndata 0.13.4, mygene 3.2.2, NetworkX 3.7 and matplotlib 3.10.0). Random seeds were fixed, and the validation signature was saved with a time stamp for reproducibility. CAMERA was implemented in Python following the pre-ranked formula of limma::cameraPR, and the single-cell sensitivity analysis used the Scrublet implementation in Scanpy and harmonypy.

Generative artificial intelligence (Claude; Anthropic, San Francisco, CA, USA; Claude Opus 5.5 during the revision) was used to assist with writing and debugging the analysis code run in Google Colab and with preparing and language-editing the manuscript draft. During the revision, Claude (Opus 5.5) was also used to write the code for the additional analyses requested by the reviewers, which the authors ran in Google Colab; to prepare the Supplementary Figures and Tables; to cross-check reported values against the analysis outputs and the references against PubMed; and to draft the language edits and the responses to the reviewers. The authors subsequently revised the text and independently verified all analyses, reported values and references.

3. Results

3.1. MMP13 Does Not Consistently Distinguish OA from Control Tissue Across the Joint

The seven datasets contained 123 OA and 66 control samples from cartilage, meniscus, synovium and subchondral bone and formed six independent cohorts (Table 1). Two of the seven datasets showed higher MMP13 in OA with a 95% confidence interval (CI) of g that excluded zero: GSE169077 cartilage (Hedges' g = 3.41) and GSE55235 synovium (g = 1.30). In the remaining five datasets, the 95% confidence intervals spanned zero (Tables 2 and S2; Figure 1). Across the six cohorts, the pooled estimate was not significant (g = 0.69, 95% CI: −0.39 to 1.77; $p = 0.16$), and heterogeneity was high ($I^2 = 67\%$, $\tau^2 = 0.43$). Without GSE169077, the pooled estimate was g = 0.41 (95% CI: −0.21 to 1.03; $p = 0.14$; k = 5), and heterogeneity fell to $I^2 = 28\%$ (Table S11).

Controls were younger than OA donors in every one of the five datasets carrying age annotations (Table 1); in GSE57218, four of the seven healthy donors were aged 15–18 years. After adjustment for age, none of these datasets retained a significant OA effect on MMP13 ($p = 0.41$–0.95; Figure S2). Group and age were strongly correlated (r = 0.57–0.80), so the adjusted estimates are imprecise and do not prove the absence of an effect. Excluding the four GSE57218 controls aged 15–18 years left three adult controls and gave g = 0.64 (95% CI: −0.55 to 1.83) for GSE57218 and a pooled estimate of g = 0.70 (95% CI: −0.42 to 1.81; $p = 0.17$; Table S11). Adolescent controls are poorly matched to patients with end-

stage OA, so the GSE57218 case–control contrast should be interpreted with caution. Sex was annotated in five of the seven datasets and inferred from expression in the other two. Adding sex to the models changed little except in GSE55235, where the OA effect fell from 1.69 log2 units ($p = 0.007$) to 1.00 ($p = 0.10$; Table S9). A sex covariate is not interpretable for GSE169077 if its arrays are pools of several donors.

Table 2. MMP13 in OA versus control samples.

Dataset	Tissue	OA/Control	log2 Fold Change	Hedges' g (95% CI)	Welch *p*
GSE169077	Cartilage	6/5	3.52	3.41 (1.55 to 5.26)	1.6×10^{-4}
GSE55235	Synovium	10/10	1.69	1.30 (0.33 to 2.26)	0.011
GSE57218	Cartilage	33/7	0.35	0.69 (−0.14 to 1.52)	0.001
GSE98918	Meniscus	12/12	0.40	0.24 (−0.56 to 1.04)	0.55
GSE117999	Cartilage	12/12	0.21	0.16 (−0.65 to 0.96)	0.70
GSE55457	Synovium	10/10	0.20	0.11 (−0.77 to 0.99)	0.80
GSE51588	Subchondral bone	40/10	0.05	0.03 (−0.67 to 0.72)	0.93
Pooled (6 cohorts)	All	—	—	0.69 (−0.39 to 1.77)	0.16

In GSE57218, the Welch test reaches significance even though the confidence interval of g includes zero, because the control variance is very small (7 donors). Pooled estimate: DerSimonian–Laird random effects with Hartung–Knapp–Sidik–Jonkman CI; GSE117999 and GSE98918 averaged into one cohort.

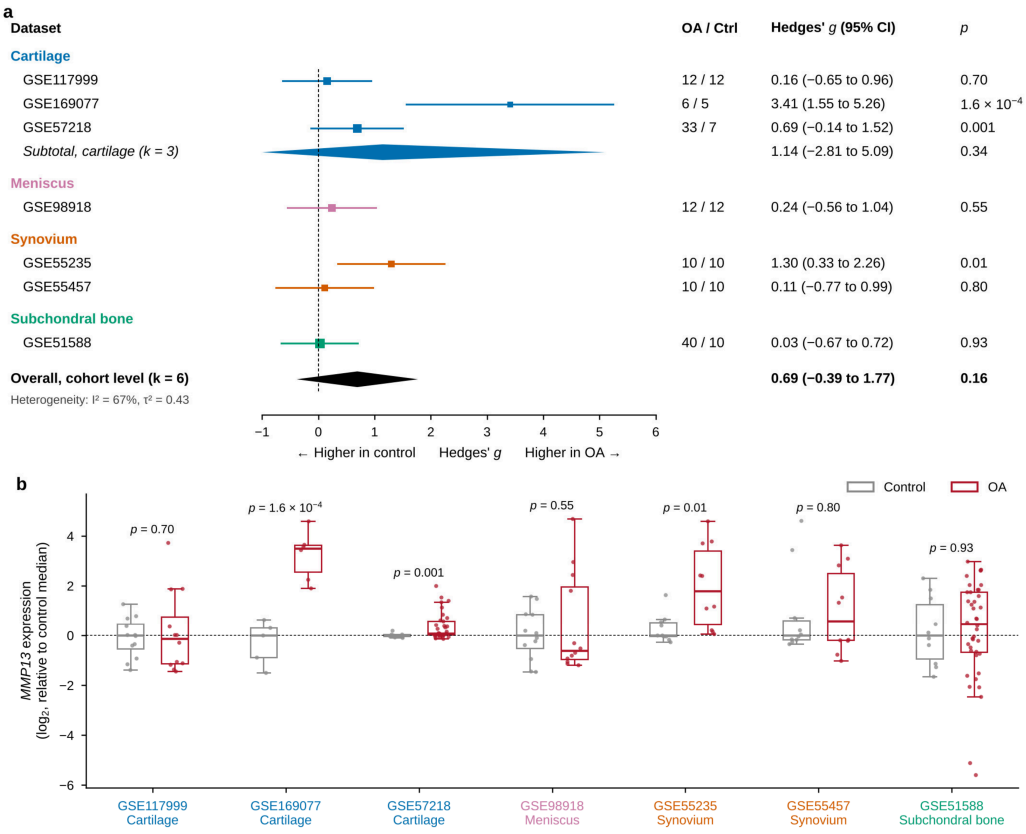

Figure 1. MMP13 expression in OA and control joint tissues. (**a**) Forest plot of Hedges' g for MMP13 (OA vs. control) in seven datasets; squares, dataset estimates (size proportional to n) with 95% CIs; diamonds, random-effects estimates (DerSimonian–Laird τ^2, Hartung–Knapp–Sidik–Jonkman 95% CI); tissue subtotals shown for ≥3 datasets; overall estimate at the cohort level (k = 6). p values, Welch's t-test. (**b**) MMP13 log2 expression relative to the control median; boxes, median and interquartile range (IQR); whiskers, 1.5 × IQR; points, samples; dashed lines, no difference (g = 0 in (**a**); control median in (**b**)).

3.2. MMP13 Is Not Higher in Lesioned than in Preserved Cartilage of the Same Joint

Among 33 OA patients with paired samples, MMP13 was higher in lesioned than in macroscopically preserved cartilage in only 13 (median: 7.15 vs. 7.38 log2 units; median paired difference: −0.03; Wilcoxon $p = 0.063$; Figure 2a). A module built from the 50 strongest positive MMP13 partners, derived without this cohort, showed no difference between sites (21/33 pairs higher in the lesion; $p = 0.56$; Figure 2b). Its negative counterpart was lower in lesioned cartilage (11/33 pairs higher; median paired difference: −0.11; $p = 0.0078$; Figure 2c). Several samples were close to the platform floor (about 7.0 log2 units), which limits resolution where MMP13 expression is low.

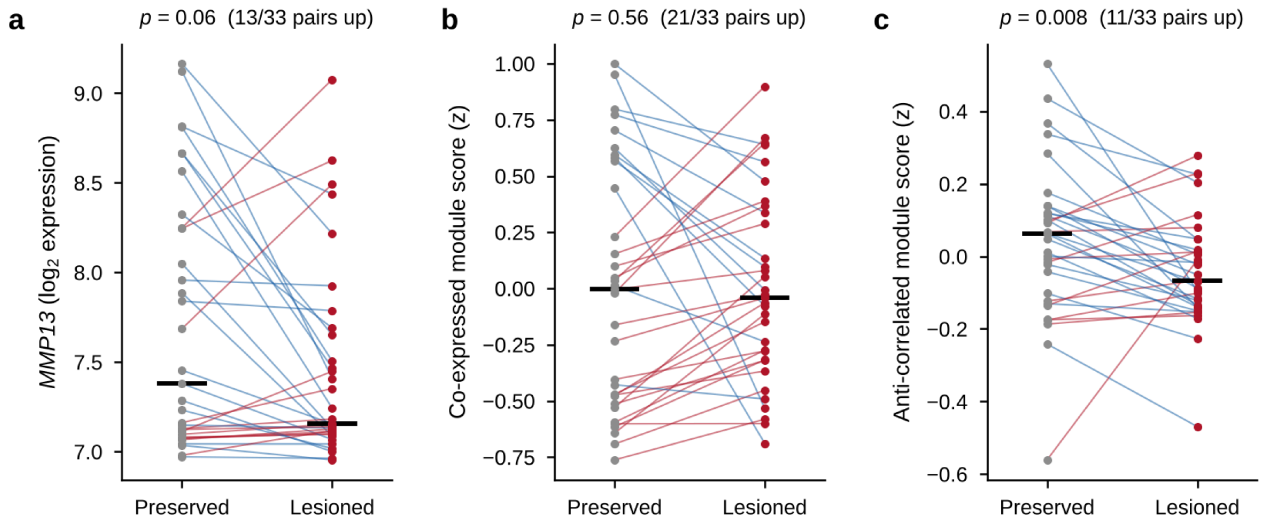

Figure 2. Within-patient comparison of lesioned and macroscopically preserved cartilage (GSE57218, 33 pairs). (**a**) MMP13 expression. (**b**,**c**) Mean z-score of the 50 strongest positively (**b**) and negatively (**c**) co-expressed MMP13 partners, derived without this cohort. Lines join samples from the same patient (red, higher in lesion; blue, lower); black bars, medians. *p* values, Wilcoxon signed-rank test.

