



■ CARTILAGE

Whole-transcriptome sequencing of knee joint cartilage from osteoarthritis patients

**H. Li,
H. H. Yang,
Z. G. Sun,
H. B. Tang,
J. K. Min**

*The First People's
Hospital of Huzhou,
Huzhou, China*

Objectives

The aim of this study was to provide a comprehensive understanding of alterations in messenger RNAs (mRNAs), long noncoding RNAs (lncRNAs), and circular RNAs (circRNAs) in cartilage affected by osteoarthritis (OA).

Methods

The expression profiles of mRNAs, lncRNAs, and circRNAs in OA cartilage were assessed using whole-transcriptome sequencing. Bioinformatics analyses included prediction and reannotation of novel lncRNAs and circRNAs, their classification, and their placement into subgroups. Gene ontology and pathway analysis were performed to identify differentially expressed genes (DEGs), differentially expressed lncRNAs (DELs), and differentially expressed circRNAs (DECs). We focused on the overlap of DEGs and targets of DELs previously identified in seven high-throughput studies. The top ten DELs were verified by quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) in articular chondrocytes, both *in vitro* and *in vivo*.

Results

In total, 739 mRNAs, 1152 lncRNAs, and 42 circRNAs were found to be differentially expressed in OA cartilage tissue. Among these, we identified 18 overlapping DEGs and targets of DELs, and the top ten DELs were screened by expression profile analysis as candidate OA-related genes. *WISP2*, *ATF3*, and *CHI3L1* were significantly increased in both normal *versus* OA tissues and normal *versus* interleukin (IL)-1 β -induced OA-like cell models, while *ADAM12*, *PRELP*, and *ASPN* were shown to be significantly decreased. Among the identified DELs, we observed higher expression of *ENST00000453554* and *MSTRG.99593.3*, and lower expression of *MSTRG.44186.2* and *NONHSAT186094.1* in normal *versus* OA cells and tissues.

Conclusion

This study revealed expression patterns of coding and noncoding RNAs in OA cartilage, which added sets of genes and noncoding RNAs to the list of candidate diagnostic biomarkers and therapeutic agents for OA patients.

Cite this article: *Bone Joint Res* 2019;8:290–303.

Keywords: Osteoarthritis, Cartilage, RNA-Seq, Whole-transcriptome sequencing, Noncoding RNAs, Articular cartilage

■ H. Li, MD, Deputy Chief Physician,
■ H. H. Yang, BMed, Chief Physician,
■ Z. G. Sun, MMed, Resident,
■ H. B. Tang, MMed, Resident,
■ J. K. Min, BMed, Chief Physician, Department of Orthopaedics, The First People's Hospital of Huzhou, The First Affiliated Hospital of Huzhou Teachers College, Huzhou, China.

Correspondence should be sent to J. K. Min; email: jikangmin2018@gmail.com

doi: 10.1302/2046-3758.87.BJR-2018-0297.R1

Bone Joint Res 2019;8:290–303.

Article focus

■ The aim of this study was to provide a comprehensive understanding of alterations in messenger RNAs (mRNAs), long noncoding RNAs (lncRNAs), and circular RNAs (circRNAs) in cartilage affected by osteoarthritis (OA).

Key messages

■ A total of 739 messenger RNAs (mRNAs), 1152 long noncoding RNAs (lncRNAs), and 42 circular RNAs (circRNAs) were

differentially expressed in OA cartilage tissue.

■ *WISP2*, *ATF3*, *CHI3L1*, *ADAM12*, *PRELP*, *ASPN*, *ENST00000453554*, *MSTRG.99593.3*, *MSTRG.44186.2*, and *NONHSAT186094.1* might be involved in the pathology of OA.

Strengths and limitations

■ Many transcripts in OA cartilage were detected by whole-transcriptome sequencing and the results were compared with previous studies.

- The overlapping set of differentially expressed genes and targets of differentially expressed lncRNAs were selected, and their expression profiles were followed *in vivo* and *in vitro*.
- Functional research of selected OA-related genes and lncRNAs, as well as the expression profiles of differentially expressed circRNAs, should be further validated.

Introduction

Osteoarthritis (OA) is a chronic joint disorder, the incidence of which is highest among elderly people. Its clinical symptoms include joint pain and dysfunction, eventually leading to disability. Pathological features of OA include degradation of articular cartilage, subchondral bone sclerosis, and osteophyte formation. During the early OA process, the surfaces of the cartilage are destroyed and depressed until the cartilage is completely exposed. This is followed by the proliferation of chondrocytes along the exposed cartilage; these chondrocytes form a group and secrete a large number of cell growth factors, including various tissue-destroying proteases, which further cause chondrocyte apoptosis and aggravate the pathological changes.¹ As OA is a multifactorial disease, the underlying pathological process is closely associated with alterations of various genes. In addition, various epigenetic modifications, such as DNA methylation and histone post-translational modification, as well as noncoding RNAs (ncRNAs), contribute to the OA process.² In particular, the connection between ncRNAs and OA development is the focus of an increasing number of studies. Noncoding RNAs represent 98% of the whole human genome and include microRNAs, long noncoding RNAs (lncRNAs), and circular RNAs (circRNAs).³

Alterations in the level of lncRNAs in OA cartilage may result in the aberrant expression of target genes, thereby disrupting cartilage homeostasis. Only a dozen lncRNAs out of more than 5000 lncRNAs identified in the human genome have, to date, been proven to be associated with OA development. However, owing to their tissue and cell specificity, stability, and detectability in various body fluids, lncRNAs have found application in personalized diagnosis, therapy, and prognosis.⁴

Circular RNAs, a novel type of ncRNA, have attracted increasing interest owing to their potential utility as reliable disease biomarkers. Circular RNAs are a large class of ncRNAs with covalently closed loop structures without a free 3' or 5' end; they are characterized by their more stable structure and tissue specificity compared with linear RNA. Circular RNAs facilitate OA development and represent promising novel candidate biomarkers of OA.⁵

In this study, whole-transcriptome sequencing was applied to examine the expression profiles of messenger RNA (mRNA), lncRNA, and circRNA in intact *versus* damaged regions of knee joint cartilage from OA patients in order to identify key features underlying transcript

alteration during the OA process. Furthermore, we screened OA-related candidate genes to assess biological function and identify avenues for clinical intervention.

Materials and Methods

Patients/tissue. Nine medial tibial plateaus (three for sequencing and six for polymerase chain reaction (PCR) validation) were obtained from patients with symptomatic OA during total knee arthroplasty. The ethics committee of the First People's Hospital of Huzhou approved the study (#201710312), and written informed consent was obtained from all subjects. All included subjects were men (median age 71 years (interquartile range 65 to 76)) of Chinese nationality and Han ethnicity who had advanced knee OA. Cartilage from damaged and undamaged regions was carefully removed to avoid contamination of bone, synovium, and blood, and was stored in liquid nitrogen until further use.

Whole-transcript sequencing analysis and identification of differentially expressed genes. Total RNA was isolated using the RNeasy Micro Kit (QIAGEN, Hilden, Germany). The quality and the integrity of the RNA were assayed using an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, California). Subsequently, ribosomal RNA (rRNA) was depleted using the RNA Clean XP Kit (Beckman Coulter, Inc., Brea, California) and RNA was fragmented using RNase-Free DNase (QIAGEN). Immediately afterwards, first-strand cDNA was synthesized using the SuperScript II kit (Invitrogen, Carlsbad, California) while adding a hexamer random primer. Double-strand DNA fragments were ligated with a TruSeq adapter (Illumina, San Diego, California) and amplified with TruSeq PCR primers (Illumina) for sequencing.

Reads that were shorter than 25 nt and those that mapped to rRNA were discarded. The remaining clean reads were genome-mapped using spliced mapping in the Hisat2 package (Anaconda, Inc., Austin, Texas). Fragments per kilobase of transcript per million mapped reads (FPKM) of EdgeR (Anaconda, Inc.) were used to screen differentially expressed genes (DEGs). The p-value was corrected by multiple hypothesis testing, and the p-value threshold was determined by controlling false discovery rate (FDR). Corrected p-values are denoted as q-values. Transcripts with a q-value < 0.05 and a fold change (FC) > 2 were categorized as DEGs (Fig. 1).

