Osteoarthritis and Cartilage 26 (2018) 1531–1538

Osteoarthritis and Cartilage

Identification of transcription factors responsible for dysregulated networks in human osteoarthritis cartilage by global gene expression analysis

K.M. Fisch † [a], R. Gamini ‡ [a], O. Alvarez-Garcia ‡ [a], R. Akagi ‡ §, M. Saito ‡ §, Y. Muramatsu ‡ §, T. Sasho §, J.A. Koziol ‡, A.I. Su ‡, M.K. Lotz ‡ [*]

† *Center for Computational Biology and Bioinformatics, Department of Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, USA*
‡ *Department of Molecular Medicine, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA, USA*
§ *Department of Orthopaedic Surgery, Chiba University Hospital 1-8-1 Inohana, Chuo-ku, Chiba, Japan*

ARTICLE INFO

Article history:
Received 5 February 2018
Accepted 13 July 2018

Keywords:
Cartilage
Gene expression
Transcription factors

SUMMARY

Objective: Osteoarthritis (OA) is the most prevalent joint disease. As disease-modifying therapies are not available, novel therapeutic targets need to be discovered and prioritized for their importance in mediating the abnormal phenotype of cells in OA-affected joints. Here, we generated a genome-wide molecular profile of OA to elucidate regulatory mechanisms of OA pathogenesis and to identify possible therapeutic targets using integrative analysis of mRNA-sequencing data obtained from human knee cartilage.

Design: RNA-sequencing (RNA-seq) was performed on 18 normal and 20 OA human knee cartilage tissues. RNA-seq datasets were analysed to identify genes, pathways and regulatory networks that were dysregulated in OA.

Results: RNA-seq data analysis revealed 1332 differentially expressed (DE) genes between OA and non-OA samples, including known and novel transcription factors (TFs). Pathway analysis identified 15 significantly perturbed pathways in OA with ECM-related, PI3K-Akt, HIF-1, FoxO and circadian rhythm pathways being the most significantly dysregulated. We selected DE TFs that are enriched for regulating DE genes in OA and prioritized these TFs by creating a cartilage-specific interaction subnetwork. This analysis revealed eight TFs, including JUN, Early growth response (EGR)1, JUND, FOSL2, MYC, KLF4, RELA, and FOS that both target large numbers of dysregulated genes in OA and are themselves suppressed in OA.

Conclusions: We identified a novel subnetwork of dysregulated TFs that represent new mediators of abnormal gene expression and promising therapeutic targets in OA.

Introduction

Osteoarthritis (OA) is the most prevalent joint disease and a leading cause of disability in the elderly[1]. In established disease, all joint tissues are affected[2] but the articular cartilage appears to be the most vulnerable to traumatic injury and aging-related changes that initiate the disease process[3,4].

There have been several important recent advances, such as genome wide association studies (GWAS), candidate gene and global gene expression analyses that have led to an improved understanding of OA pathogenesis[5,6]. However, the small number of candidate genes identified by GWAS had relatively low Odds ratio (OR)s[5,6]. Genome wide analyses of expression of RNAs offer potential for discovering new mechanisms and therapeutic targets. Only a limited number of genome wide expression analyses have been performed on human OA cartilage[7–14]. A small number of studies were performed in OA mouse models[15–19], comparing rat OA vs normal chondrocytes[20], or articular cartilage from a porcine model of OA[21].

Here we have performed RNA-seq to identify differentially expressed (DE) genes between normal and OA articular cartilage.

* Address correspondence and reprint requests to: M.K. Lotz, Department of Molecular Medicine, MEM-161, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA.
E-mail address: mlotz@scripps.edu (M.K. Lotz).
[a] These authors contributed equally to this work.

https://doi.org/10.1016/j.joca.2018.07.012
1063-4584/

We created a molecular profile of the OA transcriptome and performed functional enrichment analyses to elucidate perturbed molecular functions and pathways in OA. We identified DE transcription factors (TFs) through transcription factor binding site analysis and prioritized these using network analysis. Finally, we developed a high-resolution molecular profile of OA and elucidated a novel regulatory network of OA pathogenesis.

Methods

Cartilage donors

Normal human knee cartilage tissues were procured by tissue banks (approved by Scripps Institutional Review Board) from five female and 13 male (age 18–61, mean 38) without history of joint disease or trauma and processed within 24–48 h post mortem. Full thickness cartilage was harvested for RNA isolation from identical locations on the weightbearing regions on medial and lateral femoral condyles, and adjacent tissue sections were harvested for histology to verify the cartilage integrity. OA-affected cartilage was harvested from the tissue removed during knee replacement surgery from 12 female and eight male donors (age 52–82, mean 66). Body mass indices between the normal (BMI = 32.4 ± 8.0) and OA (BMI = 30.7 ± 8.1) were not significantly ($P = 0.506$) different.

Tissue processing, RNA and DNA isolation

Cartilage was stored at −20°C in Allprotect Tissue Reagent (Qiagen, Valencia, CA) immediately after harvest until RNA extraction. For RNA isolation, a minimum of 150 mg of cartilage (dry weight) was pulverized using a 6770 Freezer/Mill Cryogenic Grinder (SPEX SamplePrep, Metuchen, NJ), and homogenized in Qiazol Lysis Reagent (Qiagen, Valencia, CA) at a concentration of 25 mg tissue sample per 700 μl Qiazol. To remove proteins and cellular debris, a initial phenol-chloroform extraction was performed. Briefly, samples were mixed with 0.2 volumes of chloroform, incubated for 5 min in ice, and centrifuged a t 14,000 rpm for 15 min at 4°C. The aqueous phase was collected, mixed with one volume of Qiazol and incubated for 30 min in ice. Then, samples were mixed with one volume of 100% ethanol, loaded into a mRNeasy Mini kit column (Qiagen) and digested on-column with DNAse following manufacturer instructions. RNA was eluted in 15 μl of RNase-free water. RNA purity was assessed using NanoDrop (ND-1000, Thermo Scientific, Wilmington, differentially expressed (DE) genes) and RNA integrity number (RIN) was calculated using a 2100 Bioanalyzer (Agilent, Santa Clara, CA). Average RIN numbers were 6.08 ± 0.95.