3.3. MMP13 Has a Reproducible Co-Expression Neighbourhood in OA Tissue That Changes Little with Age Adjustment

Across six cohorts, 18,259 genes were tested for co-expression with MMP13 within OA samples. Of these, 1510 were robust positive and 407 robust negative partners (FDR < 0.05; same sign in ≥75% of cohorts; Figure 3a; Table S3). Ranked by meta-Z, the strongest positive partners were OLFML2B (meta ρ = 0.77), THY1 (0.75), THBS2 (0.72), GJA1 (0.71), COL5A2 (0.71), AEBP1 (0.69) and MMP9 (0.69); other top-ranked partners included CTSK (0.68), SPP1 (0.67), POSTN (0.67) and FAP (0.66), all with the same sign in every cohort in which they were measured (Figure 3b). The strongest negative partners were GRK5 (−0.58), TIMP4 (−0.56) and CIDEA (−0.56).

Controlling for donor age changed the neighbourhood little. Raw and age-controlled partial meta-correlations agreed genome-wide (Spearman r = 0.99), and the top 150 positive and top 150 negative partners all stayed significant with their original sign (Figure S3). Most OA samples came from older patients undergoing surgery for end-stage disease (median age: 64–73 years across datasets; overall range: 52–89 years). Age adjustment therefore had little effect within that range, and the analysis cannot show whether the programme is independent of age across adult life.

We then compared MMP13 with other seed genes (Table S6; Figure S5). Rerun with the same thresholds, MMP13 had 1920 robust partners (1513 positive and 407 negative; the primary analysis gave 1917, and the source of this three-gene difference between runs was not traced). Several of its own partners had larger neighbourhoods: 2709 robust partners for

FAP, 2493 for THY1, 2464 for OLFML2B and 2268 for POSTN, against 6 for MMP3. Among 1000 random genes matched to MMP13 at the expression level and variance the median was 1628, which placed MMP13 at the 63rd percentile, so the size of its neighbourhood is typical of genes with similar expression. MMP13's partners nevertheless replicated better than those of most random genes: the fraction of its top 150 positive partners that were positively correlated with MMP13 in the validation cohorts ranked at the 83rd percentile in cartilage and the 99th in synovium. In the network of the 5000 most variable genes, MMP13 ranked 366th by weighted degree and 80th by eigenvector centrality, whereas FAP, OLFML2B and THY1 were among the top ten on both measures. MMP13 is therefore a reproducible member of a fibro-stromal co-expression programme, to which FAP, OLFML2B and THY1 are more central.

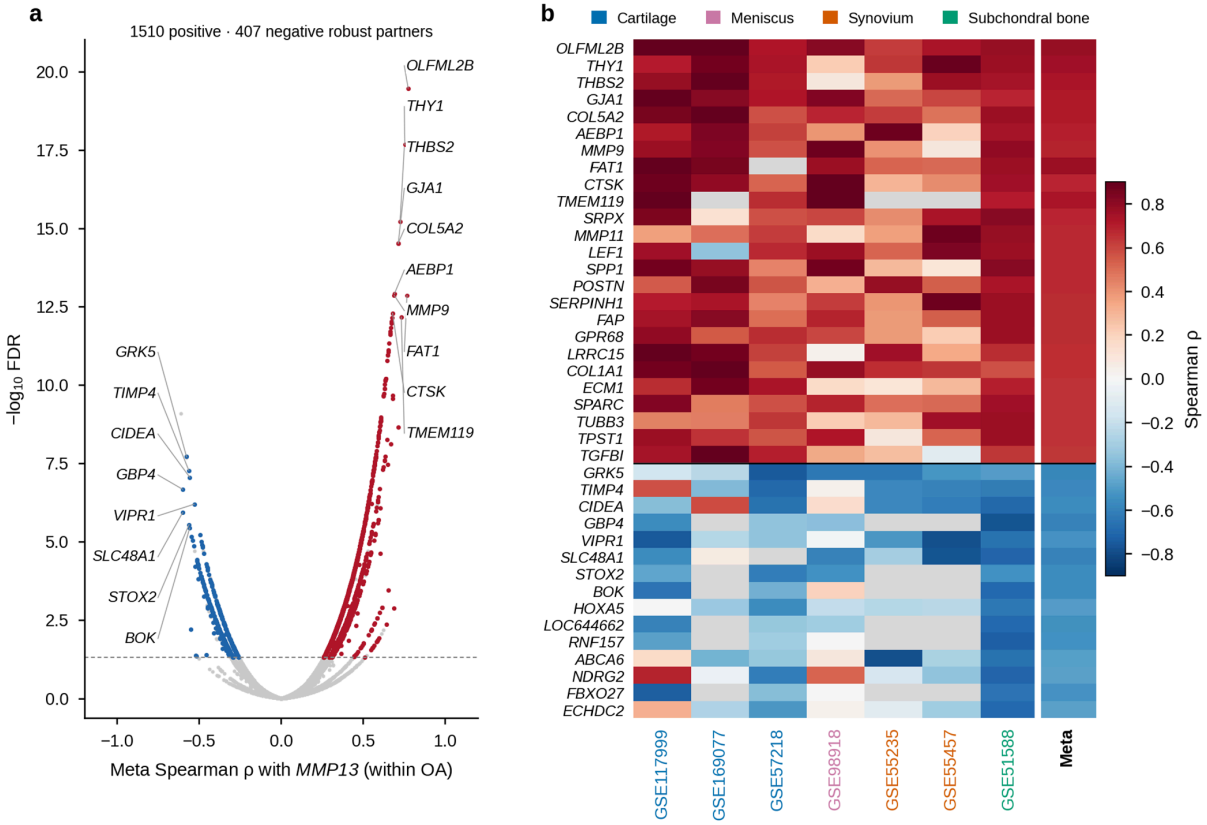

Figure 3. Consensus MMP13 co-expression neighbourhood within OA tissue. (**a**) Meta Spearman ρ with MMP13 versus −log10 FDR for 18,259 genes; red and blue, robust positive and negative partners; grey, other genes; dashed line, FDR = 0.05. (**b**) Per-dataset and meta correlations of the 25 strongest positive and 15 strongest negative partners; grey, gene not measured on that platform; colour scale, Spearman ρ; dataset labels are coloured by tissue.

Within OA tissue, MMP13 correlated with the fibroblast score (cohort-level meta ρ = 0.41) and the osteoblast score (0.42); in subchondral bone, it was also inversely correlated with the adipocyte and endothelial scores (Table S7). After adjustment for the first two principal components of the composition scores, raw and adjusted meta-correlations agreed at r = 0.74. In total, 363 positive and 79 negative genes remained robust partners (598 and 171 with one component). Of the top 150 positive partners, 124 remained significant with unchanged sign, but only 26 of the top 150 negative partners did. The strongest positive partners are thus only partly explained by the composition scores, whereas much of the negative module reflects differences in cell-type composition. The marker panels are approximations; the osteoblast panel, for example, consists largely of osteocyte genes (Table S7).

In control tissue, MMP13 co-expression was much weaker: meta-correlations agreed with those in OA tissue at r = 0.12, and only three genes met the robust-partner criteria. Of the locked positive partners, 77% were positively correlated with MMP13 in discovery controls, mostly weakly; in held-out cartilage, the fraction was 51% in controls, close to chance, against 80% in OA samples of the same series, and in held-out synovium, it was 79% in controls. Forty-one genes differed in their correlation with MMP13 between OA and control tissue at FDR < 0.05 (Table S8). This test has little power with control groups of 5–12 per dataset, and because the partners were selected in OA samples, it is biased towards larger correlations in OA.

Leaving out GSE169077, where OA and control samples had been combined, left 1383 positive and 680 negative robust partners; 1680 of the 1920 partners of the full rerun remained, and meta-correlations agreed at r = 0.98. With the Spearman-specific variance $1.06/(n-3)$, 1424 positive and 373 negative partners remained robust. Co-expression adjusted for both age and sex agreed with the unadjusted analysis at r = 0.97 and retained all 150 top positive and 149 of the 150 top negative partners (Table S11).

3.4. MMP13 Is Coupled to Hypertrophy-Associated Genes and MMP9 but Not to MMP1, MMP3, ADAMTS4 or ADAMTS5

Within OA tissue, MMP13 correlated with a hypertrophy score across 117 OA samples (z-scores within dataset; Spearman $\rho = 0.65$, $p = 2.4 \times 10^{-15}$; Figure 4b). For individual genes, meta-correlations were significant for SPP1 ($\rho = 0.67$), ALPL (0.47), PTH1R (0.45), COL10A1 (0.42) and RUNX2 (0.37; all FDR ≤ 0.0022), and also for MMP9 (0.69) and COL2A1 (0.32, FDR = 0.011). MMP1, MMP3, ADAMTS4 and ADAMTS5 showed no co-expression with MMP13 ($\rho = -0.04$ to 0.20; FDR ≥ 0.23; Figure 4a). The correlation differed between tissues: $\rho = 0.67$ (95% CI: 0.47 to 0.81) in cartilage (GSE57218 and GSE117999; $n = 45$), 0.70 (0.19 to 0.91) in meniscus ($n = 12$), 0.27 (−0.21 to 0.65) in synovium ($n = 20$) and 0.66 (0.43 to 0.81) in subchondral bone ($n = 40$) (Figure S6). The all-tissue estimate pools OA samples across datasets, excludes GSE169077 (fewer than eight OA samples) and counts the 12 patients of GSE117999 and GSE98918 twice, so its p value is descriptive. Without SPP1 and ALPL, which also mark osteoblasts, the score gave $\rho = 0.54$ across tissues and 0.63 in cartilage. The coupling therefore holds in cartilage alone and is weak in synovium.