Prediction and reannotation of novel lncRNAs, screening of differentially expressed lncRNAs, and prediction of lncRNA targets. Gffcompare (version 0.9.8; developed by the Center for Computational Biology at Johns Hopkins University, Baltimore, Maryland) was used to predict lncRNAs, which classified i, u, x categories by cufflinks: 'i', lncRNAs entirely within a reference intron; 'u', intergenic lncRNAs (lincRNAs); and 'x', lncRNAs with exonic overlap with reference on the opposite strand. A Perl script (O'Reilly Media Inc., Sebastopol, California) was used to

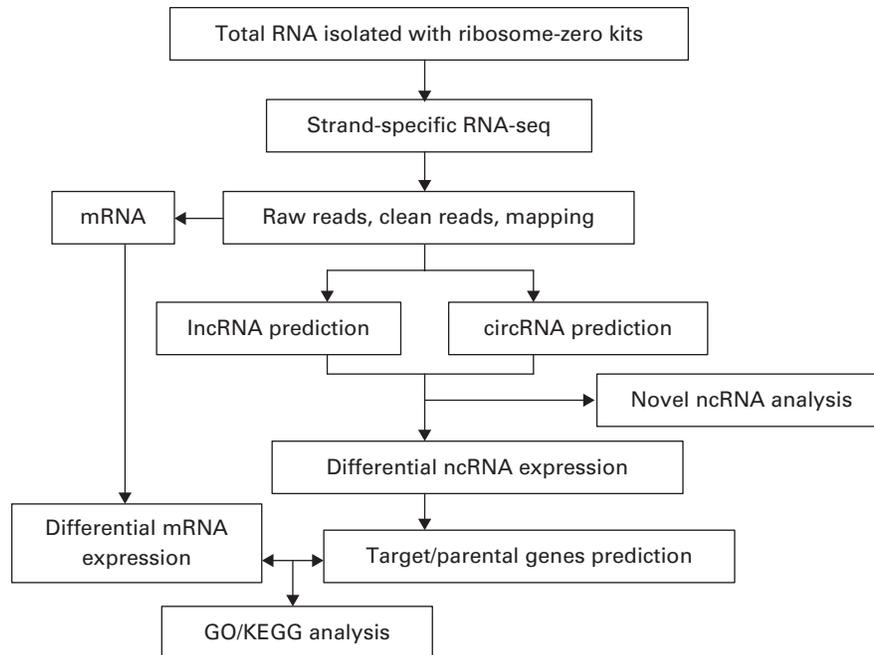


Fig. 1

Flowchart for analysis of whole-transcript sequencing analysis. RNA-seq, RNA-sequencing; mRNA, messenger RNA; lncRNA, long noncoding RNA; circRNA, circular RNA; ncRNA, noncoding RNA; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

reannotate, by tracing the genes corresponding to newly identified lncRNAs.

Stringtie (version 1.3.0; developed by the Center for Computational Biology at Johns Hopkins University) was used to calculate the expression levels of lncRNAs in NONCODE and Ensemble databases. EdgeR was used to screen differentially expressed lncRNAs (DELs) using the same criterion as DEGs.

Transregulation/cisregulation was used to predict lncRNA targets. The database used in transprediction was the same as for mRNA. BLAST (National Center for Biotechnology Information, Bethesda, Maryland) was used to select the complementary or similar sequences, RNAplex (National Center for Biotechnology Information) was used to calculate the complementary energy between the two sequences, and then sequences above the threshold were selected. For cisprediction, the distance between a lncRNA and its cisprediction gene had to be less than 10 kb.

Prediction, reannotation of novel circRNAs, screening of differentially expressed circRNAs, and prediction of parental genes of DECs. All clean reads were genome mapped using BWA-MEM⁶ to identify circRNAs. The CIRI programme⁷ was used to predict circRNAs. Based on the location of a particular circRNA, the same circRNA in each sample was merged and Blast-matched to the circBase database to identify novel circRNAs. A Perl script was used to classify circRNAs as either exonic region, intronic, or intergenic.

As most circRNAs cannot be used to acquire a complete sequence at present, the junction reads on the sites of back-splicing in circRNAs were used to calculate the expression levels and were normalized by the spliced

reads per billion mapping (SRPBM) programme.⁸ EdgeR was used to screen DECs using the same criterion as for DEGs. The parental gene (protein-encoding gene corresponding to the location of the circRNAs) was obtained according to the location of the circRNAs.

Interleukin (IL)-1 β -induced OA-like cell models. Cartilage was minced and digested at 37°C with 0.2% collagenase II (Sigma-Aldrich, St. Louis, Missouri) in complete culture medium of human articular chondrocytes (Procell, Wuhan, China) for six to seven hours, with stirring every 30 to 60 minutes after two hours. Chondrocytes were isolated after centrifugation and incubated at 37°C in a humidified atmosphere of 5% CO₂ for five to seven days until they grew to 80% confluence. Next, 5 ng/ml recombinant human IL-1 β (PeproTech, Rocky Hill, New Jersey) was treated for 18 hours to induce an OA-like phenotype in the chondrocytes.

Real-time quantitative reverse transcriptase polymerase chain reaction. Total RNA was extracted from chondrocytes and cartilage using TRIzol (Invitrogen). The PrimeScript II 1st Strand cDNA Synthesis Kit (Takara Bio Inc., Kusatsu, Japan) and SYBR Premix Ex Taq II (Takara) were used for reverse transcription and real-time quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) assay. Polymerase chain reaction primer sequences are listed in Supplementary Table i. The PCR products were verified by dissociation curves, and data were normalized to glyceraldehyde 3-phosphate dehydrogenase (GAPDH) expression to obtain Δ Ct values. Fold changes in expression were calculated using the 2^{- $\Delta\Delta$ Ct} method. Water was used as a negative and quality

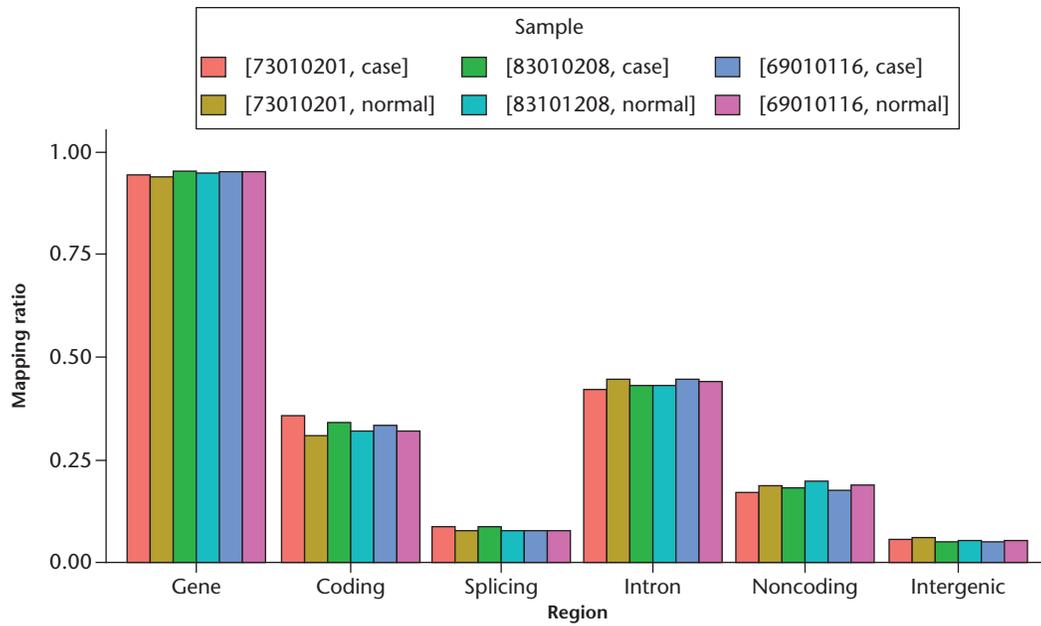


Fig. 2a

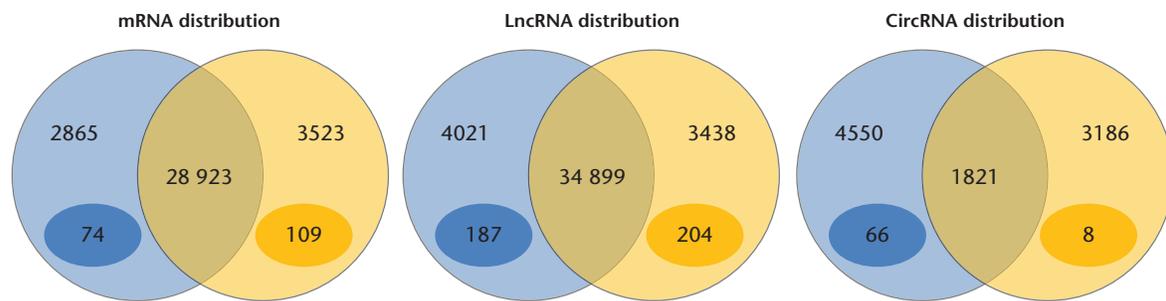


Fig. 2b

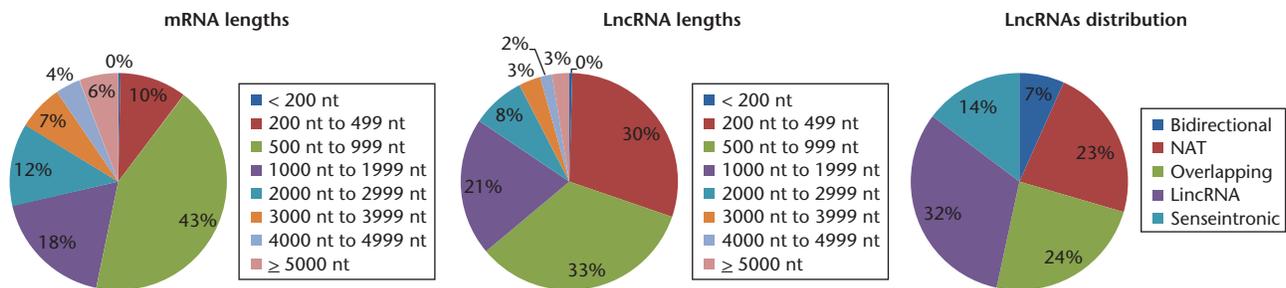


Fig. 2c

Whole-transcriptome profile of human articular cartilage. a) Chart showing the distribution of all the mapped reads in genome. b) Venn diagram of messenger RNAs (mRNAs), long noncoding RNAs (lncRNAs), and circular RNAs (circRNAs) in normal and osteoarthritic (OA) tissue of knee articular cartilage. Data in light blue and light yellow circles represent the number of genes expressed in the normal group and OA group, respectively. Data in deep blue ovals represent the number of genes that were expressed in each tissue of the normal group and in neither tissue of the OA group, while data in deep yellow ovals represent the number of genes that were expressed in each tissue of the OA group and in neither tissue of the normal group. c) Distribution and lengths of mRNAs and lncRNAs in whole-transcriptome profiles of human articular cartilage. NAT, natural antisense transcript; LincRNA, intergenic lncRNA.

control, and each sample was measured in triplicate. Genes with $2^{-\Delta\Delta ct} > 1.5$ were considered upregulation and $2^{-\Delta\Delta ct} < 0.8$ were considered downregulation.⁹

Accession codes. Whole-transcriptome sequencing data from six samples were submitted to the Sequence Read

Archive (SRA) of the National Center for Biotechnology Information (NCBI) separately, under the Bioproject accession number of PRJNA503001 and SRA run accession numbers of SRR8138531, SRR8138532, SRR8138533, SRR8138534, SRR8138535, SRR8138536.