Library preparation and sequencing

RNA samples from 18 normal and 20 OA cartilage donors were sequenced using 150 ng of total RNA as input. Sequencing mRNA libraries were prepared using the Encore Complete RNA-Seq DR Multiplex System 1–8 and 9–16 (NuGen, San Carlos, CA) with 16 unique indexed adapters (L2V6DR-BC2-L2V6DR-BC16). Two lanes of an Illumina HiSeq 2000 instrument were used to generate a total of 8–30 million 100bp reads.

The Illumina Genome Analyzer Pipeline Software (Casava v1.8.2) was used to convert the original image data generated by the sequencing machine into sequence data via base calling in order to generate fastq files and to demultiplex the samples. We performed a per base sequence quality check using the software FastQC (v0.10.1) (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) prior to read mapping. Raw RNAseq reads were aligned to the human genome (hg19) using the STAR aligner[22]. The number of reads sequenced per sample ranged from 19 to 24 million reads, which should be sufficient for gene level quantification, but only 2–12 million reads per sample mapped to protein coding genes. To account for this issue, we applied high stringency the filtering of lowly expressed genes (log counts per million (CPM) > 3) so that only the differential expression analysis included only genes that were expressed in high enough abundances to be confident in their relative gene expression values.

Global gene expression profiling (read mapping, quantification and differential expression)

HTSeq was used to count the number of reads unambiguously overlapping each gene, where each gene was considered to be the union of its exons[23] with UCSC RefSeq hg19 annotation (Release 57). Sample normalization factors were computed using the EdgeR TMM method[24]. Unless otherwise stated, all CPM and $\log_2$CPM values were computed using these normalization factors. Only genes with counts with $\log_2$CPM greater than 3.0 in one or more samples were considered expressed, resulting in a gene list of 13,102 informative transcripts. We converted the gene symbols to Entrez IDs using the annotations in the mygene.info package[25]. Of the 13,102 genes, 635 gene symbols did not have an associated Entrez ID and were discarded from this study. The resulting genes with Entrez IDs correspond to the set of 'background or detected genes' consisting of 12,463 genes. Differential expression analysis was performed with limma-voom[26] with the design ~0 + condition. Genes with an adjusted *P*-value of <0.05 (based on the moderated t-statistic using the Benjamini-Hochberg (BH) method to control the false discovery rate[27]) and a |log2FC| >1 were considered significantly DE.

Gene ontology classification of DE genes and pathway analysis

Assignment of functional categories was based on the Gene Ontology (GO) categories 'Biological process', 'Molecular function' and 'Cellular component'. Enrichment analysis of GO categories was performed in R (version 3.2.1; http://www.r-project.org) using the 'weight01' method from the Bioconductor topGO (v. 1.5.1) package (32). Node size was set to 10, and Fisher's exact test was used for assessing GO term significance. Adjusted *P*-values were computed from the *P*-values using the BH method[27]. Overrepresentation of functional categories was calculated for DE genes as compared with the 12,463 'background' genes, and significant GO terms were identified as those having BH adjusted *P*-value <0.05.

Overrepresentation analysis in WebGestalt v 2017[28] was used to identify significantly enriched KEGG[29] pathways using DE genes as input and setting the 'background genes' to all 12,463 genes expressed in our RNA-seq dataset.

Transcription factor enrichment analysis

TFs were identified by comparing the 12,463 cartilage expressed genes with all genes classified under the GO term 'transcription factor activity, sequence-specific DNA binding' (GO:0003700)[30]. To identify which TFs target DE genes, we used the oPOSSUM single site analysis[31] tool to perform a Fisher's test for overrepresentation of 478 JASPAR position weight matrixes (PWM) profiles in DE gene promoters (defined as 1 kb upstream of the starting transcription site). We used a cut-off of 0.4 and matrix score threshold of 85%. Promoters were defined as the 1 kb region upstream of TSS. For enrichment analysis, a likelihood ratio given by the OR was computed such that $\log_2$ (Odds Ratio) = $\log_2$ and Fisher's exact test was used to test for a significant association between DE status and

TF binding. Enriched TFs having and adjusted *P*-value <0.1 using the BH-method were considered significant.

Network analysis

We performed a network analysis using the HumanBase tissue-specific cartilage gene interaction network[25,26]. The network that we generated with the cartilage data was based on several data types that constituted the underlying network, including experimentally produced protein–protein interactions (http://hb.flatironinstitute.org/data), and the interaction confidence is the edge weight assigned from the algorithm used to create this compendium network. We seeded the network with TFs that were DE and enriched for binding sites in the promoter region of DE genes. We restricted the data types of the network edges to interaction, TF binding, gene GSEA microRNA targets and GSEA perturbations. We set the maximum number of genes in the network to 50 genes, prioritized by uniqueness to the query genes with a minimum interaction confidence of 0.17, a setting at which all seed genes included at least one edge in the network. This resulted in 64 genes in the subnetwork, 62 of which were considered expressed in this study, so downstream analyses were conducted on the 62 node network. Overrepresentation analysis in WebGestalt v 2017[28] was used to identify significantly enriched KEGG pathways in the 62 gene subnetwork, setting the background gene list to those 12,463 tested in the differential expression analysis. The edges in the network were filtered to high confidence edges (weight > 0.35) and the degree of each gene was calculated as the number of connections to other nodes. Cytoscape v 3.6.0[32] was used to visualize the network and annotate genes within enriched pathways in the network.

Data access

The data have been deposited into Gene Expression Omnibus (GEO), accession number GSE114007.

Results

Outline of data analysis

The objective of this study was to use **differentially expressed genes (DE)** OA vs normal cartilage ($n = 1332$) as obtained from RNA-seq for the general purpose of identifying dysregulated genes and pathways (Fig. 1). The subsequent analyses focused on **TFs**, to prioritize them by their importance as regulators of the abnormal transcriptome in OA and by their potential as OA therapeutic targets. We identified 14 TFs that were both DE ($n = 93$) and significantly enriched for binding to the promoters of DE genes ($n = 44$). Network analysis was performed by seeding a tissue-specific cartilage gene interaction network with the 14 TFs to further prioritize them for their importance in contributing to the abnormal OA transcriptome.