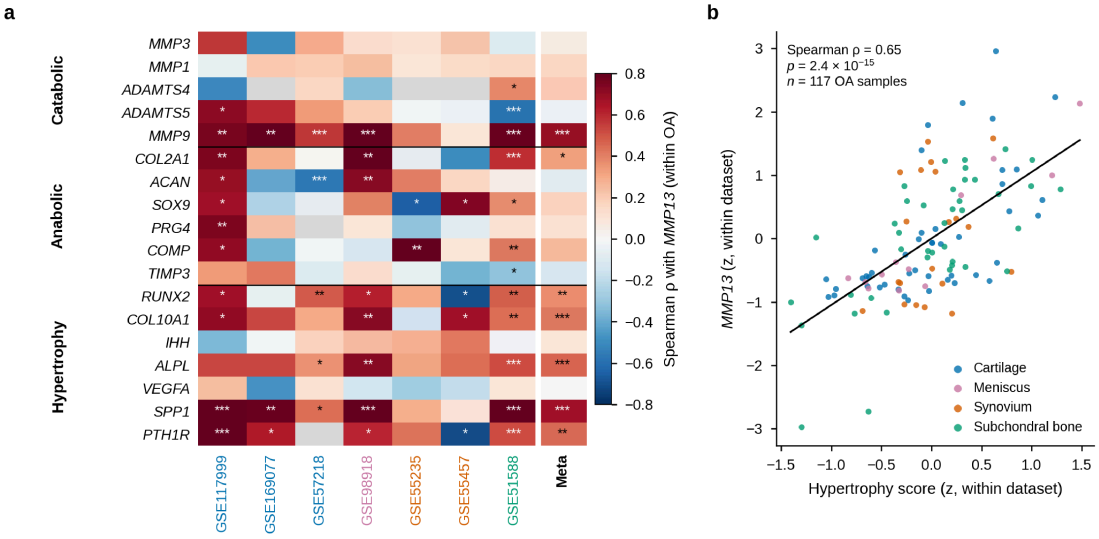

Figure 4. Coupling of MMP13 to matrix and hypertrophy programmes within OA tissue. (**a**) Spearman correlation of MMP13 with catabolic, anabolic and hypertrophy genes per dataset and as meta-estimate; *, ** and ***, p (datasets) or FDR (meta) < 0.05, <0.01 and <0.001, respectively; colour scale, Spearman

ρ. (**b**) MMP13 versus hypertrophy score in OA samples (z-scores within dataset; $n = 117$); line, least-squares fit.

3.5. The MMP13 Neighbourhood Is Enriched for Mesenchymal and Angiogenic Programmes

Genes ordered by consensus co-expression with MMP13 were enriched for epithelial–mesenchymal transition (normalised enrichment score, NES = 2.67), G2-M checkpoint (2.52), E2F targets (2.50), Myc targets V1 (2.45), mTORC1 signalling (2.17) and angiogenesis (2.17; all FDR < 0.001). Adipogenesis (NES = −1.82, FDR = 0.003), bile acid metabolism (−1.65, FDR = 0.003) and fatty acid metabolism (−1.44, FDR = 0.022; Figure 5a) were enriched in the opposite direction. In total, 32 Hallmark sets reached FDR < 0.05, 29 of them on the positive side.

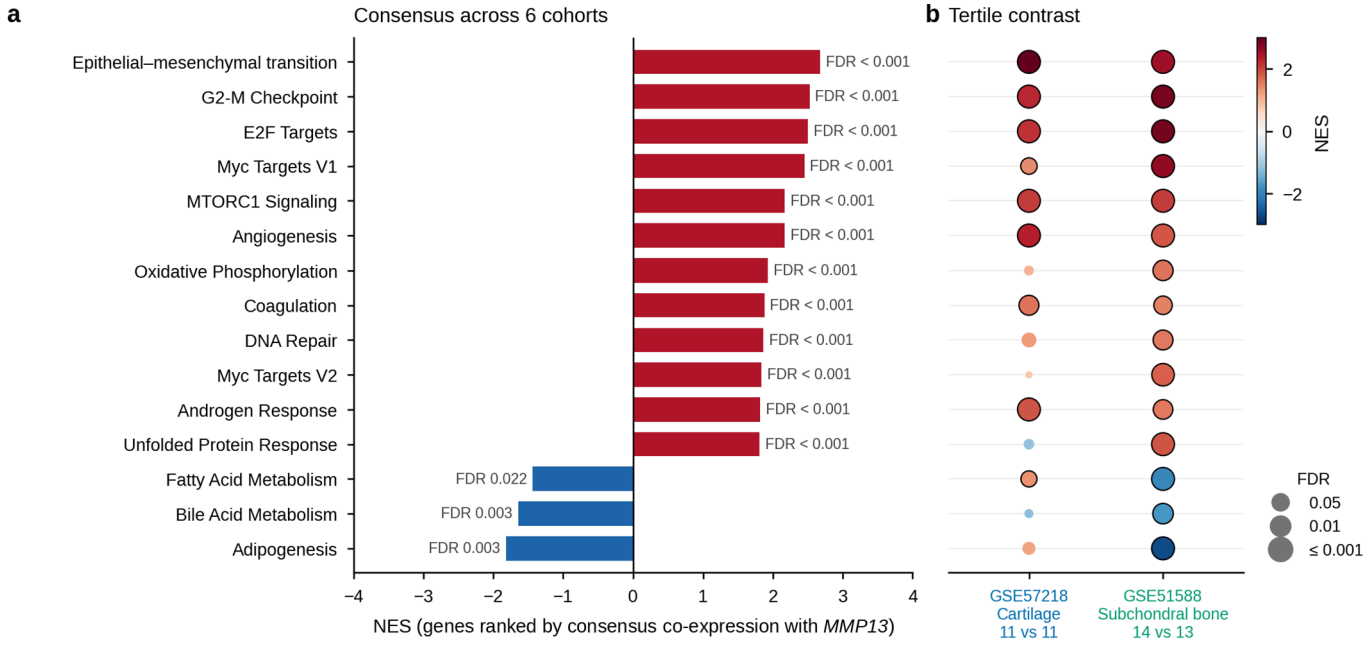

Figure 5. Hallmark programmes tracking MMP13 within OA tissue. (**a**) Pre-ranked GSEA of genes ordered by consensus co-expression with MMP13 (MSigDB Hallmark 2020, 1000 permutations); gene sets with FDR < 0.05, at most 12 per direction; red bars, positive NES; blue bars, negative NES. (**b**) MMP13-high versus MMP13-low tertiles within OA in GSE57218 (11 vs. 11) and GSE51588 (14 vs. 13); colour, NES; size, −log10 FDR; black outline, FDR < 0.05.

Tertile contrasts in the two datasets with at least 20 OA samples also showed enrichment of epithelial–mesenchymal transition, E2F targets and G2-M checkpoint (all FDR < 0.001; Figure 5b). Inflammatory programmes differed between tissues: TNF-α signalling via NF-κB was enriched in MMP13-high cartilage (NES = 3.07) but depleted in MMP13-high subchondral bone (NES = −2.47; both FDR < 0.001). With pre-ranked CAMERA, which compares each set with all other genes after allowing for inter-gene correlation, two sets remained at FDR < 0.05, epithelial–mesenchymal transition and angiogenesis, while G2-M checkpoint and E2F targets did not reach significance (FDR = 0.099 each). The sample-label permutation test, which tests whether a set is associated with MMP13 at all, without comparing it with other genes, retained 38 of 50 sets, all with a positive mean meta-Z, in line with the overall positive skew of MMP13 co-expression; no set on the negative side, including adipogenesis and fatty acid metabolism, and not TNF-α signalling via NF-κB, was retained by either correlation-aware test (Table S10). We therefore regard the mesenchymal and angiogenic enrichment as robust and the proliferative, metabolic and inflammatory enrichments as less certain.

3.6. Candidate Upstream Regulators of the Neighbourhood

The top 150 positive partners were dominated by extracellular-matrix and collagen pathways (Reactome extracellular matrix organization, adjusted $p = 5.0 \times 10^{-25}$; collagen biosynthesis and modifying enzymes, 4.3×10^{-18}). Transcription-factor libraries pointed to AP-1 components (ChEA c-JUN, 1.6×10^{-8}; TRRUST JUN, 1.0×10^{-5}; TRRUST FOS, 1.7×10^{-4}), RUNX2 (1.7×10^{-4}), SRF (1.6×10^{-4}) and TCF21 (2.4×10^{-5}; Figure 6a). Among the top 150 negative partners, enrichment was weak: the circadian regulators CLOCK (0.017) and ARNTL (0.035), with PPARA and PPARD at the threshold (0.0496; Figure 6b).