Results

Whole-transcriptome profile of human articular cartilage.

An average of 104 729 039 raw reads were generated in six samples and 98 354 813 clean reads remained after trimming with a 97.85% mapping ratio (Fig. 2a). The ratios of gene region, coding, splicing, intron, noncoding (including 5' untranslated region (UTR), 3' UTR, and ncRNA regions), and intergenic region were nearly uniform in the different samples. A total of 50868 mRNAs, 42358 lncRNAs,

and 9556 circRNAs were detected separately. Overall, 109 mRNAs, 204 lncRNAs, and eight circRNAs were identified only in OA tissues. In contrast, 74 mRNAs, 187 lncRNAs, and 64 circRNAs were uniquely expressed in normal tissues (Fig. 2b). The mean length of lncRNAs (1249 nt) was shorter than that of mRNAs (1727 nt), and the majority of lncRNAs were 200 bp to 999 bp (63%), shorter than the mRNAs (61%, 500 nt to 1999 nt; Fig. 2c). Intergenic long noncoding RNA (32%), overlapping (24%), and natural antisense

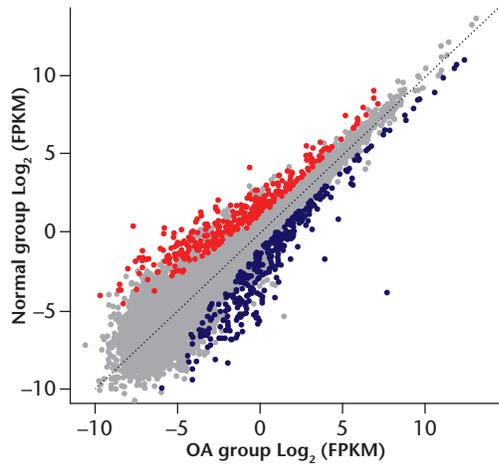


Fig. 3a

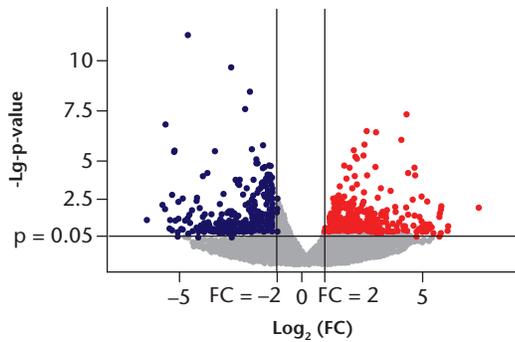


Fig. 3b

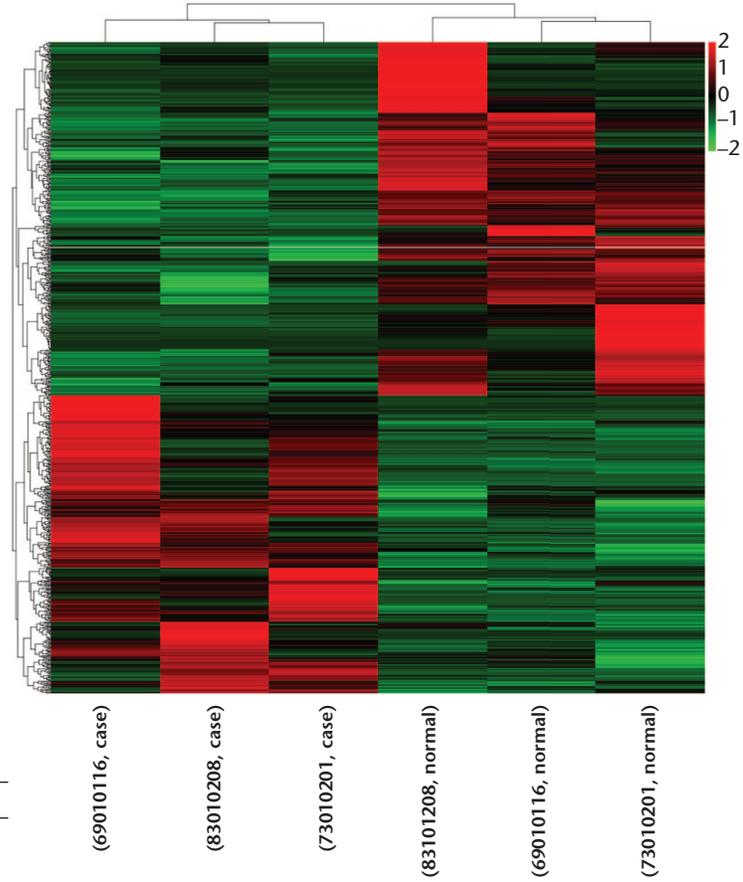


Fig. 3c

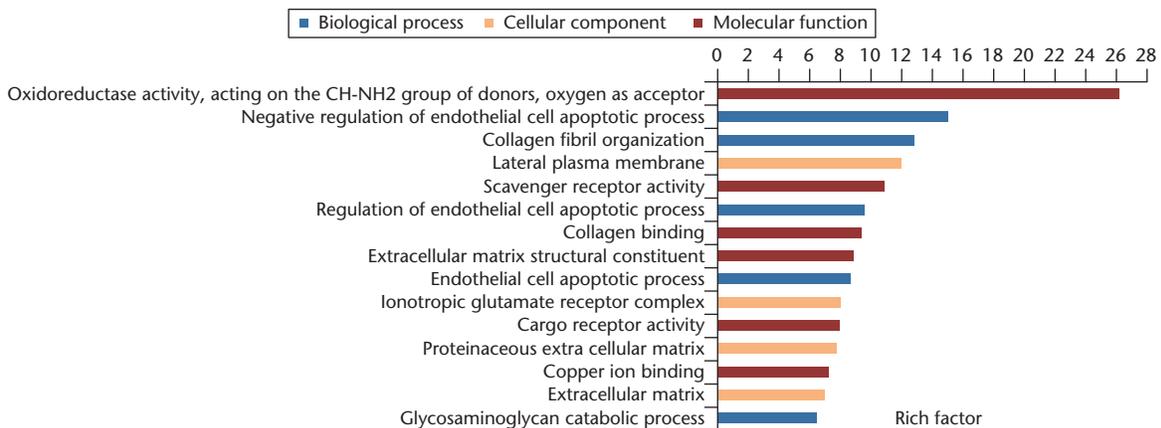


Fig. 3d

Fig.3 (Continued)

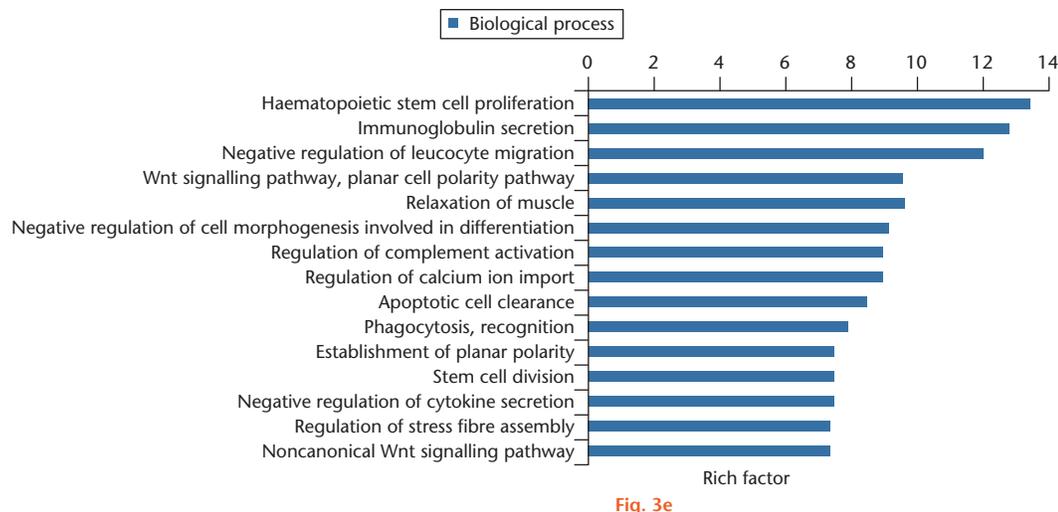


Fig. 3e

Differentially expressed genes (DEGs) in whole-transcriptome sequencing of osteoarthritic (OA) and normal tissue of knee articular cartilage. a) Scatter plot of messenger RNA (mRNA) expression. The mRNAs, represented as red points (high level) and blue points (low level), indicated a more than two-fold change of mRNAs between control normal and OA cartilage samples. b) Volcano plot of the differentially expressed mRNAs. The red points (high level) and blue points (low level) in the plot represent the differentially expressed mRNAs with statistical significance. c) Hierarchical clustering shows a difference in mRNA expression profile between the two groups and homogeneity within groups. d) The top 15 highest enriched gene ontology (GO) terms for downregulated mRNAs in the intact cartilage. e) The top 15 highest enriched GO terms for upregulated mRNAs in the intact cartilage. FPKM, fragments per kilobase of transcript per million mapped reads; FC, fold change; ESCRT, endosomal sorting complexes required for transport; cAMP, cyclic adenosine monophosphate.

transcript (NAT; 23%) account for the most abundant types of lncRNAs. Circular RNAs were classified as exonic (89.9%), intronic (6.4%), and intergenic regions (3.7%).