Transcriptome landscape of OA reveals dysregulated extracellular matrix metabolism, cell proliferation and differentiation

RNA-seq analysis was performed on human knee articular cartilage isolated from normal ($n = 18$) and OA ($n = 20$) donors. Thus, there were 38 unique subjects (statistically independent), without repeated measures. In the 38 samples, 13,102 transcripts were confidently expressed in at least one sample. For the differential expression analysis and downstream analyses, we used 12,463 of the 13,102 transcripts mapping to unique Entrez IDs. A multidimensional scaling (MDS) plot reveals distinct clustering of OA and normal samples [Fig. 2(A)].

A total of 1332 genes (Supplemental Table 1) were DE between OA and normal human cartilage (adjusted P value < 0.05 and $|\log2FoldChange| > 1$). This set of 1332 DE genes was used in all subsequent analyses.

Unsupervised hierarchical analysis of the significantly DE genes stratifies normal and OA samples into two clusters, as expected [Fig. 2(B)]. Among all DE genes, 630 genes were upregulated and 702 downregulated in OA with a range of differential expression between −4.8 and 6.8 log2FC [Fig. 2(C)]. The genes with the highest fold change in OA cartilage include POSTN, COL1A1, and TNFSF15, and the genes with the lowest fold change were ADM, DDIT4, and DUSP2 [Fig. 2(C)]. Several genes known to be involved in articular cartilage homeostasis and OA pathology were also DE[33,34]. Extracellular proteases from the ADAMTS and MMP families that included ADAMTS2, ADAMTS5, ADAMTS7, ADAMTS14, MMP2, MMP11, MMP13, and MMP19 were upregulated in OA cartilage. In addition, sixteen genes encoding collagens were also upregulated in OA samples and included COL2A1 and COL10A1. On the other hand, normal cartilage samples had significantly higher expression of genes involved in resistance to cellular stress such as DDIT4, SESN2, GADD45A, GADD45B, and GPX3.

To assess the overarching function of genes DE between normal and OA cartilage, gene set overrepresentation analyses were performed using **GO** and KEGG databases (Supplemental Table 2). GO term analysis indicated an enrichment for DE genes encoding proteins located in the extracellular space and the cell membrane [Fig. 2(D)]. DE genes were also enriched for processes involved in ECM metabolism, cell–cell interaction and intracellular signaling pathways such as G-protein coupled receptor and cAMP signaling pathways. Pathway analysis revealed 15 significantly dysregulated pathways in OA [Fig. 2(E)]. The top pathways were protein digestion and absorption, neuroactive ligand–receptor interaction and complement and coagulation cascades. Other dysregulated pathways previously linked to OA were PI3K-Akt, FoxO, HIF-1 and Circadian rhythm pathways[35,36].

Suppressed transcription factors and their impact on OA gene expression patterns

To identify TFs that may contribute to the abnormal transcriptome landscape in OA cartilage, an integrative analysis was

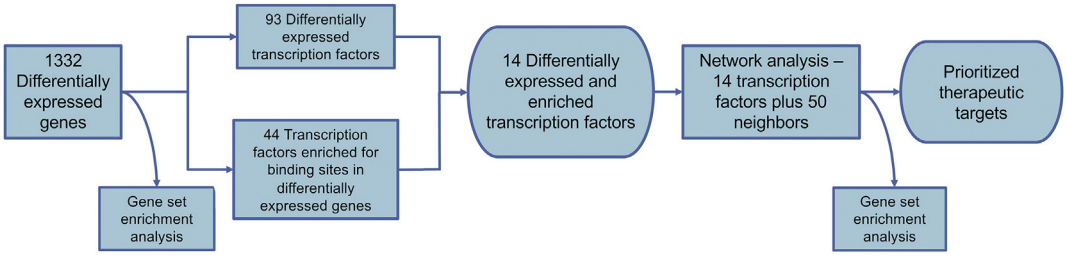

Fig. 1. Overview of the data analysis.