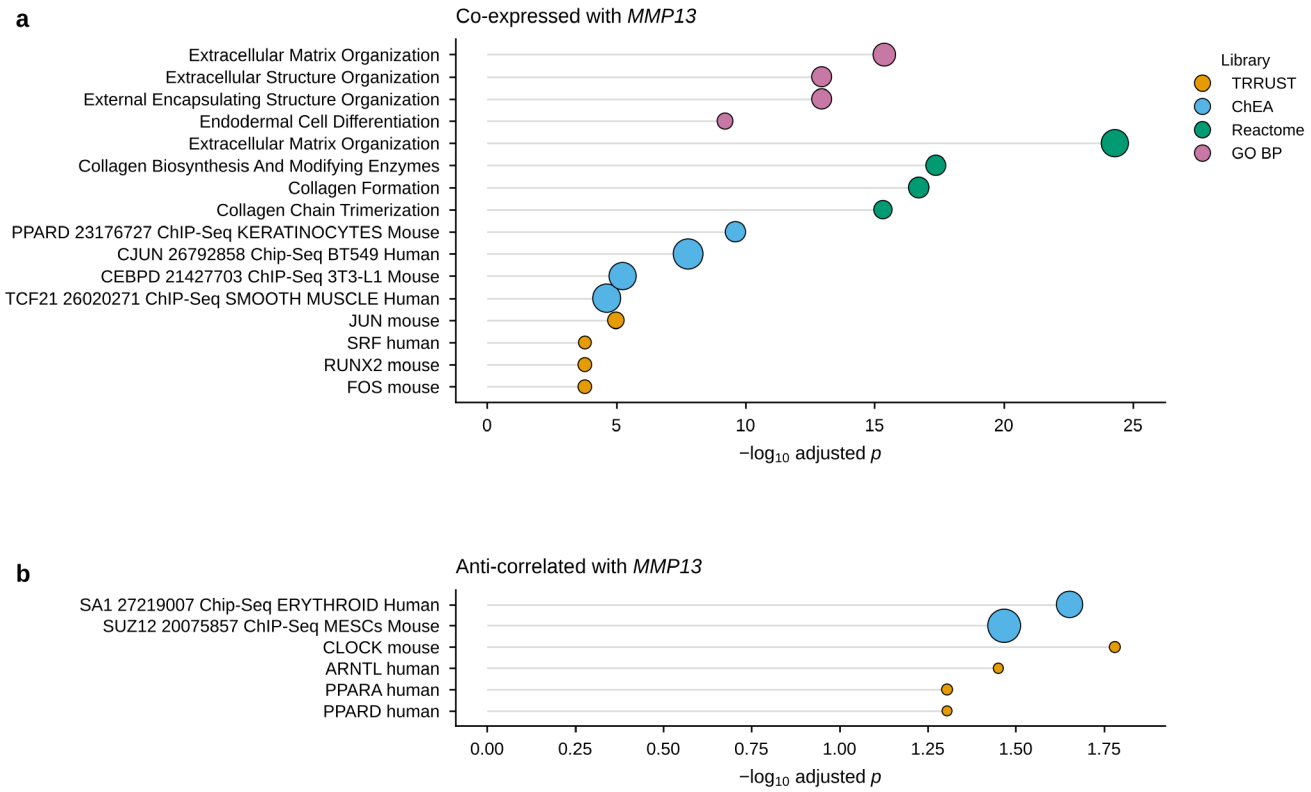

Figure 6. Pathways and transcription factors over-represented in the MMP13 neighbourhood. Enrichr analysis of the top 150 positive (**a**) and negative (**b**) partners against TRRUST 2019, ChEA 2022, Reactome 2022 and Gene Ontology (GO) Biological Process 2023, with all tested genes as the background; up to four terms per library with adjusted $p < 0.05$; dot size, number of overlapping genes.

3.7. Balance Indices Do Not Improve on MMP13 Alone

Neither composite index separated OA from control samples better than MMP13 alone. Across the seven datasets, the mean AUCs were 0.66 for MMP13 (range: 0.48–1.00), 0.54 for the Catabolic Switch Index, 0.50 for the Matrix Turnover Balance and 0.56 for the hypertrophy score (Figure S1). All of these comparisons are affected by the age imbalance described above.

3.8. Independent Validation of the MMP13 Programme

Of the pre-specified tests, those of the positive partners replicated in both held-out datasets, whereas the negative-module and hypertrophy-score correlations reached significance in synovium only (Tables 3 and S4; Figure 7). In OA knee cartilage RNA-sequencing (GSE114007; 20 OA donors), genome-wide correlations with MMP13 agreed with the discovery meta-correlations (Spearman r = 0.17, $p = 1.4 \times 10^{-113}$). Locked positive partners were shifted towards positive correlation (median $\rho = 0.29$ versus 0.04 for other genes;

79.7% positive; $p = 3.5 \times 10^{-22}$), and locked negative partners towards negative correlation (median $\rho = -0.14$; 71.1% negative; $p = 1.0 \times 10^{-18}$). Replication was stronger in OA synovium (GSE82107; 10 OA donors): r = 0.22 ($p = 3.2 \times 10^{-190}$), median $\rho = 0.52$ for positive partners (92.7% positive; $p = 9.1 \times 10^{-47}$) and −0.15 for negative partners (63.9% negative; $p = 9.6 \times 10^{-8}$). MMP13 correlated with the 50-gene positive module score in both datasets (cartilage $\rho = 0.58$, $p = 0.007$; synovium $\rho = 0.72$, $p = 0.019$) and inversely with the negative module score ($\rho = -0.42$, $p = 0.062$; $\rho = -0.67$, $p = 0.033$). Correlation with the hypertrophy score was positive in both but reached significance only in synovium (cartilage $\rho = 0.38$, $p = 0.099$; synovium $\rho = 0.67$, $p = 0.033$). Because GSE114007 was sequenced on two platforms (10 OA samples each), the tests were rerun after centring each gene within the platform. Replication was weaker but persisted: genome-wide r = 0.10; positive partners, median $\rho = 0.22$ (81.1% positive; $p = 1.6 \times 10^{-15}$); negative partners, median $\rho = -0.13$ (66.4% negative; $p = 1.7 \times 10^{-12}$); MMP13 versus the positive module score, $\rho = 0.57$ ($p = 0.009$). The inverse correlation with the negative module score did not persist ($\rho = -0.23$, $p = 0.34$). p values for the partner-set comparisons treat genes as independent and therefore overstate significance; the median correlations and the percentages of concordant partners are the more informative measures.

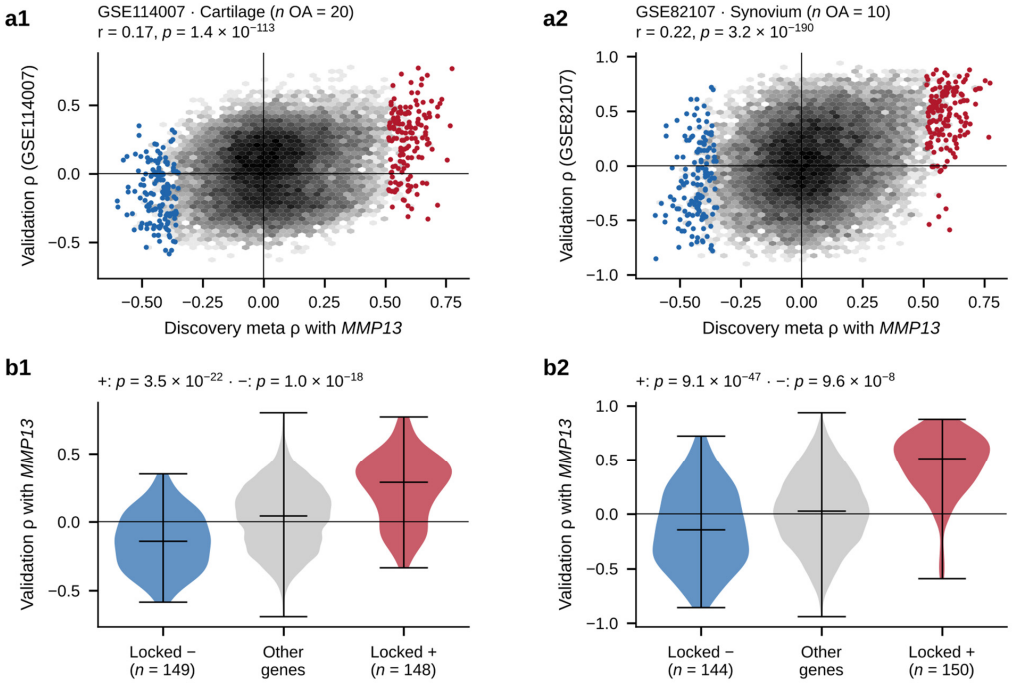

Figure 7. Pre-specified replication of the MMP13 co-expression neighbourhood in held-out data. (**a1**,**a2**) Genome-wide Spearman correlations with MMP13 within OA samples of GSE114007 cartilage ((**a1**), $n = 20$) and GSE82107 synovium ((**a2**), $n = 10$), plotted against the discovery meta-correlations (grey hexagonal bins, log-scaled density of all shared genes); locked positive (red) and negative (blue) partners are overlaid. Panel titles give Spearman r and p across all shared genes. (**b1**,**b2**) Distributions of validation correlations for locked negative partners (blue), all other genes (grey) and locked positive partners (red) in GSE114007 (b1) and GSE82107 (b2); horizontal lines mark medians and extremes; p values are one-sided Mann–Whitney tests against other genes.

Unlike most discovery datasets, both validation series showed more MMP13 in OA than in control tissue (cartilage log2 fold change 3.80, Welch $p = 3.0 \times 10^{-6}$; synovium 2.59, $p = 0.030$). In the cartilage series, controls were again much younger than OA donors (median: 34 versus 67.5 years; ranges: 18–61 and 51–82), but the OA effect remained after age adjustment ($p = 0.007$, against 2.1×10^{-6} unadjusted; $p = 0.002$ with the sequencing

platform as a further covariate). With only four controls aged 51 years or older, this adjusted estimate rests on a narrow range of overlapping ages. GSE82107 carried no age annotation.

In OA knee cartilage single-cell RNA-sequencing (GSE104782; 1597 cells after filtering), 49 of 50 positive and 44 of 50 negative module genes were detected, and Leiden clustering yielded nine clusters (Figure 8). Cluster 5 showed the highest mean MMP13 expression (0.47, 95% CI: 0.28 to 0.67), followed by cluster 8 (0.23, 0.09 to 0.40) and ≤0.08 in the others (Figure 8e). Its markers included IBSP, COL10A1, CHAD and MEF2C, and it also had the highest hypertrophy score (0.53 versus ≤0.03), to which COL10A1 contributes. The fraction of cells with detectable MMP13 was highest in cluster 7 and ranged across clusters (8.2–28.3%). The positive module score was highest in a small cluster of 43 cells marked by TMSB10, TMSB4X, COL1A1 and IGFBP7 (cluster 8; 0.94 versus ≤0.44), which also had the lowest negative module score. Cell by cell, MMP13 correlated with the positive module score ($\rho = 0.29$; $p = 8.2 \times 10^{-32}$, descriptive because cells are not independent); across the nine cluster means, the correlation was positive but did not reach significance ($\rho = 0.35$, $p = 0.36$). In the sensitivity analysis, neither the mitochondrial filter nor Scrublet flagged any cell, so the analysis did not test robustness to these filters, and re-clustering reproduced the primary clustering exactly (1597 cells, nine clusters, adjusted Rand index = 1.00); cluster 5 again had the highest mean MMP13 expression, with IBSP, COL10A1, CHAD and MEF2C among its markers (Figure S7; Table S12).