DEG screening, gene ontology, and Kyoto Encyclopedia of Genes and Genomes pathway analysis. In total, 338 mRNAs were significantly downregulated ($\log_2FC \leq -1$; $q\text{-value} < 0.05$) and 401 mRNAs were significantly upregulated ($\log_2FC \geq 1$; $q\text{-value} < 0.05$) in the intact *versus* damaged cartilage. The top 15 upregulated and downregulated DEGs are shown in Supplementary Table ii. Differentially expressed genes were hierarchically clustered, as shown in Figures 3a to 3c, and were enriched for several gene ontology (GO) pathways closely related to the anabolic and catabolism of cartilage matrix (e.g. collagen binding, extracellular matrix structural constituent, proteinaceous extracellular matrix, and extracellular matrix; Figs 3d and 3e). A total of 78 and 31 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways with enrichment factor > 1 that corresponded to up- and downregulated DEGs, respectively, were identified. These included extracellular matrix (ECM)-receptor interaction, the transforming growth factor (TGF)- β signalling pathway, and the osteoclast differentiation and insulin signalling pathway, which are associated with OA development (Supplementary Tables iii and iv).

Prediction and reannotation of novel lncRNAs. A total of 2589 novel lncRNAs were predicted, including 'i' class (1822), 'x' class (541), and 'u' class (226). For 'x' and 'u' classes, a Perl script was used to identify the corresponding genes and the genes at both ends, while for 'i' class, a Perl script revealed the nearest genes at both ends. The

2589 predicted novel lncRNAs were reannotated as 7777 genes, with an average of three genes per lncRNAs.

DEL screening, GO and KEGG pathway analysis of target genes, overlapping collection of DEGs, and target genes of DELs. In all, 506 and 645 DELs were significantly down- and upregulated ($\log_2FC \leq -1$ or $\log_2FC \geq 1$; $q\text{-value} < 0.05$), respectively, in intact *versus* damaged cartilage, which were hierarchically clustered as shown in Figure 4. The top 15 DELs are presented in Supplementary Table v. Furthermore, 45 294 target genes were predicted for 979 DELs. Of these, the majority were repeatedly predicted by different lncRNAs. As such, a total of 5288 target genes were obtained.

The predicted target genes of DELs were highly enriched for GO, including collagen binding involved in the cell matrix, AP-2 adaptor complex binding, tubulin deacetylation, and leucine catabolic process (Figs 4d and 4e). Briefly, 176 and 198 KEGG pathways were enriched by the predicted target genes of down- and upregulated DELs separately, including mineral absorption, TGF- β signalling, endocrine- and other factor-regulated calcium reabsorption and mineral absorption pathway. These KEGG pathways may be related to OA development (Supplementary Tables vi and vii).

To increase the likelihood of identifying key genes involved in the OA process, a set of overlapping DEGs and target genes of DELs was constructed. A total of 165 overlapping genes were identified, of which 64 were found to be downregulated genes and 101 were found to be upregulated genes, in intact *versus* damaged cartilage (Supplementary Table viii). The down- and upregulated overlapping genes were enriched for GO terms

including myotube differentiation, striated muscle cell differentiation, and cell adhesion molecule binding (Figs 4f and 4g). Extracellular matrix, proteinaceous extracellular matrix, cellular response to extracellular stimulus, cartilage development, and connective tissue development were also enriched and may be associated with OA progress. Furthermore, 13 and 60 KEGG pathways of down- and upregulated overlapping genes, respectively, were enriched and included mucin type O-glycan biosynthesis, calcium signalling pathway, and insulin signalling pathway (Supplementary Tables ix and x).

Screening of DECs and GO analysis of their parental genes. A total of 9556 circRNAs were predicted, of which 5536 (57.9%) were novel, based on circBase. Four circRNAs (including two novel circRNAs) were significantly down-regulated and 38 (including 12 novel circRNAs) were significantly upregulated in intact *versus* damaged cartilage ($FC \geq 2$; p -value < 0.05 ; Fig. 5a). The top 15 DECs are shown in Supplementary Table xi. The parental genes of DECs were highly enriched for regulation of small GTPase-mediated signal transduction, GTPase activator activity, and regulation of intracellular signal transduction (Fig. 5b).

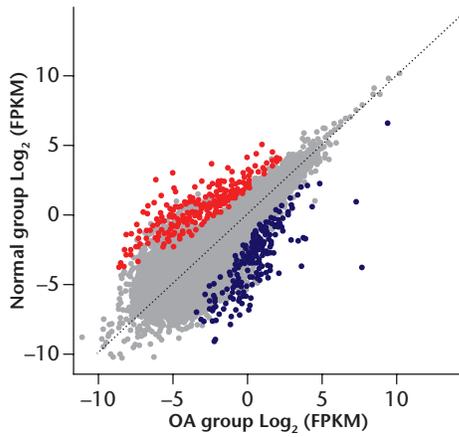


Fig. 4a

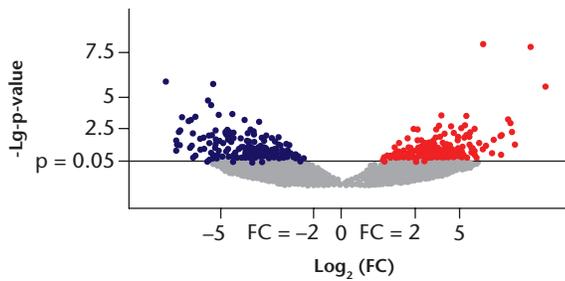


Fig. 4b

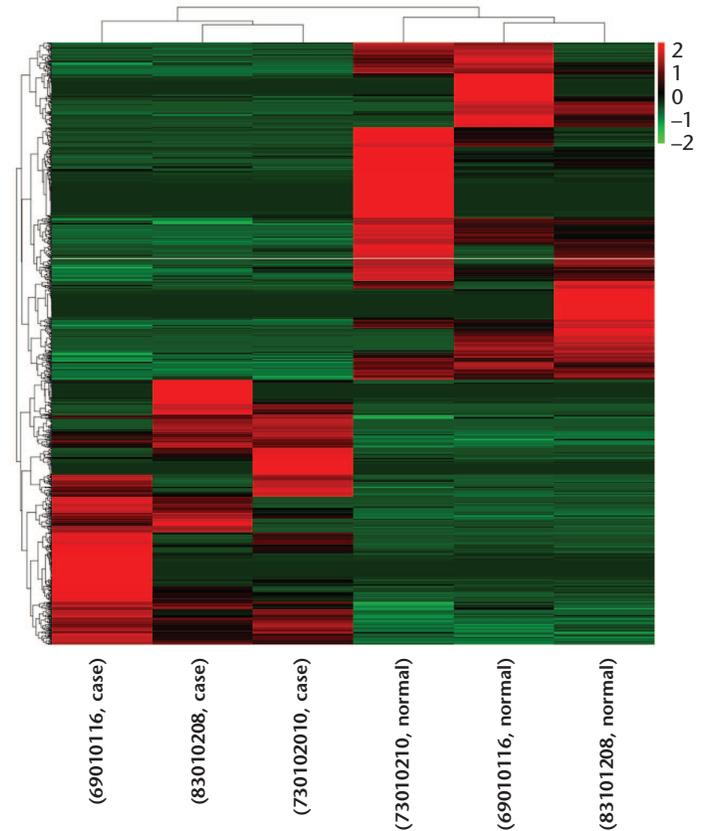


Fig. 4c

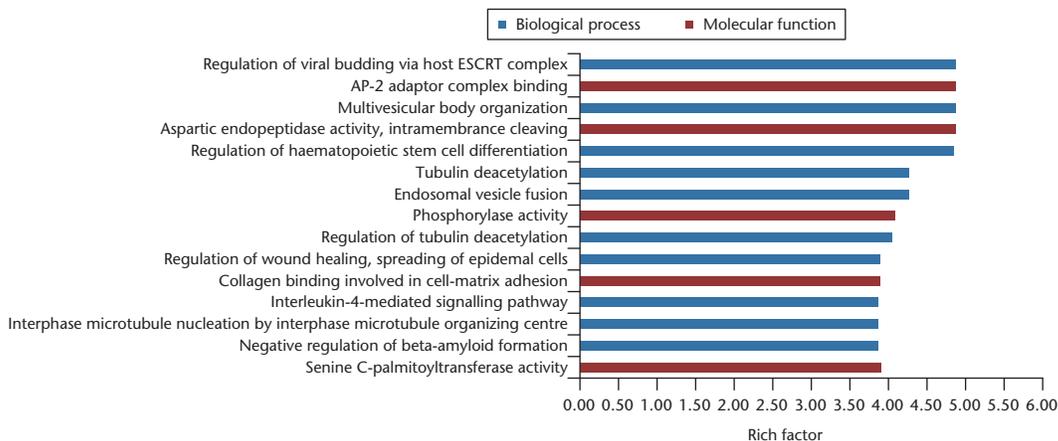


Fig. 4d

Fig.4 (Continued)