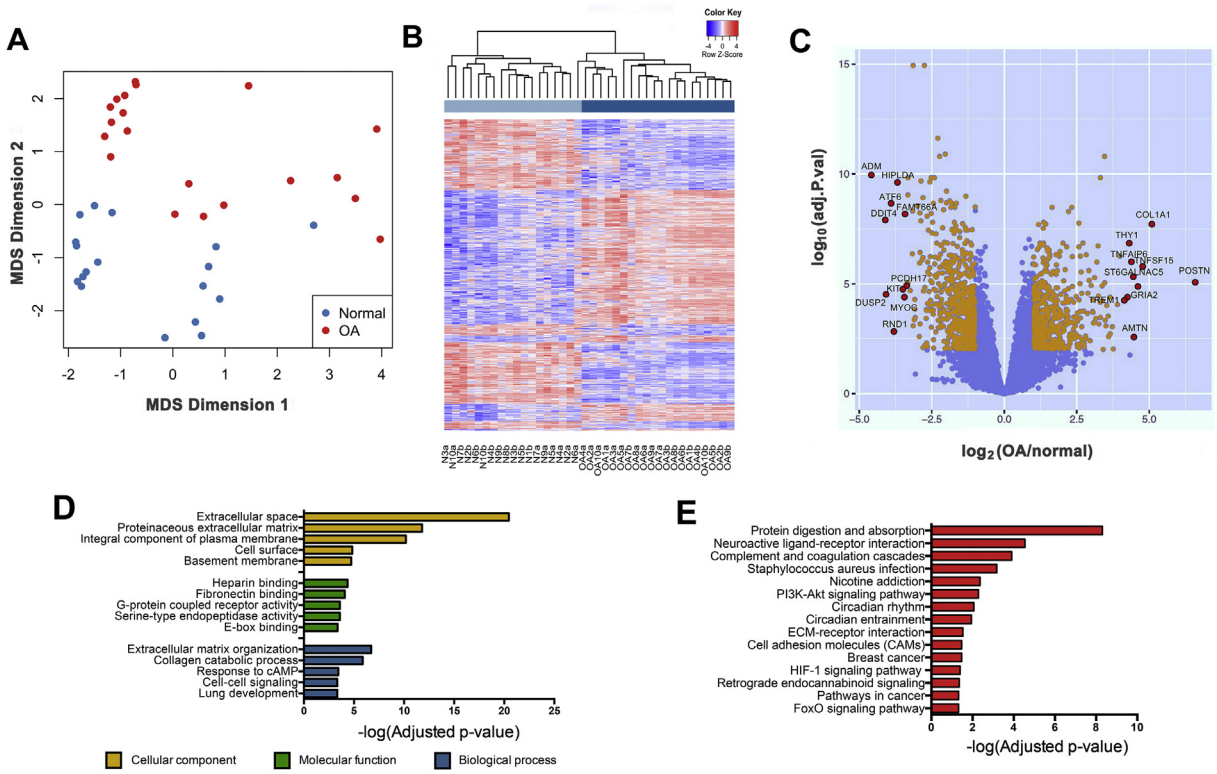

Fig. 2. Transcriptomic landscape of normal and OA knee articular cartilage. A) Multidimensional scaling (MDS) plot of gene expression (lcpm) in normal (blue) and OA articular cartilage samples (red) reveals strong clustering of samples by phenotype. B) Unsupervised hierarchical clustering of normal and OA articular cartilage samples based on the expression levels of the top 1000 differentially expressed (DE) genes ranked by adjusted *P*-value. C) Volcano plot representation of gene expression analysis in normal and OA articular cartilage samples highlighting the most DE genes. **D**) Bar plot representing top results of gene ontology (GO) enrichment analysis of DE genes between normal and OA articular cartilage samples. E) Bar plot representing all 15 significantly enriched KEGG pathways in DE genes between normal and OA articular cartilage samples.

performed using two complementary approaches. First, the 12,463 cartilage expressed genes were compared with all genes classified under the GO term 'transcription factor activity, sequence-specific DNA binding' (GO:0003700)[30] to identify TFs expressed in cartilage, which resulted in 865 genes (Supplemental Table 3). Of the 865 TFs, 93 were DE, with 76 being downregulated and 17 upregulated in OA cartilage (Supplemental Table 4). Among this large number of TFs with reduced expression in OA was the master regulator of chondrogenesis SOX9[37]. In addition, seven members of the Krüppel-like family of transcription factors (KLF), all three members of the Nur (nuclear receptor subfamily 4) family of orphan nuclear receptors, and two members of the Early growth response (EGR) family had significantly lower expression in OA. Comparison with previous studies that analyzed the transcriptome of normal and OA human cartilage[8,14] revealed substantial overlap between DE TFs (48/93 vs Soul *et al.*[14], and 23/93 vs Karlsson *et al.*[8]) and high concordance in the direction of the changes (44/48 vs Soul *et al.*[14], and 17/23 vs Karlsson *et al.*[8]).

As a complementary approach, the 1332 DE genes were analyzed using the oPOSSUM single site analysis[31] tool and the JASPAR[38] database to identify TFs enriched for binding sites in the promoter regions of these DE genes. This analysis identified 44 TFs that were significantly enriched (Supplementary Table 5) with JDP2, NFYA, RELA, SP1 and KLF4 exhibiting the highest enrichment. Integration of both analyses delineated a subset of 14 TFs that were DE and also enriched for binding sites in the promoter regions of DE genes [Fig. 3(A)]. Surprisingly, all 14 TFs had significantly lower expression in OA than in normal cartilage [Fig. 3(B)]. The DE TFs that targeted the largest number of DE genes were KLF4, KLF5, TBX4, TBX5 and EGR1 [Fig. 3(C)].

Network analysis reveals master regulators and key molecular mechanisms of OA pathogenesis

To identify subsets of genes with functional enrichment for OA, we performed a network analysis using the HumanBase tissue-specific cartilage gene interaction network, which is a genome-scale protein function and interaction map of human tissues derived from integrating data sets from thousands of experiments[39,40]. We seeded the network with the 14 TFs that were DE and enriched for binding sites in the promoter region of DE genes. We included 50 nearest neighbors of these genes prioritized by uniqueness to the query genes with a minimum interaction confidence of 0.17, a setting at which all seed genes included at least one edge in the network. This resulted in a subnetwork of 64 genes, 44 of which were DE in our RNA-seq dataset and 62 of which were considered expressed in cartilage [Fig. 4(A); Supplemental Table 6[. Most of the genes in the subnetwork were downregulated in OA and only three were upregulated (SERPINE1, IL11, and ARL4C), indicating that this subnetwork is suppressed in OA cartilage.

To determine which specific processes are regulated by this subnetwork, we performed overrepresentation analysis of KEGG pathways and identified 40 significantly enriched pathways including HIF-1, NF-κB, TGF-β, FoxO and Wnt signaling pathways (Supplemental Table 7). The subnetwork analysis expanded the TF enrichment analysis to reveal additional dysregulated pathways involved in OA pathogenesis. The pathways that were enriched in this subnetwork analysis and also in the prior pathway analysis based on the 1332 DE genes (Fig. 2) included HIF-1 signaling, pathways in cancer and FoxO signaling [Fig. 4(B)]. This analysis highlights that these pathways are both focal to OA pathogenesis

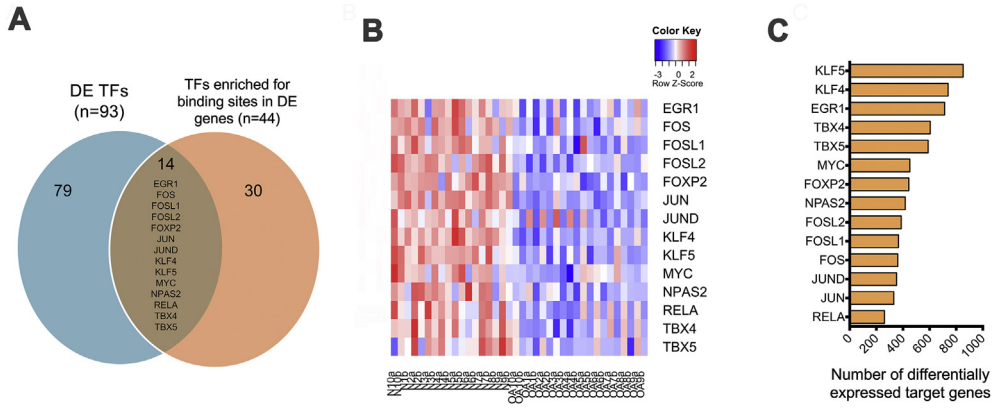

Fig. 3. Transcription factor analysis. A) Overlap between DE transcription factors (TFs) and TFs enriched for binding sites in the promoter region of DE genes in normal and OA articular cartilage results in 14 diferentially expressed genes (DE) enriched TFs. B) Heatmap of normal and OA articular cartilage samples based on the expression of 14 DE enriched TFs. C) Top five DE TFs with highest percentage of target genes that are DE in normal and OA articular cartilage.