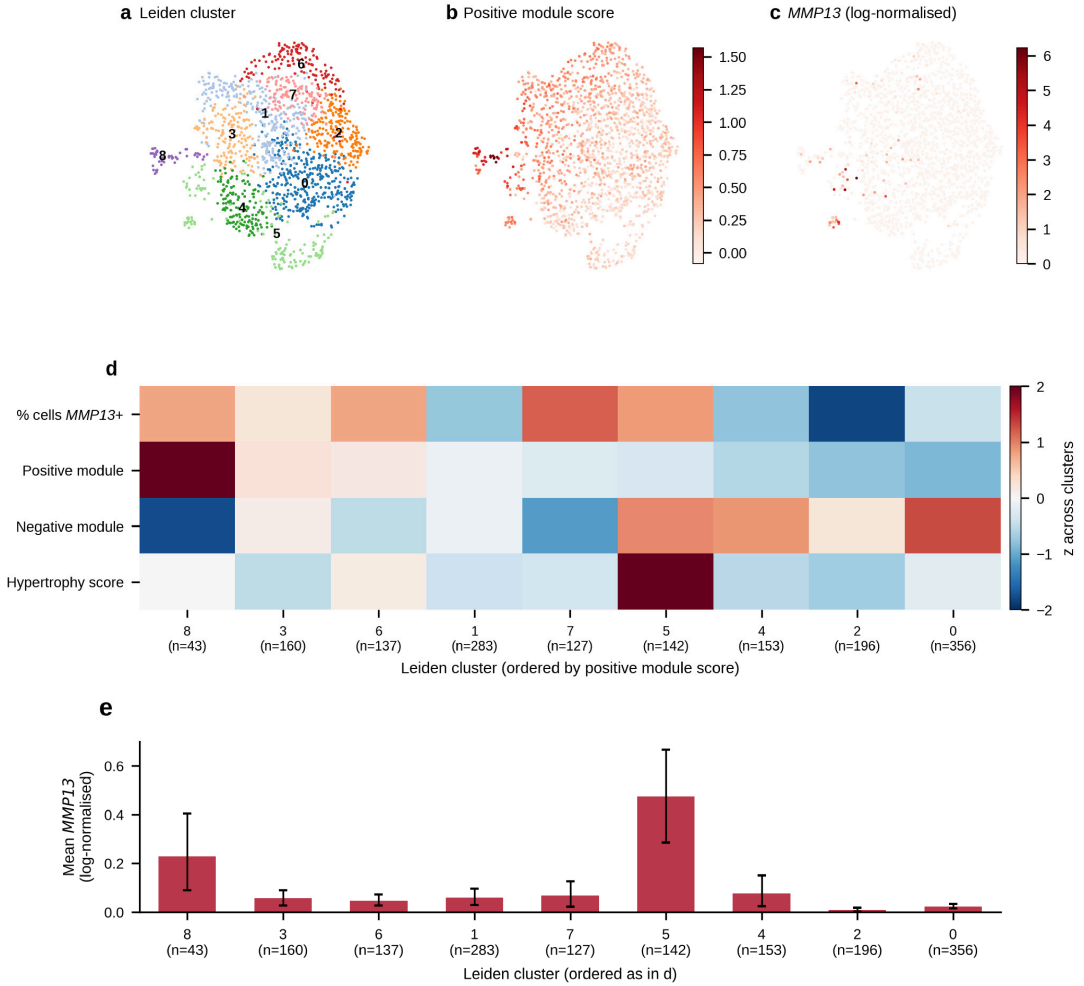

Figure 8. Single-cell localisation of MMP13 and the locked modules in OA knee cartilage (GSE104782; 1597 cells). (**a**) Uniform manifold approximation and projection (UMAP) embedding coloured by Leiden cluster. (**b**) Per-cell positive module score. (**c**) Log-normalised MMP13 expression. (**d**) Cluster

profiles, z-scored across clusters, of the percentage of MMP13-detected cells, the positive and negative module scores and the hypertrophy score; clusters are ordered by positive module score, with cell numbers in parentheses. (**e**) Mean log-normalised MMP13 expression per cluster; error bars, 95% bootstrap confidence intervals over cells. Clusters are described by marker genes only; no cell-type labels were assigned.

Table 3. Pre-specified validation of the locked MMP13 signature in held-out data.

Dataset	Tissue, Technology	OA/Control	Genome-Wide r	Positive Partners: Median ρ (% Positive)	Negative Partners: Median ρ (% Negative)	MMP13 vs. Positive Module, ρ (p)	MMP13 OA vs. Control, log2FC (p)
GSE114007	Knee cartilage, RNA-seq	20/18	0.17	0.29 (79.7%)	−0.14 (71.1%)	0.58 (0.007)	3.80 (3.0×10^{-6}); age-adjusted $p = 0.007$; age- and platform-adjusted $p = 0.002$
GSE82107	Synovium, microarray	10/7	0.22	0.52 (92.7%)	−0.15 (63.9%)	0.72 (0.019)	2.59 (0.030); age not annotated
GSE89408	Synovium, RNA-seq	—	—	—	—	—	Not used: expression table could not be matched to samples

Correlations are Spearman coefficients within OA samples. Other genes reached median ρ = 0.04 (GSE114007) and 0.03 (GSE82107). Shifts of locked partners relative to other genes: one-sided Mann–Whitney $p = 3.5 \times 10^{-22}$ (positive) and 1.0×10^{-18} (negative) in GSE114007; 9.1×10^{-47} and 9.6×10^{-8} in GSE82107.

4. Discussion

MMP13 expression in human OA transcriptomes behaved differently as a case–control marker and as a variable within OA tissue. As a marker that might separate OA from control tissue, it proved inconsistent and heterogeneous from one discovery dataset to the next, and its case–control differences could not be separated from donor age, which was lower in controls in every age-annotated cohort, the independent cartilage series included. Within OA tissue, MMP13 belonged to a reproducible co-expression programme that links hypertrophic differentiation, fibro-stromal matrix remodelling and, less certainly, proliferation, and that changed little with age adjustment. Locked before testing, its positive partners replicated in independent cartilage and synovium cohorts. MMP13 was not its hub: fibro-stromal partners such as FAP, THY1 and OLFML2B had larger and more central neighbourhoods, although MMP13's partners replicated better than those of most expression-matched genes.

A weak and heterogeneous OA-versus-control signal is consistent with the experimental evidence that MMP13 drives cartilage erosion [5,6]. It suggests that bulk mRNA levels measured across heterogeneous tissues are a poor readout of that role. Two features of public datasets contribute to this. First, controls were younger in every age-annotated dataset, and ageing alone alters joint tissue biology [23]. Second, control definitions vary, from healthy donors to non-OA knees with meniscal tears, which already carry degenerative change [20]. Ranking OA genes by differential expression between such groups therefore should be accompanied by explicit reporting of age and control sourcing.

The paired analysis adds a within-patient comparison. MMP13 was no higher in lesioned than in preserved cartilage from the same joint, even though in the same series it was higher in OA than in healthy donor cartilage by Welch's test ($p = 0.001$), although the confidence interval of Hedges' g included zero and four of the seven controls were adolescents. In their original work, Ramos et al. [18] observed that genes differing between lesioned and preserved cartilage overlapped with genes previously reported to differ between preserved OA cartilage and healthy cartilage. Our MMP13 result is more consistent with a joint-wide state than with a lesion-restricted event, although no formal preserved-versus-healthy test was run. Because preserved and lesioned samples share donor age, the paired comparison avoids the age confounding that affects OA-versus-control contrasts. A larger RNA-sequencing study of 35 paired samples from the same cohort reported 2387 differentially expressed mRNAs between lesioned and preserved cartilage [56]; our microarray-based paired result,

which approached but did not reach significance and was constrained by values near the platform floor, should be interpreted alongside that larger study.

The co-expression neighbourhood was the strongest and most reproducible finding. MMP13 tracked SPP1, ALPL, PTH1R, COL10A1 and RUNX2, genes of the hypertrophy-like programme that is proposed to push OA progression through protease-mediated matrix degradation [14]. This agrees with the observation that hypertrophy appears even in MMP13-deficient mice [6], which is compatible with MMP13 acting downstream of, or in parallel with, hypertrophic differentiation. MMP1, MMP3, ADAMTS4 and ADAMTS5, by contrast, showed no co-expression with MMP13 within OA tissue, which suggests that catabolic enzymes are not regulated as a single module across patients. In cartilage alone, the coupling was as strong as across tissues ($\rho = 0.67$), whereas in synovium it was weak ($\rho = 0.27$). In subchondral bone and synovium, SPP1, ALPL and RUNX2 also mark osteoblast-lineage cells, so there the association is better described as osteogenic than as chondrocyte hypertrophy (Figure S6).