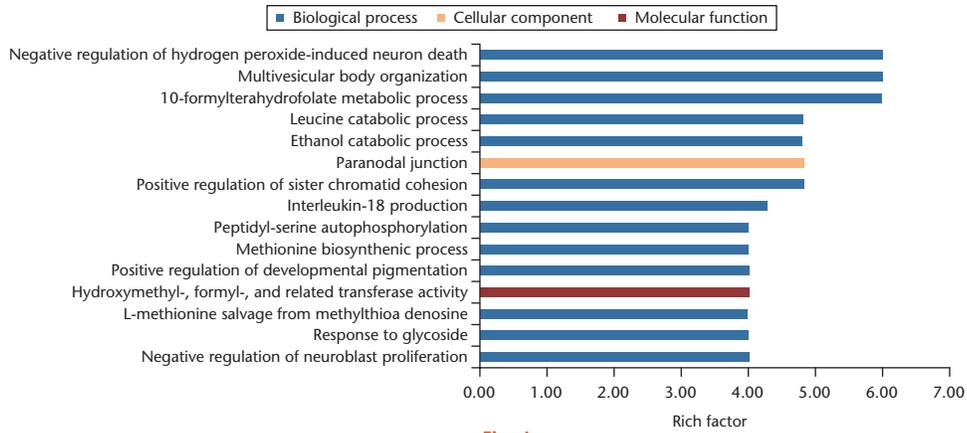


Fig. 4e

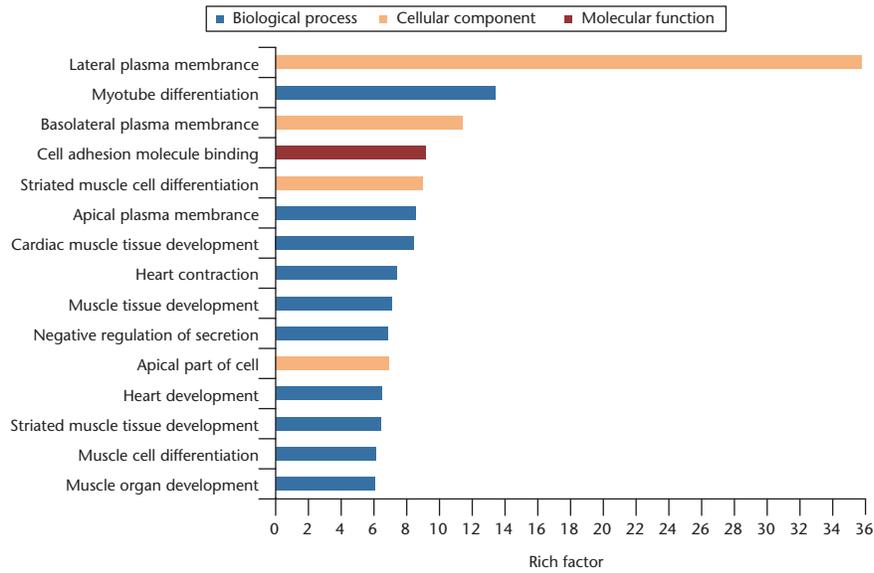


Fig. 4f

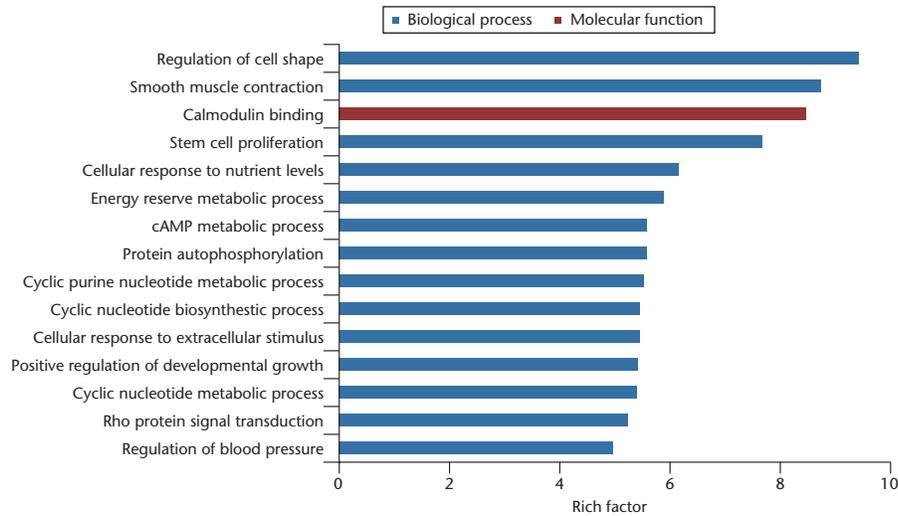


Fig. 4g

Differentially expressed long noncoding RNAs (lncRNAs) (DEs) in whole-transcriptome sequencing of osteoarthritic (OA) and normal tissue of knee articular cartilage. a) Scatter plot of lncRNA expression. The lncRNAs, represented as red points (high level) and blue points (low level), indicated a more than two-fold change of lncRNAs between control normal and OA cartilage samples. b) Volcano plot of the DEs. The red points (high level) and blue points (low level) in the plot represent the DEs with statistical significance. c) Hierarchical clustering shows a difference in lncRNA expression profile between the two groups and homogeneity within groups. d) The top 15 highest enriched gene ontology (GO) terms for target messenger RNAs (mRNAs) of downregulated DEs. e) The top 15 highest enriched GO terms for target mRNAs of upregulated DEs. f) Top 15 highest enriched GO terms for downregulated overlapping genes of differentially expressed genes (DEGs) and target genes of DEs. g) Top 15 highest enriched GO terms for upregulated overlapping genes of DEGs and target genes of DEs. FPKM, fragments per kilobase of transcript per million mapped reads; FC, fold change; cAMP, cyclic adenosine monophosphate.