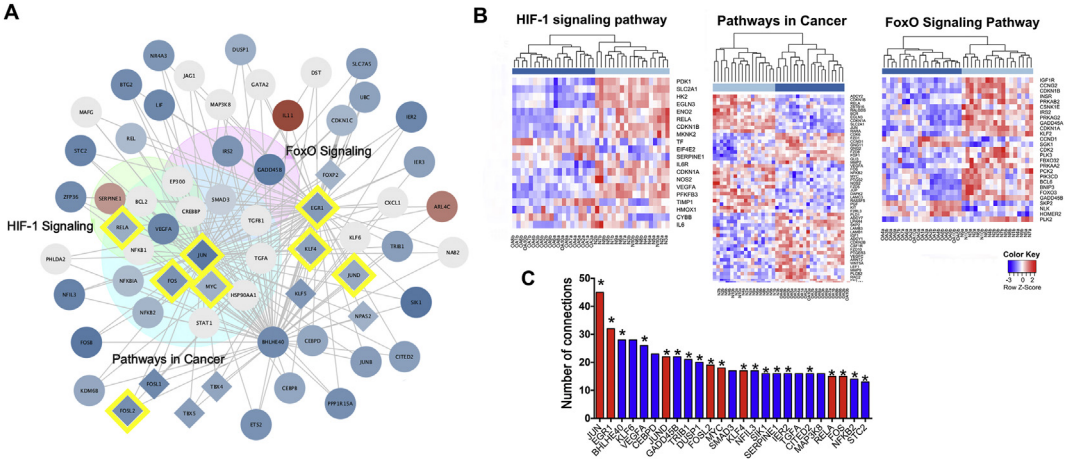

Fig. 4. Network-based prioritization of DE genes in OA. A) Network diagram of the 64 gene subnetwork from the HumanBase cartilage-specific network seeded with the 14 DE enriched TFs. Node shape represented by diamonds represent the 14 DE enriched TFs, node color indicates significantly up- (red) or down- (blue) regulated in OA vs Normal articular cartilage and node size corresponds to the degree (larger nodes have more connections to other nodes). Edges not drawn between nodes to aid in visualization of the membership of each gene within significantly enriched pathways, labeled with text and highlighted by background color. B) Unsupervised hierarchical clustering of normal and OA articular cartilage samples based on the expression levels of DE genes that belong to the HIF-1 signaling pathway, pathways cancer, and FoxO signaling pathway. C) Ranking of genes in the network based on their degree. DE genes are marked with an asterisk. TFs that are DE and enriched for binding to DE genes are highlighted in red.

and that these prioritized TFs play a central role in their dysregulation. Interestingly, most of the genes from the HIF-1 and FoxO signaling pathways were downregulated in OA cartilage [Fig. 4(B)].

To further prioritize DE genes between OA and normal cartilage and identify targets for further study, the edges in the network were filtered to high confidence edges (weight > 0.35) and the degree of each gene was calculated as the number of connections to other nodes [Fig. 4(C)]. This resulted in a ranked list where JUN and EGR1 were the genes with the highest number of connections (45 and 32 connections, respectively). In addition to JUN and EGR1, the top 25 highest degree genes included six TFs what also were DE in OA and enriched for targeting DE genes (JUND, FOSL2, MYC, KLF4, RELA, FOS). These analyses therefore suggest that these eight TFs might constitute a key transcriptional network that could drive the dysregulated transcriptomic landscape of OA cartilage. In particular, KLF4 and EGR1 are the TFs that target a larger number of DE genes [Fig. 3(C)] and may thus represent high priority targets for therapeutic interventions.

Discussion

This study described the transcriptomic landscape of normal and OA cartilage. Unbiased analyses of these datasets were performed to identify critical regulators of abnormal gene expression changes in OA with the objective to prioritize therapeutic targets. We identified a subnetwork of dysregulated TF that may largely contribute to the abnormal gene expression pattern in OA.

The data presented here confirms and expands previous findings from genome-wide gene expression analyses of normal and OA cartilage using RNA-seq[13,14,21] and DNA arrays[7–10]. A common finding of these studies in human cartilage and also in animal models of experimental OA[15–19] is a differential expression of genes involved in ECM metabolism. Our data analysis showed increased expression of ECM components (COL10A1, COL13A1, COL15A1, COL1A2) and proteases (MMP13, ADAMTS5) which is suggestive of active remodeling of ECM during OA pathogenesis. In addition, genes involved in activation of the complement cascade were also dysregulated in our study in agreement with previous reports[13,14], thus strengthening the notion that complement factors may play a significant role in OA pathogenesis. On the other hand, our data analysis did not reveal abnormal expression of pathways previously reported to be involved in OA pathophysiology such as TGF-β/WNT signaling, inflammation or angiogenesis[14,18].

Functional enrichment analyses showed that DE genes were also involved in pathways recently linked to the maintenance of articular cartilage homeostasis, namely PI3K-Akt, FoxO, HIF-1 and

circadian rhythm signaling pathways. PI3K-Akt signaling pathway is involved in cartilage ECM degradation, modulation of autophagy and cell death[41]. Involvement of the PI3K-Akt signaling pathway in OA pathogenesis has been observed in hip OA[10] and also in knee OA cartilage when comparing unaffected and affected regions of the joint[13]. FoxO TFs are negative regulators of the PI3K-Akt pathway that have been recently reported to be decreased in OA and aged cartilage and to promote resistance to cellular stress in chondrocytes[42,43]. HIF-1α is the central regulator of cellular responses to hypoxia and has been shown to promote chondrocyte differentiation and survival[35]. Moreover, transcriptomic differences in genes involved on cellular responses to hypoxia have been reported in hip[10] and knee[13] OA cartilage. Lastly, several genes of the circadian rhythm pathway have been shown to be dysregulated in OA cartilage (21, 27) and to regulate pathways in OA pathogenesis, including the TGFβ pathway[36,44]. Collectively, these observations suggest that dysregulation of these pathways would compromise tissue homeostasis and promote cartilage degeneration.