The positive partners replicated in two cohorts that were independent of the discovery data, and one of them was profiled by RNA sequencing rather than microarrays, which argues against an artefact of one platform or cohort. In cartilage, the positive partners still replicated after centring for sequencing platforms, but the inverse correlation of MMP13 with the negative module did not, and the pre-specified hypertrophy-score correlation was significant in synovium but not in cartilage ($p = 0.099$), although hypertrophy coupling in discovery synovium was weak. Positive partners replicated more strongly than negative ones, and synovium replicated more strongly than cartilage, in keeping with the fibrostromal character of the top partners. The case–control result was less consistent: in the cartilage RNA-sequencing series of Fisch et al. [50], MMP13 was clearly higher in end-stage OA than in normal cartilage, and the difference persisted after age adjustment. This is compatible with the heterogeneous pooled estimate from the discovery arrays and shows that a strong case–control contrast can appear in some cohorts; since only four controls fell within the age range of the OA donors, no available dataset can fully separate disease from ageing. At single-cell resolution, the mean MMP13 expression was highest in the cluster expressing COL10A1 and IBSP, the clearest hypertrophic-like signal in the data, although its confidence interval overlapped that of a small COL1A1^{+} cluster in which the positive module peaked, even with cell-level intervals that ignore donor structure; the programme may therefore span more than one cell state. Ji et al. [17] defined chondrocyte populations in this dataset by their own criteria; we did not map our unsupervised clusters onto their labels, and these single-cell observations rest on cells from a single study.

Our results also bear on how public OA datasets are used. Bioinformatic studies have repeatedly nominated hub genes from case–control contrasts in small series such as GSE169077 [24–26]. That series produced the largest MMP13 effect of any dataset in our analysis ($g = 3.41$), yet the pooled effect across cohorts fell short of significance and every age-annotated dataset had younger controls. Case–control rankings drawn from such series should therefore be treated as hypotheses until they replicate across cohorts and are tested for sensitivity to age. GSE169077 also carried no age annotation, and its arrays may represent pooled RNA [34], which would lower within-group variance and inflate standardised effect sizes. Without it, the pooled case–control estimate fell to $g = 0.41$ (95% CI: −0.21 to 1.03), heterogeneity fell from 67% to 28%, and 1680 of the 1920 robust partners of the rerun remained (Table S11).

The within-OA perspective is consistent with evidence that OA patients form molecular subgroups. Unsupervised clustering of knee cartilage transcriptomes has identified two [28], two [27] and four [29] patient subgroups. Soul et al. [28] went on to report suppressed circadian regulators in both of their subgroups, which agrees with the pres-

ence of CLOCK and ARNTL targets among genes anti-correlated with MMP13 here. Our approach complements that work: without partitioning patients into subgroups, it relates variation among patients to a single therapeutic target and asks which programme varies with it in every cohort. Mechanistic studies of the MMP13–hypertrophy coupling are also consistent with our results, showing that HIF-2α transactivates COL10A1, MMP13 and VEGFA [30], and that Indian hedgehog tracks COL10A1, MMP13 and severity in human OA cartilage [31]. The present analysis adds that this coupling is reproducible across cohorts and tissues and changes little with age adjustment in patients with end-stage OA. Fisch et al. [50] found AP-1 family transcription factors, JUN and FOS among them, targeting many dysregulated genes in OA cartilage; their target sets were also over-represented among MMP13 partners here.

The positive partners included fibro-stromal genes (THY1, POSTN, FAP, THBS2 and COL5A2) and bone-remodelling genes (CTSK, MMP9 and SPP1). Periostin is one such partner, and silencing it reduced MMP13 expression through NF-κB in a mouse model of post-traumatic OA [57]. Single-cell studies have identified fibrocartilage and hypertrophic chondrocyte populations in OA cartilage [17]. Bulk co-expression cannot assign our module to a cell type, but it predicts that MMP13-high tissue is enriched for these states. Our single-cell analysis supports this in part (Figure 8), and spatial data could test it directly.

The top partners, which include collagens, THY1, POSTN, FAP and genes of epithelial–mesenchymal transition, overlap with a fibroblast activation axis seen in many tissues. A cross-tissue single-cell atlas identified fibroblast states, including $SPARC^{+}COL3A1^{+}$ vascular-interacting fibroblasts, that are expanded in inflamed synovium, intestine, lung and salivary gland [58]. Part of the MMP13 neighbourhood may therefore reflect how much activated stroma a sample contains rather than a programme specific to OA. Two analyses address this. After adjustment for marker-based composition scores, 124 of the top 150 positive partners remained but only 26 of the top 150 negative partners, and the number of robust partners fell from 1920 in the rerun to 442, so composition explains part of the positive module and much of the negative one, for example the adipocyte content of subchondral bone. In control tissue, the neighbourhood was much weaker ($r = 0.12$ with the OA meta-correlations; three robust partners). In held-out cartilage, 80% of locked positive partners were positively correlated with MMP13 in OA samples compared with 51% in control samples, which suggests that the coupling is stronger in OA tissue. This comparison cannot establish that the coupling is specific to OA: control groups were small, MMP13 is lower and may vary less in control tissue, which weakens correlations, and the partners were selected in OA samples. Marker scores are crude proxies for cell proportions, reference-based deconvolution was not performed, and neither analysis replaces cell-resolved data.

Regulator analyses suggest several links to known biology. AP-1 (JUN, FOS) and RUNX2 targets were over-represented among positive partners, in line with the known transcriptional control of MMP13 [15]. MMP9, a strong partner, is among the catabolic genes directly induced by HIF-2α in chondrocytes [16]. Lipid-metabolic and adipogenic sets, and more weakly PPAR and circadian targets, ranked on the negative side, although these sets did not survive correlation-aware testing. Cartilage-specific PPARγ deficiency leads to accelerated OA with aberrant mTOR signalling [59], and mTORC1 signalling ranked on the MMP13-high side in pre-ranked GSEA, though not in CAMERA. Loss of the chondrocyte clock gene Bmal1 (ARNTL) disrupts cartilage homeostasis [60]. These links are hypotheses; over-representation of curated targets does not demonstrate regulator activity in these samples.

In two discovery datasets, TNF-α/NF-κB signalling ranked in opposite directions with MMP13 in cartilage and in subchondral bone. This was not confirmed by the correlation-aware tests and may reflect differences between datasets as much as between tissues.

This study has several limitations. Discovery data were bulk microarrays from public series with small control groups (5–12 per dataset) and heterogeneous control definitions. Donor age was annotated in five of seven discovery datasets and in one of two validation datasets, and group and age were strongly correlated, so age-adjusted OA effects, including the significant adjusted effect in GSE114007, remain imprecise. Most OA donors were older patients undergoing surgery for end-stage disease, so the small effect of age adjustment on co-expression holds only within that age range. Two discovery series come from the same patients and were treated as one cohort. In GSE57218, many MMP13 values were near the platform floor. One pre-specified validation series (GSE89408) could not be used because its expression table could not be matched to the sample annotations, and the validation cohorts were small (20 and 10 OA donors). The single-cell data came from one study, so cluster-level associations were underpowered. The datasets were selected from a list compiled by the authors without a systematic database search, so other eligible GEO series may have been missed and selection bias cannot be excluded. Control tissues were of uneven quality: GSE57218 included adolescent donors, control synovium came from post-mortem or trauma donors without histological grading, and GSE57218 combined hip and knee joints. Sex had to be inferred from expression in two datasets, and marker-based composition scores are only proxies for cell proportions. The deposited single-cell matrix carries no donor labels, so donor effects could not be corrected and cluster structure may partly reflect donors. In GSE169077, which has fewer than eight OA samples, co-expression was computed across OA and control samples, so its contribution partly reflects the case–control difference (the analysis without it is given in Table S11). Many tests were reported without correction across analyses, and p values from comparisons that treat genes or cells as independent overstate significance. Transcription-factor and drug–gene results rely on external databases and indicate associations, not mechanisms. Finally, co-expression does not establish causality; perturbing MMP13 and its partners, along with spatial or protein-level localisation, will be required.

5. Conclusions

Across four joint tissues, MMP13 was an inconsistent marker of OA versus control tissue in the discovery data, its case–control differences could not be separated from donor age, and it was no higher in lesioned than in preserved cartilage of the same joint. Within OA tissue, it was part of a reproducible programme of fibro-stromal matrix remodelling and hypertrophy-associated genes and, less certainly, proliferation, while its negative partners largely reflected tissue composition. MMP13 was a consistent member of this programme, but FAP, THY1 and OLFML2B were more central. Its positive partners replicated in independent cartilage and synovium cohorts, the pre-specified hypertrophy-score correlation only in synovium, and in single-cell data, the mean MMP13 expression was highest in a cluster with hypertrophy-associated markers. Stratifying patients by this MMP13-associated programme, rather than by MMP13 level alone, may prove more informative for mechanistic studies and for selecting patients for MMP13-targeted therapy.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cimb48121292/s1.

Author Contributions: Conceptualization, S.E., E.S.A. and H.Ş.; methodology, S.E. and H.Ş.; software, S.E.; validation, S.E., H.G. and H.Ş.; formal analysis, S.E., H.G. and H.Ş.; investigation, S.E., E.D.A., H.G. and H.Ş.; resources, E.S.A. and H.Ş.; data curation, S.E., E.D.A. and H.G.; writing—original draft preparation, S.E., E.D.A. and H.Ş.; writing—review and editing, S.E., E.S.A., H.G. and H.Ş.; visualization, S.E. and H.G.; supervision, E.S.A. and H.Ş.; project administration, H.Ş. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were not required for this study because it involved only secondary analysis of publicly available, anonymised data from the Gene Expression Omnibus, which contain no identifiable personal information; no new data were collected from human participants. Under Article 28(1)(b) of the Turkish Personal Data Protection Law No. 6698, the Law does not apply to personal data processed in anonymised form for research purposes, and under the U.S. Common Rule (45 CFR 46.102(e)(1)) the analysis of publicly available, non-identifiable data does not constitute human subjects' research. Ethical approval and informed consent for the original samples were obtained by the original investigators, as described in the source publications and GEO records.

Informed Consent Statement: Patient consent was waived because this study used only publicly available, anonymised data. Under Article 28 (1)(b) of the Turkish Personal Data Protection Law No. 6698 and the U.S. Common Rule (45 CFR 46.102 (e)(1)), analysis of anonymised public data does not require new informed consent. Ethical approval and informed consent for the original samples were obtained by the original investigators, as described in the source publications and GEO records.

Data Availability Statement: The data presented in this study are available in GitHub at https://github.com/Sajjadeslamkhah/mmp13-oa-coexpression (accessed on 19 November 2026), release v1.0.0 (analysis notebooks, locked validation signature, result tables and figure source data). These data were derived from the following resources available in the public domain: the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/, accessed on 19 November 2026), series GSE57218, GSE117999, GSE98918, GSE169077, GSE51588, GSE55235, GSE55457, GSE114007, GSE82107 and GSE104782.