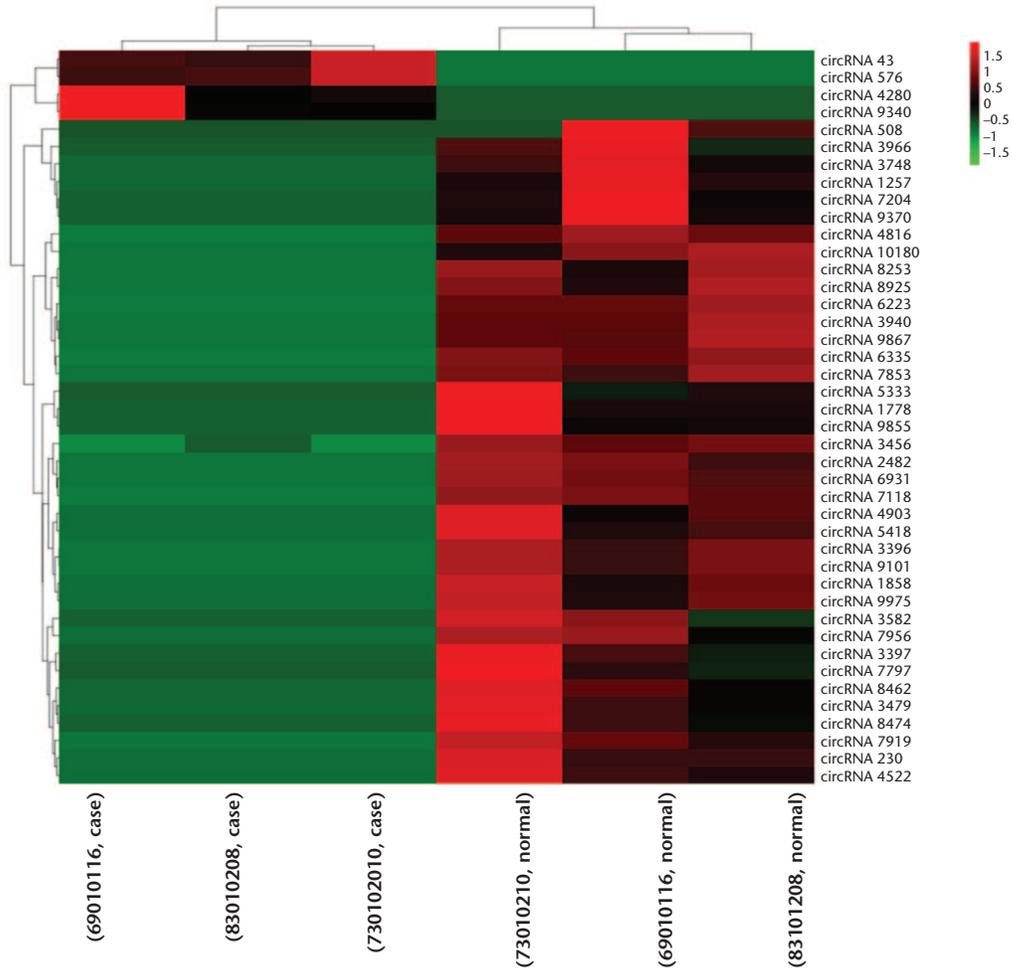


Fig. 5a

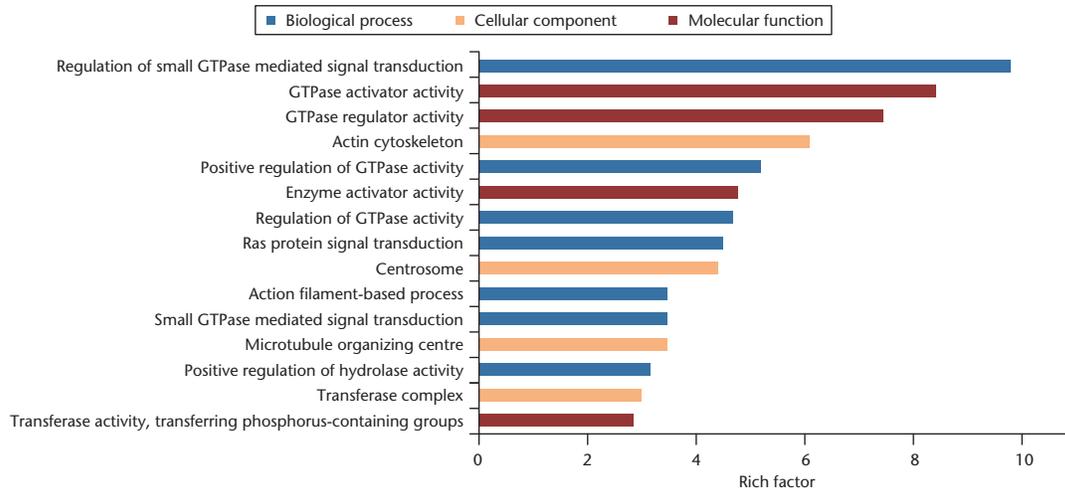


Fig. 5b

Differentially expressed circRNAs (DECs) in whole-transcriptome sequencing of osteoarthritic (OA) and normal tissue of knee articular cartilage and gene ontology (GO) analysis of their parental genes. a) DECs in whole-transcriptome sequencing of OA and normal tissue of knee articular cartilage. Hierarchical clustering shows a difference in circRNA expression profile between the two groups and homogeneity within groups. b) GO analysis for parental genes of DECs. The top 15 highest enriched GO terms in three domains were demonstrated.

Table I. Candidate genes with biological function related to bone and cartilage

Gene name	Biological function	Log ₂ FC	Regulation
<i>ACE</i> ¹⁰	Gene polymorphism correlated with OA	3.86	Up
<i>ADAM12</i> ¹¹	Cartilage degradation	-1.33	Down
<i>ASPN</i> ¹²	Susceptibility gene for OA	-1.11	Down
<i>ATF3</i> ¹³	OA development, bone formation	2.13	Up
<i>BMP3</i> ¹⁴	Cartilage repair, bone formation	2.72	Up
<i>CHI3L1</i> ¹⁵	OA development	2.17	Up
<i>GALNT16</i> ¹⁶	TGF-beta/BMP signalling	-2.89	Down
<i>GAS7</i> ¹⁷	Chondrogenesis	1.51	Up
<i>LEPR</i> ¹⁸	Susceptibility gene for OA	2.03	Up
<i>NFAM1</i> ¹⁹	Osteoclast formation and bone resorption	1.60	Up
<i>PRELP</i> ²⁰	Function on osteoclasts, osteoblasts, and cartilage	-1.13	Down
<i>WISP2</i> ²¹	Rheumatoid arthritis	1.76	Up
<i>C3</i> ²²	Rheumatoid arthritis, osteoporosis	3.08	Up
<i>C7</i> ²³	Rheumatoid arthritis	4.35	Up
<i>FAM83G</i> ²⁴	Type I BMP	-1.16	Down
<i>JUN</i> ²⁵	OA development	1.38	Up
<i>EVX1</i> ²⁶	Joint formation	2.64	Up
<i>BARX2</i> ²⁷	Cartilage formation	-1.58	Down

FC, fold change; OA, osteoarthritis; TGF, transforming growth factor; BMP, bone morphogenetic protein

Table II. The overlapping differentially expressed genes (DEGs) screened out from seven different research studies, including the present study

Resources	Number	List of overlapping DEGs
From all seven results	1	<i>FN1</i>
From six results including present study	4	<i>ADAM12, TNFAIP6, CRLF1, COL5A1</i>
From five results including present study	8	<i>HTRA1, SOX11, GFRA2, ST6GALNAC5, MSX2, DNER, OGN, CRTAC1</i>
From four results including present study	15	<i>TSPAN2, BAALC, SLC7A5, NKX2-5, GIPR, ADAMTS6, UROCT1, IL11, TF, P4HA3, ASPN, AOC2, COL15A1, PTGES, TPPP3</i>
From three results including present study	27	<i>SORT1, PTGER1, BVES, C3ORF52, ENOX1, RNF152, TNFRSF11B, GLIS1, PSAT1, ANK3, GRIA2, SLC16A10, SRPX2, LOXL3, NOG, ANGPTL, CILP 4, OMD, SIPA1L2, GJB2, MAMDC2, HHIPL1, SPSB4, TMEM59L, LY6D, SLC2A12, COL3A1</i>

Identifying candidate genes for PCR-validation. The sets of 165 overlapping DEGs and target genes of DELs (Supplementary Table viii) were further searched in the PubMed literature database. Each “Gene symbol” (gene name) was used as a query keyword and searched in the “title/abstract”. A total of 18 genes were previously implicated in bone and cartilage diseases, while 111 genes were found to be associated with other conditions such as tumours, Alzheimer’s, heart disease, and diabetes. The remaining 36 genes lacked experimental or clinical validations. As their expression profiles in OA cartilage were still unknown, these 18 genes were selected to be further analyzed (Table I).¹⁰⁻²⁷

Previous research has revealed the expression pattern of mRNAs in OA cartilage using high-throughput commercialized²⁸⁻³¹ or laboratory-developed^{32,33} gene microarrays and RNA-sequencing technology.³⁴ The DEGs of the present study were compared with the data from six independent laboratories. Table II shows the overlapping DEGs identified in seven research studies.²⁸⁻³⁴ Fibronectin 1 (*FN1*) was the only gene present in all seven datasets. This gene encodes one of the most abundant proteins in OA synovium; its decrease in mechanically stimulated human chondrocyte cells was shown to mediate the chondroprotective action of diacerein.^{35,36} Four genes were identified in six DEG sets: *ADAM12*, which is among

the 18 genes selected for validation; *TNFAIP6* (also known as *TSG-6*), which is enhanced in OA synovial fluids,³⁷ is thought to function by blocking matrix assembly and futile synthesis, and is correlated with increased OA risk;³⁸ *COL5A1*, which is involved in type V collagen biosynthesis, and may have roles in regulating the width of fibrils and controlling the assembly of other types of collagen into fibrils in joint cartilage;³⁹ and cytokine receptor-like factor 1 (*CRLF1*). The role of *CRLF1* is less well known⁴⁰ and it was therefore chosen for further validation.

Finally, since very little information was available on 1151 DELs, the top five upregulated and top five down-regulated DELs were selected for further analysis.

Expression of DEGs and DELs in IL-1 β -induced OA-like cell models. Real-time qRT-PCR was used to validate the expression of the 18 chosen genes in human articular chondrocytes *in vitro*. *ADAM12* (FC = 0.296), *ASPN* (FC = 0.57), and *PRELP* (FC = 0.40) were significantly decreased in normal *versus* IL-1 β -treated OA-like cell lines (Fig. 6a). In contrast, *CHI3L1* (FC = 4.48), *NFAM1* (FC = 4.03), *ATF3* (FC = 2.56), *WISP2* (FC = 1.86), and *LEPR* (FC = 1.76) were significantly increased in normal chondrocytes. *EVX1*, *BARX2*, and *FAM83G* showed opposing trends based on sequencing and qRT-PCR analysis, although they were significantly differentially expressed between the OA and normal groups. The other eight

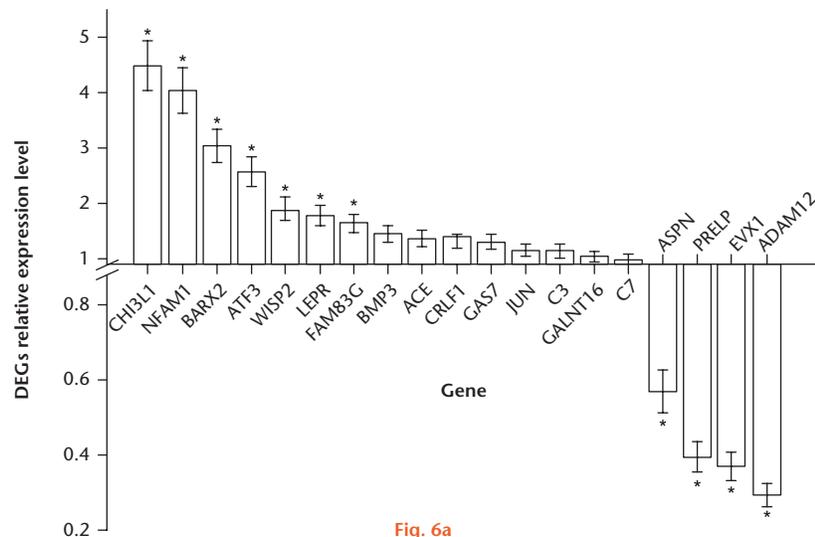


Fig. 6a

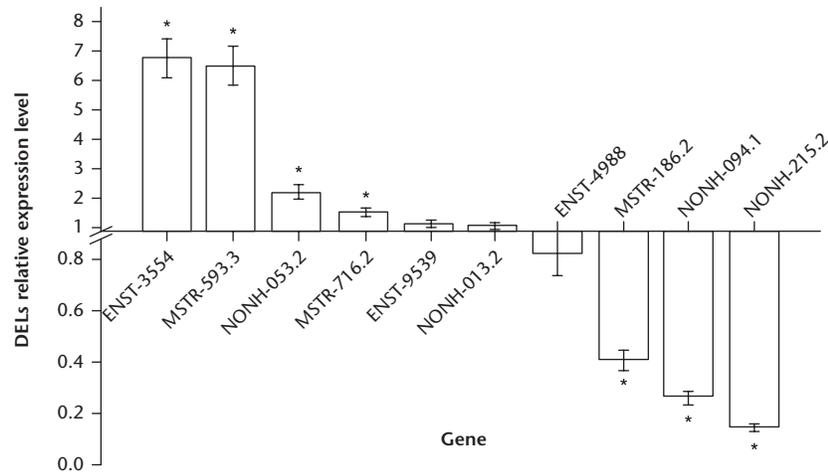


Fig. 6b

Validation of candidate genes and long noncoding RNAs (lncRNAs) by real-time reverse transcriptase polymerase chain reaction (RT-PCR) in interleukin (IL)-1 β -induced osteoarthritis (OA)-like cell models. a) Relative expression levels of candidate genes in normal *versus* IL-1 β -treated OA chondrocytes. b) Relative expression levels of candidate differentially expressed lncRNAs (DELS) in normal *versus* IL-1 β -treated OA cell lines. *Significantly expressed genes and DELs. DEG, differentially expressed gene.

genes seemed to be independent of OA progress in IL-1 β -induced OA-like cell models.