Whereas general analysis of gene expression changes offers broad insight into dysregulated processes that might contribute to OA pathobiology, it is challenging to select specific candidates with greater potential for therapeutic targeting. For this reason, a main focus of the present study was to identify TFs as key regulators of DE genes. Starting with the complete list of TFs that are expressed in human cartilage (Supplemental Table 3), we identified 93 TF that were DE in OA (Supplemental Table 4) and interestingly, the majority were downregulated in OA cartilage. We also compared the 93 DE TFs with previous genome-wide transcriptomic studies using normal and OA human cartilage[8,14]. We found substantial overlap between DE TFs and we identified 13 TFs that were DE in all three studies, with GLI3, RCAN1, and SOX11 being upregulated and ATOH8, BCL6, CEBPB, CITED2, FOSL2, HIF3A, HMGB2, KLF15, SOX13 and ZBTB16 being downregulated in OA cartilage. Since most of these TF have not been previously investigated in articular cartilage, they are attractive targets. Further studies are needed to validate their function in chondrocyte homeostasis and OA pathogenesis.

Our integrative data analysis prioritized TFs based on their differential expression, enrichment for binding sites in DE genes, and number of connections in a cartilage-specific subnetwork (Supplemental Fig. 1). The resulting eight transcription factors (JUN, EGR1, JUND, FOSL2, MYC, KLF4, RELA and FOS) thus represent high priority candidates for therapeutic intervention. To date, there is no evidence about the role of KLF4 in articular cartilage. JUN, JUND, FOS and FOSL2 encode different subunits of the AP-1[45]. The function of AP-1 in chondrocytes remains to be more clearly elucidated. There is evidence that IL-1 induces expression of MMP13 proteases via activation of JUN/FOS heterodimers[46–48]. On the other hand, AP-1 heterodimers containing FOS have been shown to be required for chondrocyte differentiation[49]. RELA/p65 is a key subunit of the NFκB protein complex that is involved in a wide range of biological processes[50]. Whereas activation of the NFkB pathway by proinflammatory cytokines is considered a key driver of cartilage destruction in OA[51], a recent report showed that mice with homozygous deletion of RELA exhibit accelerated cartilage degeneration[52]. Only a small number of studies investigated the role of EGR1[53] in chondrocytes[54–56]. EGR1 regulates terminal differentiation of chondrocytes and also mediates the catabolic response to proinflammatory cytokines. Overall, the limited available information about the functions of these eight TFs in articular cartilage suggests that they are involved in chondrocyte differentiation, metabolism, stress responses and cell survival.

It is important to note that, in addition to the eight TFs mentioned above, our subnetwork analysis also highlighted other DE genes with high numbers of connections that may contribute to the dysregulated gene expression in OA cartilage. For instance, the transcription factor BHLHE40 (also known as DEC1) is a component of the circadian molecular clock[57,58] that also acts a chondrogenic factor[59,60]. On the other hand, SERPINE1 is markedly upregulated in OA cartilage. It encodes for plasminogen activation inhibitor 1 (PAI-1), a serine protease inhibitor that inhibits tissue plasminogen activator and urokinase, and that has additional actions that may be dependent on or independent of its protease inhibitory effects[61]. Interestingly, increased SERPINE1 expression was found in OA affected areas[9] and in a subgroup of OA patients[14].

The main objective of the present study was to analyze OA-related changes in gene expression in cartilage. As control group we used joints that had no macroscopic or microscopic evidence of cartilage damage. While this represents a clear distinction between normal and OA, the limitation is that the normal donors were much younger than the OA patients who had knee arthroplasty. It is thus possible that some of the differences in gene expression are manifestations of cartilage aging. The sample collection for the normal cartilage donors was standardized and tissue for RNA isolation obtained from the weightbearing regions on medial and lateral femoral condyles. In human OA joints this is difficult since the cartilage in the most weight bearing areas is often completely eroded. To examine the effect of mechanical load additional studies are needed where cartilage from areas that different in loading are analyzed.

In conclusion, the present study is the first to analyze TFs on a genome-wide scale for their role in cartilage homeostasis and potential in OA pathogenesis. We used a novel approach to prioritize the TFs for their potential as therapeutic targets for OA. The involvement of these TFs in cartilage homeostasis and their value as therapeutic targets need to be validated in studies on their function *in vitro* and *in vivo*.

Author contributions

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Lotz had full access to all data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design: Fisch, Gamini, Alvarez-Garcia, Sasho, Su, Lotz.

Acquisition of data: Saito, Akagi, Muramatsu, Alvarez-Garcia.

Analysis and interpretation of data: Fisch, Gamini, Alvarez-Garcia, Sasho, Koziol, Su, Lotz.

Competing financial interests

Martin K. Lotz and Andrew Su receive grant support from NIH.

Role of the funding source

This study was supported by NIH grants AG007996, AG049617 and TR001114.

Acknowledgments

We appreciate the technical support from Stuart F. Duffy, Josan Chung, and Merissa Olmer.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.joca.2018.07.012.