Acknowledgments: The authors thank Google Colaboratory (Google LLC, Mountain View, CA, USA; hosted service without version numbers) for the cloud computing environment in which all analyses were run. During the preparation of this study and its revision, the authors used Claude (Anthropic, San Francisco, CA, USA; Claude Opus 5.5 during the revision) to assist with writing and debugging the analysis code, with the additional analyses, figures and Supplementary Tables of the revision, and with preparing and language-editing the manuscript and the responses to the reviewers. The authors subsequently revised the text, independently verified all analyses, reported values and references, and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. GBD 2021 Osteoarthritis Collaborators. Global, regional, and national burden of osteoarthritis, 1990–2020 and projections to 2050: A systematic analysis for the Global Burden of Disease Study 2021. *Lancet Rheumatol.* **2023**, *5*, e508–e522. [CrossRef] [PubMed]
2. Hunter, D.J.; Bierma-Zeinstra, S. Osteoarthritis. *Lancet* **2019**, *393*, 1745–1759. [CrossRef] [PubMed]
3. Loeser, R.F.; Goldring, S.R.; Scanzello, C.R.; Goldring, M.B. Osteoarthritis: A disease of the joint as an organ. *Arthritis Rheum.* **2012**, *64*, 1697–1707. [CrossRef] [PubMed]
4. Mitchell, P.G.; Magna, H.A.; Reeves, L.M.; Lopresti-Morrow, L.L.; Yocum, S.A.; Rosner, P.J.; Geoghegan, K.F.; Hambor, J.E. Cloning, expression, and type II collagenolytic activity of matrix metalloproteinase-13 from human osteoarthritic cartilage. *J. Clin. Investig.* **1996**, *97*, 761–768. [CrossRef] [PubMed]
5. Neuhold, L.A.; Killar, L.; Zhao, W.; Sung, M.L.; Warner, L.; Kulik, J.; Turner, J.; Wu, W.; Billinghurst, C.; Meijers, T.; et al. Postnatal expression in hyaline cartilage of constitutively active human collagenase-3 (MMP-13) induces osteoarthritis in mice. *J. Clin. Investig.* **2001**, *107*, 35–44. [CrossRef] [PubMed]
6. Little, C.B.; Barai, A.; Burkhardt, D.; Smith, S.M.; Fosang, A.J.; Werb, Z.; Shah, M.; Thompson, E.W. Matrix metalloproteinase 13–deficient mice are resistant to osteoarthritic cartilage erosion but not chondrocyte hypertrophy or osteophyte development. *Arthritis Rheum.* **2009**, *60*, 3723–3733. [CrossRef] [PubMed]
7. Krzeski, P.; Buckland-Wright, C.; Bálint, G.; Cline, G.A.; Stoner, K.; Lyon, R.; Beary, J.; Aronstein, W.S.; Spector, T.D. Development of musculoskeletal toxicity without clear benefit after administration of PG-116800, a matrix metalloproteinase inhibitor, to patients with knee osteoarthritis: A randomized, 12-month, double-blind, placebo-controlled study. *Arthritis Res. Ther.* **2007**, *9*, R109. [CrossRef] [PubMed]

8. Johnson, A.R.; Pavlovsky, A.G.; Ortwine, D.F.; Prior, F.; Man, C.F.; Bornemeier, D.A.; Banotai, C.A.; Mueller, W.T.; McConnell, P.; Yan, C.; et al. Discovery and characterization of a novel inhibitor of matrix metalloprotease-13 that reduces cartilage damage in vivo without joint fibroplasia side effects. *J. Biol. Chem.* **2007**, *282*, 27781–27791. [CrossRef] [PubMed]
9. Baragi, V.M.; Becher, G.; Bendele, A.M.; Biesinger, R.; Bluhm, H.; Boer, J.; Deng, H.; Dodd, R.; Essers, M.; Feuerstein, T.; et al. A new class of potent matrix metalloproteinase 13 inhibitors for potential treatment of osteoarthritis: Evidence of histologic and clinical efficacy without musculoskeletal toxicity in rat models. *Arthritis Rheum.* **2009**, *60*, 2008–2018. [CrossRef] [PubMed]
10. Bedingfield, S.K.; Colazo, J.M.; Yu, F.; Liu, D.D.; Jackson, M.A.; Himmel, L.E.; Cho, H.; Crofford, L.J.; Hasty, K.A.; Duvall, C.L. Amelioration of post-traumatic osteoarthritis via nanoparticle depots delivering small interfering RNA to damaged cartilage. *Nat. Biomed. Eng.* **2021**, *5*, 1069–1083. [CrossRef] [PubMed]
11. Hu, Q.; Ecker, M. Overview of MMP-13 as a promising target for the treatment of osteoarthritis. *Int. J. Mol. Sci.* **2021**, *22*, 1742. [CrossRef] [PubMed]
12. Tchetina, E.; Mwale, F.; Poole, A.R. Distinct phases of coordinated early and late gene expression in growth plate chondrocytes in relationship to cell proliferation, matrix assembly, remodeling, and cell differentiation. *J. Bone Miner. Res.* **2003**, *18*, 844–851. [CrossRef] [PubMed]
13. Behonick, D.J.; Xing, Z.; Lieu, S.; Buckley, J.M.; Lotz, J.C.; Marcucio, R.S.; Werb, Z.; Miclau, T.; Colnot, C. Role of matrix metalloproteinase 13 in both endochondral and intramembranous ossification during skeletal regeneration. *PLoS ONE* **2007**, *2*, e1150. [CrossRef] [PubMed]
14. van der Kraan, P.M.; van den Berg, W.B. Chondrocyte hypertrophy and osteoarthritis: Role in initiation and progression of cartilage degeneration? *Osteoarthr. Cartil.* **2012**, *20*, 223–232. [CrossRef] [PubMed]
15. Vincenti, M.P.; Brinckerhoff, C.E. Transcriptional regulation of collagenase (MMP-1, MMP-13) genes in arthritis: Integration of complex signaling pathways for the recruitment of gene-specific transcription factors. *Arthritis Res.* **2002**, *4*, 157–164. [CrossRef] [PubMed]
16. Yang, S.; Kim, J.; Ryu, J.H.; Oh, H.; Chun, C.H.; Kim, B.J.; Min, B.H.; Chun, J.S. Hypoxia-inducible factor-2α is a catabolic regulator of osteoarthritic cartilage destruction. *Nat. Med.* **2010**, *16*, 687–693. [CrossRef] [PubMed]
17. Ji, Q.; Zheng, Y.; Zhang, G.; Hu, Y.; Fan, X.; Hou, Y.; Wen, L.; Li, L.; Xu, Y.; Wang, Y.; et al. Single-cell RNA-seq analysis reveals the progression of human osteoarthritis. *Ann. Rheum. Dis.* **2019**, *78*, 100–110. [CrossRef] [PubMed]
18. Ramos, Y.F.; den Hollander, W.; Bovée, J.V.; Bomer, N.; van der Breggen, R.; Lakenberg, N.; Keurentjes, J.C.; Goeman, J.J.; Slagboom, P.E.; Nelissen, R.G.; et al. Genes involved in the osteoarthritis process identified through genome wide expression analysis in articular cartilage; the RAAK study. *PLoS ONE* **2014**, *9*, e103056. [CrossRef] [PubMed]
19. Brophy, R.H.; Zhang, B.; Cai, L.; Wright, R.W.; Sandell, L.J.; Rai, M.F. Transcriptome comparison of meniscus from patients with and without osteoarthritis. *Osteoarthr. Cartil.* **2018**, *26*, 422–432. [CrossRef] [PubMed]
20. Rai, M.F.; Tycksen, E.D.; Cai, L.; Yu, J.; Wright, R.W.; Brophy, R.H. Distinct degenerative phenotype of articular cartilage from knees with meniscus tear compared to knees with osteoarthritis. *Osteoarthr. Cartil.* **2019**, *27*, 945–955. [CrossRef] [PubMed]
21. Woetzel, D.; Huber, R.; Kupfer, P.; Pohlers, D.; Pfaff, M.; Driesch, D.; Häupl, T.; Koczan, D.; Stiehl, P.; Guthke, R.; et al. Identification of rheumatoid arthritis and osteoarthritis patients by transcriptome-based rule set generation. *Arthritis Res. Ther.* **2014**, *16*, R84. [CrossRef] [PubMed]
22. Chou, C.H.; Wu, C.C.; Song, I.W.; Chuang, H.P.; Lu, L.S.; Chang, J.H.; Kuo, S.Y.; Lee, C.H.; Wu, J.Y.; Chen, Y.T.; et al. Genome-wide expression profiles of subchondral bone in osteoarthritis. *Arthritis Res. Ther.* **2013**, *15*, R190. [CrossRef] [PubMed]
23. Loeser, R.F.; Collins, J.A.; Diekman, B.O. Ageing and the pathogenesis of osteoarthritis. *Nat. Rev. Rheumatol.* **2016**, *12*, 412–420. [CrossRef] [PubMed]
24. Li, Z.; Xie, L.; Zeng, H.; Wu, Y. PDK4 inhibits osteoarthritis progression by activating the PPAR pathway. *J. Orthop. Surg. Res.* **2024**, *19*, 109. [CrossRef] [PubMed]
25. You, X.; Ye, Y.; Lin, S.; Zhang, Z.; Guo, H.; Ye, H. Identification of key genes and immune infiltration in osteoarthritis through analysis of zinc metabolism-related genes. *BMC Musculoskelet. Disord.* **2024**, *25*, 227. [CrossRef] [PubMed]
26. Zhong, J.; Xiang, D.; Ma, X. Prediction and analysis of osteoarthritis hub genes with bioinformatics. *Ann. Transl. Med.* **2023**, *11*, 66. [CrossRef] [PubMed]
27. Coutinho de Almeida, R.; Mahfouz, A.; Mei, H.; Houtman, E.; den Hollander, W.; Soul, J.; Suchiman, E.; Lakenberg, N.; Meessen, J.; Huetink, K.; et al. Identification and characterization of two consistent osteoarthritis subtypes by transcriptome and clinical data integration. *Rheumatology* **2021**, *60*, 1166–1175. [CrossRef] [PubMed]
28. Soul, J.; Dunn, S.L.; Anand, S.; Serracino-Inglott, F.; Schwartz, J.M.; Boot-Handford, R.P.; Hardingham, T.E. Stratification of knee osteoarthritis: Two major patient subgroups identified by genome-wide expression analysis of articular cartilage. *Ann. Rheum. Dis.* **2018**, *77*, 423–430. Correction in *Ann. Rheum. Dis.* **2018**, *77*, 960. https://doi.org/10.1136/annrheumdis-2017-212603corr1. [CrossRef] [PubMed]
29. Yuan, C.; Pan, Z.; Zhao, K.; Li, J.; Sheng, Z.; Yao, X.; Liu, H.; Zhang, X.; Yang, Y.; Yu, D.; et al. Classification of four distinct osteoarthritis subtypes with a knee joint tissue transcriptome atlas. *Bone Res.* **2020**, *8*, 38. [CrossRef] [PubMed]

30. Saito, T.; Fukai, A.; Mabuchi, A.; Ikeda, T.; Yano, F.; Ohba, S.; Nishida, N.; Akune, T.; Yoshimura, N.; Nakagawa, T.; et al. Transcriptional regulation of endochondral ossification by HIF-2α during skeletal growth and osteoarthritis development. *Nat. Med.* **2010**, *16*, 678–686. [CrossRef] [PubMed]
31. Wei, F.; Zhou, J.; Wei, X.; Zhang, J.; Fleming, B.C.; Terek, R.; Pei, M.; Chen, Q.; Liu, T.; Wei, L. Activation of Indian hedgehog promotes chondrocyte hypertrophy and upregulation of MMP-13 in human osteoarthritic cartilage. *Osteoarthr. Cartil.* **2012**, *20*, 755–763. [CrossRef] [PubMed]
32. Barrett, T.; Wilhite, S.E.; Ledoux, P.; Evangelista, C.; Kim, I.F.; Tomashevsky, M.; Marshall, K.A.; Phillippy, K.H.; Sherman, P.M.; Holko, M.; et al. NCBI GEO: Archive for functional genomics data sets—Update. *Nucleic Acids Res.* **2013**, *41*, D991–D995. [CrossRef] [PubMed]
33. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [CrossRef] [PubMed]
34. Attur, M.; Yang, Q.; Shimada, K.; Tachida, Y.; Nagase, H.; Mignatti, P.; Statman, L.; Palmer, G.; Kirsch, T.; Beier, F.; et al. Elevated expression of periostin in human osteoarthritic cartilage and its potential role in matrix degradation via matrix metalloproteinase-13. *FASEB J.* **2015**, *29*, 4107–4121. [CrossRef] [PubMed]
35. DerSimonian, R.; Laird, N. Meta-analysis in clinical trials. *Control. Clin. Trials* **1986**, *7*, 177–188. [CrossRef] [PubMed]
36. IntHout, J.; Ioannidis, J.P.; Borm, G.F. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med. Res. Methodol.* **2014**, *14*, 25. [CrossRef] [PubMed]
37. Srikanth, V.K.; Fryer, J.L.; Zhai, G.; Winzenberg, T.M.; Hosmer, D.; Jones, G. A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis. *Osteoarthr. Cartil.* **2005**, *13*, 769–781. [CrossRef]
38. Whitlock, M.C. Combining probability from independent tests: The weighted Z-method is superior to Fisher's approach. *J. Evol. Biol.* **2005**, *18*, 1368–1373. [CrossRef] [PubMed]
39. Fieller, E.C.; Hartley, H.O.; Pearson, E.S. Tests for rank correlation coefficients. I. *Biometrika* **1957**, *44*, 470–481. [CrossRef]
40. Becht, E.; Giraldo, N.A.; Lacroix, L.; Buttard, B.; Elarouci, N.; Petitprez, F.; Selves, J.; Laurent-Puig, P.; Sautès-Fridman, C.; Fridman, W.H.; et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* **2016**, *17*, 218. Erratum in *Genome Biol.* **2016**, *17*, 249. https://doi.org/10.1186/s13059-016-1113-y. [CrossRef] [PubMed]
41. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550. [CrossRef] [PubMed]
42. Fang, Z.; Liu, X.; Peltz, G. GSEApy: A comprehensive package for performing gene set enrichment analysis in Python. *Bioinformatics* **2023**, *39*, btac757. [CrossRef] [PubMed]
43. Liberzon, A.; Birger, C.; Thorvaldsdóttir, H.; Ghandi, M.; Mesirov, J.P.; Tamayo, P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **2015**, *1*, 417–425. [CrossRef] [PubMed]
44. Kuleshov, M.V.; Jones, M.R.; Rouillard, A.D.; Fernandez, N.F.; Duan, Q.; Wang, Z.; Koplev, S.; Jenkins, S.L.; Jagodnik, K.M.; Lachmann, A.; et al. Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **2016**, *44*, W90–W97. [CrossRef] [PubMed]
45. Han, H.; Cho, J.W.; Lee, S.; Yun, A.; Kim, H.; Bae, D.; Yang, S.; Kim, C.Y.; Lee, M.; Kim, E.; et al. TRRUST v2: An expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res.* **2018**, *46*, D380–D386. [CrossRef] [PubMed]
46. Lachmann, A.; Xu, H.; Krishnan, J.; Berger, S.I.; Mazloom, A.R.; Ma'ayan, A. ChEA: Transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics* **2010**, *26*, 2438–2444. [CrossRef] [PubMed]
47. Gillespie, M.; Jassal, B.; Stephan, R.; Milacic, M.; Rothfels, K.; Senff-Ribeiro, A.; Griss, J.; Sevilla, C.; Matthews, L.; Gong, C.; et al. The Reactome pathway knowledgebase 2022. *Nucleic Acids Res.* **2022**, *50*, D687–D692. [CrossRef] [PubMed]
48. Wu, D.; Smyth, G.K. Camera: A competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Res.* **2012**, *40*, e133. [CrossRef] [PubMed]
49. Cannon, M.; Stevenson, J.; Stahl, K.; Basu, R.; Coffman, A.; Kiwala, S.; McMichael, J.F.; Kuzma, K.; Morrissey, D.; Cotto, K.; et al. DGIdb 5.0: Rebuilding the drug–gene interaction database for precision medicine and drug discovery platforms. *Nucleic Acids Res.* **2024**, *52*, D1227–D1235. [CrossRef] [PubMed]
50. Fisch, K.M.; Gamini, R.; Alvarez-Garcia, O.; Akagi, R.; Saito, M.; Muramatsu, Y.; Sasho, T.; Koziol, J.A.; Su, A.I.; Lotz, M.K. Identification of transcription factors responsible for dysregulated networks in human osteoarthritis cartilage by global gene expression analysis. *Osteoarthr. Cartil.* **2018**, *26*, 1531–1538. [CrossRef] [PubMed]
51. Xin, J.; Mark, A.; Afrasiabi, C.; Tsueng, G.; Juchler, M.; Gopal, N.; Stupp, G.S.; Putman, T.E.; Ainscough, B.J.; Griffith, O.L.; et al. High-performance web services for querying gene and variant annotation. *Genome Biol.* **2016**, *17*, 91. [CrossRef] [PubMed]

52. Wolf, F.A.; Angerer, P.; Theis, F.J. SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* **2018**, *19*, 15. [CrossRef] [PubMed]
53. Traag, V.A.; Waltman, L.; van Eck, N.J. From Louvain to Leiden: Guaranteeing well-connected communities. *Sci. Rep.* **2019**, *9*, 5233. [CrossRef] [PubMed]
54. Wolock, S.L.; Lopez, R.; Klein, A.M. Scrublet: Computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* **2019**, *8*, 281–291.e9. [CrossRef] [PubMed]
55. Korsunsky, I.; Millard, N.; Fan, J.; Slowikowski, K.; Zhang, F.; Wei, K.; Baglaenko, Y.; Brenner, M.; Loh, P.R.; Raychaudhuri, S. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **2019**, *16*, 1289–1296. [CrossRef] [PubMed]
56. 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. RNA sequencing data integration reveals an miRNA interactome of osteoarthritis cartilage. *Ann. Rheum. Dis.* **2019**, *78*, 270–277. [CrossRef] [PubMed]
57. Duan, X.; Cai, L.; Pham, C.T.N.; Abu-Amer, Y.; Pan, H.; Brophy, R.H.; Wickline, S.A.; Rai, M.F. Amelioration of posttraumatic osteoarthritis in mice using intraarticular silencing of periostin via nanoparticle-based small interfering RNA. *Arthritis Rheumatol.* **2021**, *73*, 2249–2260. [CrossRef] [PubMed]
58. Korsunsky, I.; Wei, K.; Pohin, M.; Kim, E.Y.; Barone, F.; Major, T.; Taylor, E.; Ravindran, R.; Kemble, S.; Watts, G.F.M.; et al. Cross-tissue, single-cell stromal atlas identifies shared pathological fibroblast phenotypes in four chronic inflammatory diseases. *Med* **2022**, *3*, 481–518.e14. [CrossRef] [PubMed]
59. Vasheghani, F.; Zhang, Y.; Li, Y.H.; Blati, M.; Fahmi, H.; Lussier, B.; Roughley, P.; Lagares, D.; Endisha, H.; Saffar, B.; et al. PPARγ deficiency results in severe, accelerated osteoarthritis associated with aberrant mTOR signalling in the articular cartilage. *Ann. Rheum. Dis.* **2015**, *74*, 569–578. [CrossRef] [PubMed]
60. Dudek, M.; Gossan, N.; Yang, N.; Im, H.J.; Ruckshanthi, J.P.; Yoshitane, H.; Li, X.; Jin, D.; Wang, P.; Boudiffa, M.; et al. The chondrocyte clock gene Bmal1 controls cartilage homeostasis and integrity. *J. Clin. Investig.* **2016**, *126*, 365–376. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.