Some lncRNAs, including *MSTRG.44186.2* (FC = 0.40), *NONHSAT186094.1* (FC = 0.26), and *NONHSAT123215.2* (FC = 0.14) were found to be significantly decreased in normal *versus* OA-like chondrocytes. In contrast, *ENST00000453554* (FC = 6.76), *MSTRG.99593.3* (FC = 6.52), *MSTRG.138716.2* (FC = 1.52), and *NONHSAT044053.2* (FC = 2.21) were significantly increased in normal chondrocytes. Finally, *NONHSAT044053.2* showed an opposing trend based on sequencing and qRT-PCR analysis (Fig. 6b).

In total, the expression profiles of 14 genes comprising eight DEGs (*ADAM12*, *ASPN*, *PRELP*, *CHI3L1*, *NFAM1*, *ATF3*, *WISP2*, and *LEPR*) and six DELs (*MSTRG.44186.2*, *NONHSAT186094.1*, *NONHSAT123215.2*, *ENST00000453554*, *MSTRG.99593.3*, and *MSTRG.138716.2*) were shown to be influenced by IL-1 β in chondrocytes cultured *in vitro*, and therefore may be involved in chondrocyte apoptosis and OA development.

Expressions of DEGs and DELs in OA articular cartilage.

Based on our *in vivo* PCR validation, the expression patterns of 14 genes were found to be consistent with the sequencing results, although four out of the 14 were not significantly different between normal and OA tissue. Among the DEGs, *WISP2* (FC = 5.04), *ATF3* (FC = 2.13), and *CHI3L1* (FC = 1.63) were significantly increased in normal *versus* OA tissue, while *ADAM12* (FC = 0.56), *PRELP* (FC = 0.31), and *ASPN* (FC = 0.30) were significantly decreased in normal *versus* OA tissue. Among the DELs, we observed increased expression of *ENST00000453554* (FC = 6.82) and *MSTRG.99593.3* (FC = 6.15), and decreased expression of *MSTRG.44186.2* (FC = 0.65) and *NONHSAT186094.1* (FC = 0.10) in normal *versus* damaged tissue (Fig. 7).

Discussion

The term 'transcriptome' has been expanded to include several types of transcripts, including mRNAs and

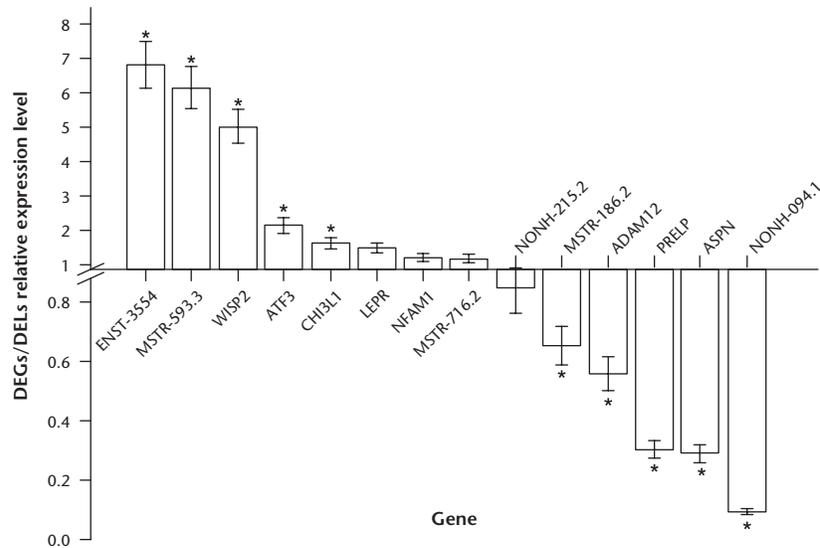


Fig. 7

Validation of candidate genes and long noncoding RNAs (lncRNAs) in paired osteoarthritic (OA) articular cartilage by real-time reverse transcriptase polymerase chain reaction (RT-PCR). *Significantly expressed genes and differentially expressed lncRNAs (DElS). DEGs, differentially expressed genes.

noncoding RNAs, which affords novel opportunities for comprehensive RNA expression profiling, as well as discovering novel lncRNAs and circRNAs. In this study, we simultaneously detected mRNAs, lncRNAs, and circRNAs in OA articular cartilage for the first time. A total of 50 868 mRNAs, 42 358 lncRNAs, and 9556 circRNAs were identified, among which 109 mRNAs, 204 lncRNAs, and eight circRNAs were exclusively detected in damaged OA tissue, suggesting that they may be associated with OA progress. Additionally, 2589 lncRNAs (6.11%) and 5536 circRNAs (57.9%) were identified as novel genes, providing a resource of new genes to explore the function of ncRNA in OA progression.

Whole-transcriptome sequencing results in 739 DEGs, 1151 DElS, and 42 DECs. From this, DEGs, target genes of DElS, and parental genes of DECs were separately assessed for enrichment of GO terms and KEGG pathways. Of primary focus were pathways involved in cartilage synthesis and metabolism, including collagen fibril organization, extracellular matrix, collagen binding, proteinaceous extracellular matrix, TGF-beta signalling pathway, osteoclast differentiation, mineral absorption, and insulin signalling pathway. In addition, some pathways, such as negative regulation of endothelial cell apoptotic process, immunoglobulin secretion, and regulation of calcium ion import seemed less relevant to OA progress, but may provide insight into novel OA pathogenic factors.

As whole-transcriptome sequencing can simultaneously identify DEGs and DElS, target genes of DElS that also belonged to the identified DEGs were of particular interest. If they were involved in OA progress, their functions may be controlled by corresponding lncRNAs. *WISP2*, *ATF3*, *CHI3L1*, *ADAM12*, *PRELP*, and *ASPN*

were previously suggested to be involved in OA progress. For these genes, elucidating their interactions with corresponding lncRNAs identified in this study will contribute to a better understanding of the lncRNA-regulated molecular mechanism underlying OA development. Interestingly, *ADAM12* and *ASPN* were previously identified as DEGs in seven independent OA-related mRNA microarrays or sequencing experiments.²⁸⁻³⁴ Nonetheless, how these genes are regulated by lncRNAs in the progression of OA is still unclear.

In addition, the expression profiles of *ENST0000045354*, *MSTRG.99593.3*, *MSTRG.44186.2*, and *NONHSAT186094.1* were found to be significantly altered in OA progress, which provided new OA-related lncRNA candidates for further research. In previous studies, Fu et al,³⁰ Xing et al,⁴¹ and Liu et al⁴² independently identified 4707, 121, and 152 DElS, respectively, in OA *versus* normal cartilage using lncRNA microarrays. The vast differences in the number of identified DElS depended largely on how they defined their threshold of differential genes: Liu et al⁴² and Xing et al⁴¹ used eight-fold changes, while Fu et al³⁰ used two-fold changes as the threshold, which was similar to our study. Furthermore, because we used a sequencing method rather than a microarray (as in most previous studies), 2589 novel lncRNAs specific to knee articular cartilage were identified. Moreover, 42 DECs in OA *versus* normal cartilage and eight circRNAs exclusively expressed in OA cartilage were reported, which provided useful noncoding gene candidates for further work on OA pathogenesis.

Our study has some limitations. First, experimental validation of selected OA-related genes and lncRNAs merits further research, as does an in-depth examination of differentially expressed circRNAs. Likewise, efforts should be made to identify the specific roles of other DEGs

identified in this study and to place them in the larger context of OA progression. Finally, future studies should focus on obtaining larger datasets from OA patients, so that risk ratios of candidate genes can be analyzed, and useful OA biomarkers can be utilized in clinical diagnostic and treatment applications.

In conclusion, we identified four important lncRNAs and six genes that were targeted by DELs that regulate OA progress. Further research into the functions of these genes and their interaction with lncRNAs may lead to a better understanding of their biogenesis and the mechanism underlying OA. The filtered lncRNAs and circRNAs pool of this study contributes to an increased understanding of the basic features of human knee cartilage tissue and may eventually be used as diagnostic biomarkers and therapeutic agents for OA patients.

Supplementary Material

 Tables illustrating data related to: polymerase chain reaction (PCR) primer sequences; gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses; screening of differentially expressed genes (DEGs), differentially expressed long noncoding RNAs (DELs), and differentially expressed circular RNAs (DECs); and the top 15 up- and downregulated DEGs, DELs, and DECs.