References

1. Dillon CF, Rasch EK, Gu Q, Hirsch R. Prevalence of knee osteoarthritis in the United States: arthritis data from the Third National Health and Nutrition Examination Survey 1991-94. J Rheumatol 2006;33:2271–9.
2. Loeser RF, Goldring SR, Scanzello CR, Goldring MB. Osteoarthritis: a disease of the joint as an organ. Arthritis Rheum 2012;64:1697–707.
3. Lotz M, Loeser RF. Effects of aging on articular cartilage homeostasis. Bone 2012;51:241–8.
4. Goldring MB. Articular cartilage degradation in osteoarthritis. HSS J 2012;8:7–9.
5. Rodriguez-Fontenla C, Calaza M, Evangelou E, Valdes AM, Arden N, Blanco FJ, *et al.* Assessment of osteoarthritis candidate genes in a meta-analysis of nine genome-wide association studies. Arthritis Rheum 2014;66:940–9.
6. Reynard LN. Analysis of genetics and DNA methylation in osteoarthritis: What have we learnt about the disease? Semin Cell Dev Biol 2017;62:57–66.
7. Aigner T, Fundel K, Saas J, Gebhard PM, Haag J, Weiss T, *et al.* Large-scale gene expression profiling reveals major pathogenetic pathways of cartilage degeneration in osteoarthritis. Arthritis Rheum 2006;54:3533–44.
8. Karlsson C, Dehne T, Lindahl A, Brittberg M, Pruss A, Sittinger M, *et al.* Genome-wide expression profiling reveals new candidate genes associated with osteoarthritis. Osteoarthritis Cartilage 2010;18:581–92.
9. Ramos YF, den Hollander W, Bovee JV, Bomer N, van der Breggen R, Lakenberg N, *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.
10. Xu Y, Barter MJ, Swan DC, Rankin KS, Rowan AD, Santibanez-Koref M, *et al.* Identification of the pathogenic pathways in osteoarthritic hip cartilage: commonality and discord between hip and knee OA. Osteoarthritis Cartilage 2012;20:1029–38.
11. Dunn SL, Soul J, Anand S, Schwartz JM, Boot-Handford RP, Hardingham TE. Gene expression changes in damaged osteoarthritic cartilage identify a signature of non-chondrogenic and mechanical responses. Osteoarthritis Cartilage 2016;24:1431–40.
12. Snelling S, Rout R, Davidson R, Clark I, Carr A, Hulley PA, *et al.* A gene expression study of normal and damaged cartilage in anteromedial gonarthrosis, a phenotype of osteoarthritis. Osteoarthritis Cartilage 2014;22:334–43.
13. Steinberg J, Ritchie GRS, Roumeliotis TI, Jayasuriya RL, Clark MJ, Brooks RA, *et al.* Integrative epigenomics, transcriptomics and proteomics of patient chondrocytes reveal genes and pathways involved in osteoarthritis. Sci Rep 2017;7:8935.
14. Soul J, Dunn SL, Anand S, Serracino-Inglott F, Schwartz JM, Boot-Handford RP, *et al.* Stratification of knee osteoarthritis: two major patient subgroups identified by genome-wide expression analysis of articular cartilage. Ann Rheum Dis 2017;77:423, https://doi.org/10.1136/annrheumdis-2017-212603. Published Online First: 22 Dec 2017.
15. Burleigh A, Chanalaris A, Gardiner MD, Driscoll C, Boruc O, Saklatvala J, *et al.* Joint immobilization prevents murine osteoarthritis and reveals the highly mechanosensitive nature of protease expression in vivo. Arthritis Rheum 2012;64:2278–88.
16. Loeser RF, Olex AL, McNulty MA, Carlson CS, Callahan MF, Ferguson CM, *et al.* Microarray analysis reveals age-related differences in gene expression during the development of osteoarthritis in mice. Arthritis Rheum 2012;64:705–17.
17. Poulet B, Ulici V, Stone TC, Pead M, Gburcik V, Constantinou E, *et al.* Time-series transcriptional profiling yields new perspectives on susceptibility to murine osteoarthritis. Arthritis Rheum 2012;64:3256–66.
18. Gardiner MD, Vincent TL, Driscoll C, Burleigh A, Bou-Gharios G, Saklatvala J, *et al.* Transcriptional analysis of micro-dissected articular cartilage in post-traumatic murine osteoarthritis. Osteoarthritis Cartilage 2015;23:616–28.
19. Loeser RF, Olex AL, McNulty MA, Carlson CS, Callahan M, Ferguson C, *et al.* Disease progression and phasic changes in gene expression in a mouse model of osteoarthritis. PLoS One 2013;8:e54633.
20. Appleton CT, Pitelka V, Henry J, Beier F. Global analyses of gene expression in early experimental osteoarthritis. Arthritis Rheum 2007;56:1854–68.
21. Sieker JT, Proffen BL, Waller KA, Chin KE, Karamchedu NP, Akelman MR, *et al.* Transcriptional profiling of articular cartilage in a porcine model of early post-traumatic osteoarthritis. J Orthop Res 2018 Jan;36(1):318–29, https://doi.org/10.1002/jor.2364. Epub 2017 Aug 3.
22. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, *et al.* STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.
23. Anders S, Pyl PT, Huber W. HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics 2015;31:166–9.
24. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139–40.
25. Xin J, Mark A, Afrasiabi C, Tsueng G, Juchler M, Gopal N, *et al.* High-performance web services for querying gene and variant annotation. Genome Biol 2016;17:91.
26. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 2014;15:R29.
27. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. Behav Brain Res 2001;125:279–84.
28. Zhang B, Kirov S, Snoddy J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. Nucleic Acids Res 2005;33:W741–8.
29. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28:27–30.
30. Ardlie KG, Deluca DS, Segré AV, Sullivan TJ, Young TR, Gelfand ET. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 2015;348:648–60, https://www.ncbi.nlm.nih.gov/pubmed/25954001.
31. Kwon AT, Arenillas DJ, Worsley Hunt R, Wasserman WW. oPOSSUM-3: advanced analysis of regulatory motif over-representation across genes or ChIP-Seq datasets. G3 (Bethesda) 2012;2:987–1002.
32. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498–504.
33. Sandy JD, Chan DD, Trevino RL, Wimmer MA, Plaas A. Human genome-wide expression analysis reorients the study of inflammatory mediators and biomechanics in osteoarthritis. Osteoarthritis Cartilage 2015;23:1939–45.
34. Rogers EL, Reynard LN, Loughlin J. The role of inflammation-related genes in osteoarthritis. Osteoarthritis Cartilage 2015;23:1933–8.

35. Wu L, Huang X, Li L, Huang H, Xu R, Luyten W. Insights on biology and pathology of HIF-1alpha/-2alpha, TGFbeta/BMP, Wnt/beta-catenin, and NF-kappaB pathways in osteoarthritis. Curr Pharmaceut Des 2012;18:3293–312.
36. Akagi R, Akatsu Y, Fisch KM, Alvarez-Garcia O, Teramura T, Muramatsu Y, *et al*. Dysregulated circadian rhythm pathway in human osteoarthritis: NR1D1 and BMAL1 suppression alters TGF-beta signaling in chondrocytes. Osteoarthritis Cartilage 2017;25:943–51.
37. Bi W, Deng JM, Zhang Z, Behringer RR, de Crombrugghe B. Sox9 is required for cartilage formation. Nat Genet 1999;22:85–9.
38. Sandelin A, Alkema W, Engstrom P, Wasserman WW, Lenhard B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res 2004;32:D91–4.
39. Greene CS, Krishnan A, Wong AK, Ricciotti E, Zelaya RA, Himmelstein DS, *et al*. Understanding multicellular function and disease with human tissue-specific networks. Nat Genet 2015;47:569–76.
40. Krishnan A, Zhang R, Yao V, Theesfeld CL, Wong AK, Tadych A, *et al*. Genome-wide prediction and functional characterization of the genetic basis of autism spectrum disorder. Nat Neurosci 2016;19:1454–62.
41. Beier F, Loeser RF. Biology and pathology of Rho GTPase, PI-3 kinase-Akt, and MAP kinase signaling pathways in chondrocytes. J Cell Biochem 2010;110:573–80.
42. Akasaki Y, Alvarez-Garcia O, Saito M, Carames B, Iwamoto Y, Lotz MK. FoxO transcription factors support oxidative stress resistance in human chondrocytes. Arthritis Rheum 2014;66:3349–58.
43. Akasaki Y, Hasegawa A, Saito M, Asahara H, Iwamoto Y, Lotz MK. Dysregulated FOXO transcription factors in articular cartilage in aging and osteoarthritis. Osteoarthritis Cartilage 2014;22:162–70.
44. Dudek M, Gossan N, Yang N, Im HJ, Ruckshanthi JP, Yoshitane H, *et al*. The chondrocyte clock gene Bmal1 controls cartilage homeostasis and integrity. J Clin Invest 2016;126:365–76.
45. Hess J, Angel P, Schorpp-Kistner M. AP-1 subunits: quarrel and harmony among siblings. J Cell Sci 2004;117:5965–73.
46. Hui A, Min WX, Tang J, Cruz TF. Inhibition of activator protein 1 activity by paclitaxel suppresses interleukin-1-induced collagenase and stromelysin expression by bovine chondrocytes. Arthritis Rheum 1998;41:869–76.
47. Fahmi H, Di Battista JA, Pelletier JP, Mineau F, Ranger P, Martel-Pelletier J. Peroxisome proliferator–activated receptor gamma activators inhibit interleukin-1beta-induced nitric oxide and matrix metalloproteinase 13 production in human chondrocytes. Arthritis Rheum 2001;44:595–607.
48. Schmucker AC, Wright JB, Cole MD, Brinckerhoff CE. Distal interleukin-1beta (IL-1beta) response element of human matrix metalloproteinase-13 (MMP-13) binds activator protein 1 (AP-1) transcription factors and regulates gene expression. J Biol Chem 2012;287:1189–97.
49. Ionescu AM, Schwarz EM, Vinson C, Puzas JE, Rosier R, Reynolds PR, *et al*. PTHrP modulates chondrocyte differentiation through AP-1 and CREB signaling. J Biol Chem 2001;276:11639–47.
50. Nolan GP, Ghosh S, Liou HC, Tempst P, Baltimore D. DNA binding and I kappa B inhibition of the cloned p65 subunit of NF-kappa B, a rel-related polypeptide. Cell 1991;64:961–9.
51. Saito T, Tanaka S. Molecular mechanisms underlying osteoarthritis development: Notch and NF-kappaB. Arthritis Res Ther 2017;19:94.
52. Kobayashi H, Chang SH, Mori D, Itoh S, Hirata M, Hosaka Y, *et al*. Biphasic regulation of chondrocytes by Rela through induction of anti-apoptotic and catabolic target genes. Nat Commun 2016;7:13336.
53. Duclot F, Kabbaj M. The role of early growth response 1 (EGR1) in brain plasticity and neuropsychiatric disorders. Front Behav Neurosci 2017;11:35.
54. Nebbaki SS, El Mansouri FE, Afif H, Kapoor M, Benderdour M, Duval N, *et al*. Egr-1 contributes to IL-1-mediated down-regulation of peroxisome proliferator-activated receptor gamma expression in human osteoarthritic chondrocytes. Arthritis Res Ther 2012;14:R69.
55. Rockel JS, Bernier SM, Leask A. Egr-1 inhibits the expression of extracellular matrix genes in chondrocytes by TNFalpha-induced MEK/ERK signalling. Arthritis Res Ther 2009;11:R8.
56. Tan L, Peng H, Osaki M, Choy BK, Auron PE, Sandell LJ, *et al*. Egr-1 mediates transcriptional repression of COL2A1 promoter activity by interleukin-1beta. J Biol Chem 2003;278:17688–700.
57. Honma S, Kawamoto T, Takagi Y, Fujimoto K, Sato F, Noshiro M, *et al*. Dec1 and Dec2 are regulators of the mammalian molecular clock. Nature 2002;419:841–4.
58. Nakashima A, Kawamoto T, Honda KK, Ueshima T, Noshiro M, Iwata T, *et al*. DEC1 modulates the circadian phase of clock gene expression. Mol Cell Biol 2008;28:4080–92.
59. Shen M, Kawamoto T, Yan W, Nakamasu K, Tamagami M, Koyano Y, *et al*. Molecular characterization of the novel basic helix-loop-helix protein DEC1 expressed in differentiated human embryo chondrocytes. Biochem Biophys Res Commun 1997;236:294–8.
60. Shen M, Yoshida E, Yan W, Kawamoto T, Suardita K, Koyano Y, *et al*. Basic helix-loop-helix protein DEC1 promotes chondrocyte differentiation at the early and terminal stages. J Biol Chem 2002;277:50112–20.
61. Eddy AA. Plasminogen activator inhibitor-1 and the kidney. Am J Physiol Ren Physiol 2002;283:F209–20.