References

- Tian H. Detection of differentially expressed genes involved in osteoarthritis pathology. *J Orthop Surg Res* 2018;13:49-56.
- He A, Ning Y, Wen Y, et al. Use of integrative epigenetic and mRNA expression analyses to identify significantly changed genes and functional pathways in osteoarthritic cartilage. *Bone Joint Res* 2018;7:343-350.
- Cen X, Huang XQ, Sun WT, Liu Q, Liu J. Long noncoding RNAs: a new regulatory code in osteoarthritis. *Am J Transl Res* 2017;9:4747-4755.
- Zhang G, Wu Y, Xu D, Yan X. Long noncoding RNA UFC1 promotes proliferation of chondrocyte in osteoarthritis by acting as a sponge for miR-34a. *DNA Cell Biol* 2016;35:691-695.
- Liu Q, Zhang X, Hu X, et al. Emerging roles of circRNA related to the mechanical stress in human cartilage degradation of osteoarthritis. *Mol Ther Nucleic Acids* 2017;7:223-230.
- Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* 2013:1303.3997. <https://arxiv.org/abs/1303.3997> (date last accessed 12 June 2019).
- Gao Y, Wang J, Zhao F. CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. *Genome Biol* 2015;16:4.
- Zheng Q, Bao C, Guo W, et al. Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs. *Nat Commun* 2016;7:11215.
- Xu L, Li L, Li J, et al. Overexpression of miR-1260b in non-small cell lung cancer is associated with lymph node metastasis. *Aging Dis* 2015;6:478-485.
- Poornima S, Subramanyam K, Khan IA, Hasan Q. The insertion and deletion (I28005D) polymorphism of the angiotensin I converting enzyme gene is a risk factor for osteoarthritis in an Asian Indian population. *J Renin Angiotensin Aldosterone Syst* 2015;16:1281-1287.
- Poonpet T, Tammachote R, Tammachote N, Kanitnate S, Honsawek S. Association between ADAM12 polymorphism and knee osteoarthritis in Thai population. *Knee* 2016;23:357-361.
- Liang W, Gao B, Xu G, et al. Association between single nucleotide polymorphisms of asporin (ASPN) and BMP5 with the risk of knee osteoarthritis in a Chinese Han population. *Cell Biochem Biophys* 2014;70:1603-1608.
- Iezaki T, Ozaki K, Fukasawa K, et al. ATF3 deficiency in chondrocytes alleviates osteoarthritis development. *J Pathol* 2016;239:426-437.
- Zhang Z, Yang W, Cao Y, et al. The functions of BMP3 in rabbit articular cartilage repair. *Int J Mol Sci* 2015;16:25934-25946.
- Szyclynska MA, Trovato FM, Di Rosa M, et al. Co-expression and co-localization of cartilage glycoproteins CHI3L1 and lubricin in osteoarthritic cartilage: morphological, immunohistochemical and gene expression profiles. *Int J Mol Sci* 2016;17:359.
- Voglmeir J, Laurent N, Flitsch SL, Oelgeschläger M, Wilson IB. Biological and biochemical properties of two *Xenopus laevis* N-acetylgalactosaminyltransferases with contrasting roles in embryogenesis. *Camp Biochem Physiol B Biochem Mol Biol* 2015;180:40-47.
- Chang Y, Ueng SW, Lin-Chao S, Chao CC. Involvement of Gas7 along the ERK1/2 MAP kinase and SOX9 pathway in chondrogenesis of human marrow-derived mesenchymal stem cells. *Osteoarthritis Cartilage* 2008;16:1403-1412.
- Yang J, Du H, Lv J, Zhang L. Association of rs1137101 polymorphism in LEPR and susceptibility to knee osteoarthritis in a Northwest Chinese Han population. *BMC Musculoskelet Disord* 2016;17:311.
- Sambandam Y, Sundaram K, Saigusa T, Balasubramanian S, Reddy SV. NFAM1 signaling enhances osteoclast formation and bone resorption activity in Paget's disease of bone. *Bone* 2017;101:236-244.
- Li H, Cui Y, Luan J, et al. PRELP promotes osteoblastic differentiation of preosteoblastic MC3T3-E1 cells by regulating the beta-catenin pathway. *Biochem Biophys Res Commun* 2016;470:558-562.
- Tanaka I, Morikawa M, Okuse T, Shirakawa M, Imai K. Expression and regulation of WISP2 in rheumatoid arthritic synovium. *Biochem Biophys Res Commun* 2005;334:973-978.
- MacKay DL, Kean TJ, Bernardi KG, et al. Reduced bone loss in a murine model of postmenopausal osteoporosis lacking complement component 3. *J Orthop Res* 2018;36:118-128.
- Gang X, Sun Y, Li F, et al. Identification of key genes associated with rheumatoid arthritis with bioinformatics approach. *Medicine (Baltimore)* 2017;96:e7673.
- Vogt J, Dingwell KS, Herhaus L, et al. Protein associated with SMAD1 (PAWS1/FAM83G) is a substrate for type I bone morphogenetic protein receptors and modulates bone morphogenetic protein signalling. *Open Biol* 2014;4:130210.
- Rhee J, Park SH, Kim SK, et al. Inhibition of BATF/JUN transcriptional activity protects against osteoarthritic cartilage destruction. *Ann Rheum Dis* 2017;76:427-434.
- Schulte CJ, Allen C, England SJ, Juárez-Morales JL, Lewis KE. Evx1 is required for joint formation in zebrafish fin dermoskeleton. *Dev Dyn* 2011;240:1240-1248.
- Tsau C, Ito M, Gromova A, et al. Barx2 and Fgf10 regulate ocular glands branching morphogenesis by controlling extracellular matrix remodeling. *Development* 2011;138:3307-3317.
- Karlsson C, Dehne T, Lindahl A, et al. Genome-wide expression profiling reveals new candidate genes associated with osteoarthritis. *Osteoarthritis Cartilage* 2010;18:581-592.
- Geyer M, Grässel S, Straub RH, et al. Differential transcriptome analysis of intraarticular lesional vs intact cartilage reveals new candidate genes in osteoarthritis pathophysiology. *Osteoarthritis Cartilage* 2009;17:328-335.
- Fu M, Huang G, Zhang Z, et al. Expression profile of long noncoding RNAs in cartilage from knee osteoarthritis patients. *Osteoarthritis Cartilage* 2015;23:423-432.
- Snelling S, Rout R, Davidson R, et al. A gene expression study of normal and damaged cartilage in anteromedial gonarthrosis, a phenotype of osteoarthritis. *Osteoarthritis Cartilage* 2014;22:334-343.
- Aigner T, Fundel K, Saas J, et al. Large-scale gene expression profiling reveals major pathogenetic pathways of cartilage degeneration in osteoarthritis. *Arthritis Rheum* 2006;54:3533-3544.
- Aigner T, Saas J, Zien A, et al. Analysis of differential gene expression in healthy and osteoarthritic cartilage and isolated chondrocytes by microarray analysis. *Methods Mol Med* 2004;100:109-128.
- Dunn SL, Soul J, Anand S, et al. Gene expression changes in damaged osteoarthritic cartilage identify a signature of non-chondrogenic and mechanical responses. *Osteoarthritis Cartilage* 2016;24:1431-1440.
- Steinecker-Frohnwieser B, Weigl L, Kullich W, Lohberger B. The disease modifying osteoarthritis drug diacerein is able to antagonize pro inflammatory state of chondrocytes under mild mechanical stimuli. *Osteoarthritis Cartilage* 2014;22:1044-1052.
- Okabe T, Ohmori Y, Tanigami A, et al. Detection of gene expression in synovium of patients with osteoarthritis using a random sequencing method. *Acta Orthop* 2007;78:687-692.
- Wisniewski HG, Colón E, Liublinska V, et al. TSG-6 activity as a novel biomarker of progression in knee osteoarthritis. *Osteoarthritis Cartilage* 2014;22:235-241.
- Chou CH, Attarian DE, Wisniewski HG, Band PA, Kraus VB. TSG-6 - a double-edged sword for osteoarthritis (OA). *Osteoarthritis Cartilage* 2018;26:245-254.
- Tseou A. Osteoarthritis year in review 2014: genetics and genomics. *Osteoarthritis Cartilage* 2014;22:2017-2024.

40. Piras R, Chiappe F, Torraca IL, et al. Expanding the mutational spectrum of CRLF1 in Crisponi/CISS1 syndrome. *Hum Mutat* 2014;35:424-433.
41. Xing D, Liang JQ, Li Y, et al. Identification of long noncoding RNA associated with osteoarthritis in humans. *Orthop Surg* 2014;6:288-293.
42. Liu Q, Zhang X, Dai L, et al. Long noncoding RNA related to cartilage injury promotes chondrocyte extracellular matrix degradation in osteoarthritis. *Arthritis Rheumatol* 2014;66:969-978.

Author contributions

- H. Li: Wrote the manuscript, Conducted the experiment.
- H. H. Yang: Conducted the experiment, Analyzed the data.
- Z. G. Sun: Conducted the experiment, Analyzed the data.
- H. B. Tang: Conducted the experiment, Analyzed the data.
- J. K. Min: Designed the experiment, Analyzed the data.

Funding statement

- This work was supported financially by: Science and Technology Project of Huzhou City (grant number 2016GY26); Zhejiang Provincial Technological Research Project for Public Welfare (grant number 2017C33227); Zhejiang Provincial Natural Science Foundation of China (grant number LY14H060001).
- No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

Ethical review statement

- The ethics committee of the First People's Hospital of Huzhou approved the study (#201710312), and written informed consent was obtained from all subjects.

© 2019 Author(s) et al. This article is distributed under the terms of the Creative Commons Attribution-Non Commercial 4.0 International (CC BY-NC 4.0) licence (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed.