



OPEN ACCESS

## EXTENDED REPORT

# Stratification of knee osteoarthritis: two major patient subgroups identified by genome-wide expression analysis of articular cartilage

Jamie Soul,<sup>1</sup> Sara L Dunn,<sup>1</sup> Sanjay Anand,<sup>2</sup> Ferdinand Serracino-Inglott,<sup>3</sup> Jean-Marc Schwartz,<sup>4</sup> Ray P Boot-Handford,<sup>1</sup> Tim E Hardingham<sup>1</sup>**Handling editor** Josef S Smolen

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/annrheumdis-2017-212603>).

<sup>1</sup>Wellcome Trust Centre for Cell-Matrix Research, Division of Cell-Matrix Biology and Regenerative Medicine, School of Biological Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK  
<sup>2</sup>Department of Orthopaedic Surgery, Stockport NHS Foundation Trust, Stockport, UK  
<sup>3</sup>Department of Vascular Surgery, Central Manchester NHS Foundation Trust, Manchester, UK  
<sup>4</sup>Division of Evolution and Genomic Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK

**Correspondence to**

Professor Tim E Hardingham, Wellcome Trust Centre for Cell-Matrix Research, University of Manchester, Manchester M13 9PT, UK; [tim.hardingham@manchester.ac.uk](mailto:tim.hardingham@manchester.ac.uk)

Received 26 October 2017  
Revised 21 November 2017  
Accepted 28 November 2017  
Published Online First  
22 December 2017



**To cite:** Soul J, Dunn SL, Anand S, et al. *Ann Rheum Dis* 2018;**77**:423–430.

**ABSTRACT**

**Introduction** Osteoarthritis (OA) is a heterogeneous and complex disease. We have used a network biology approach based on genome-wide analysis of gene expression in OA knee cartilage to seek evidence for pathogenic mechanisms that may distinguish different patient subgroups.

**Methods** Results from RNA-Sequencing (RNA-Seq) were collected from intact knee cartilage at total knee replacement from 44 patients with OA, from 16 additional patients with OA and 10 control patients with non-OA. Results were analysed to identify patient subsets and compare major active pathways.

**Results** The RNA-Seq results showed 2692 differentially expressed genes between OA and non-OA. Analysis by unsupervised clustering identified two distinct OA groups: Group A with 24 patients (55%) and Group B with 18 patients (41%). A 10 gene subgroup classifier was validated by RT-qPCR in 16 further patients with OA. Pathway analysis showed increased protein expression in both groups. PhenomeExpress analysis revealed group differences in complement activation, innate immune responses and altered Wnt and TGFβ signalling, but no activation of inflammatory cytokine expression. Both groups showed suppressed circadian regulators and whereas matrix changes in Group A were chondrogenic, in Group B they were non-chondrogenic with changes in mechanoreceptors, calcium signalling, ion channels and in cytoskeletal organisers. The gene expression changes predicted 478 potential biomarkers for detection in synovial fluid to distinguish patients from the two groups.

**Conclusions** Two subgroups of knee OA were identified by network analysis of RNA-Seq data with evidence for the presence of two major pathogenic pathways. This has potential importance as a new basis for the stratification of patients with OA for drug trials and for the development of new targeted treatments.

**INTRODUCTION**

Osteoarthritis (OA) is a complex joint disease with variable aetiology, symptoms and outcome. It is highly prevalent with major risk factors of ageing, joint trauma and obesity. Although showing no simple pattern of Mendelian inheritance,<sup>1</sup> OA is estimated to have ~50% genetic contribution and GWAS studies have revealed multiple low risk genes.<sup>2</sup> OA involves pathological changes in most joint tissues including cartilage, bone, synovium, ligaments and adjunct joint tissues and includes

variable development of osteophytes, subchondral bone cysts and sclerosis.<sup>3–5</sup> The degenerative changes lead to the failure of the joint, which through pain and incapacity causes major morbidity. Progress in the development of disease modifying drugs has been slow and there remains a large unmet clinical need.<sup>6</sup> Stratification on histopathological features has revealed subgroups of OA, but these have not readily given new insights into the molecular processes of the disease or allowed the development of new drugs.<sup>7</sup>

To develop new treatments, it would help to gain more understanding of OA heterogeneity. New omics technologies offer the possibility to search for different pathogenic mechanisms and thereby identify OA subtypes in an unbiased way. We have used genome-wide sequencing to investigate articular cartilage gene expression. Cartilage was selected as it forms the major load-bearing surface in the joint and it is exposed via synovial fluid (SF) to systemic and local factors within the joint. The pattern of gene expression in OA cartilage therefore reflects intrinsic chondrocyte responses and the influence of extrinsic systemic, synovial, biochemical and biomechanical factors. The analysis reported here is on intact OA cartilage and thus avoids the gene expression changes that occur in damaged cartilage, on which we have previously reported.<sup>8</sup>

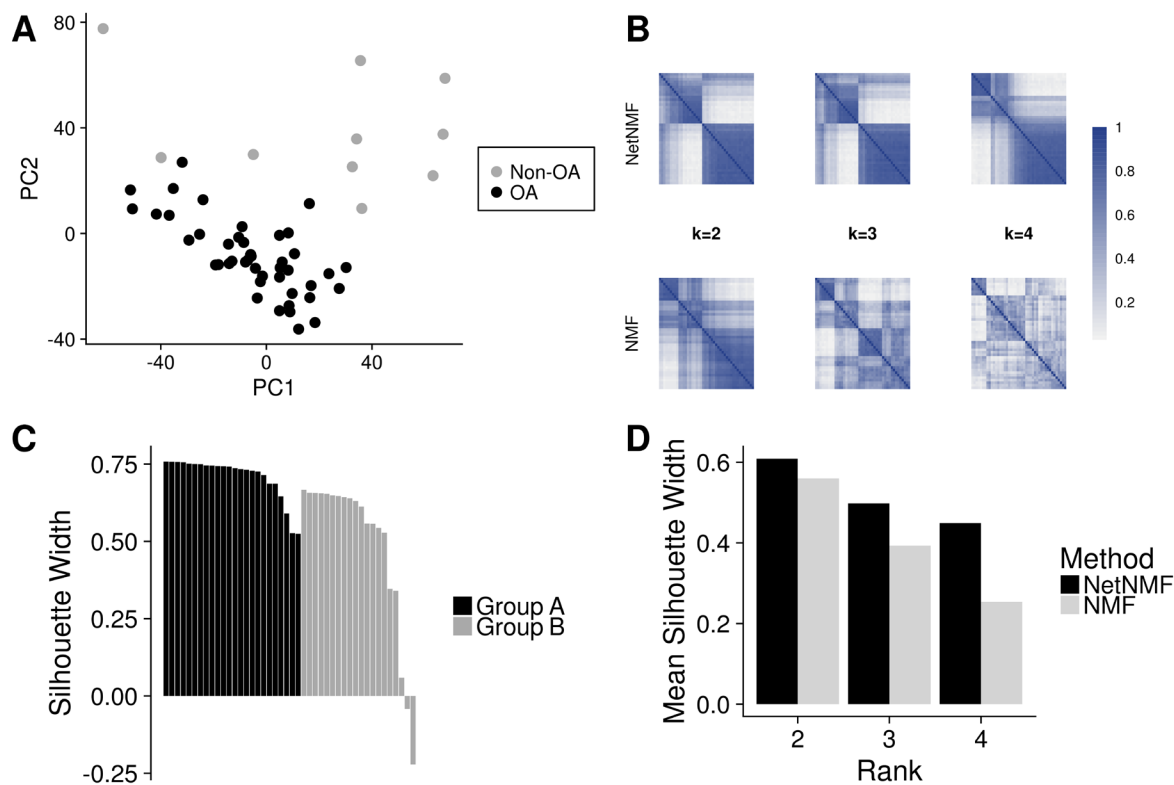
This is the largest scale gene expression analysis of cartilage of patients with OA to date and provides new insight into the heterogeneity of OA.

**METHODS****Study design**

Intact cartilage from the posterior lateral condyle (PLC) was obtained under Ethics Committee approval with prior informed consent at total knee replacement (TKR) from patients with predominant medial compartmental OA and from non-OA age-matched control samples (online supplementary methods). Histology grading (modified Mankin score), DNA and GAG analysis, extraction of RNA and submission for RNA-Sequencing (RNA-Seq) was as previously described (online supplementary methods).<sup>8–10</sup>

**RNA sequencing and analysis**

RNA sequencing, quality control, read mapping and quantification were performed by standard bioinformatic methods.<sup>11 12</sup> Fold changes were calculated with DESeq2 with P values adjusted for



**Figure 1** Unsupervised clustering of gene expression from patients with OA. RNA-Seq data from 44 PLC cartilage samples taken at total knee replacement and 10 non-OA PLC samples were normalised and batch effect corrected. Principal component analysis of the data was performed and the first and second (PC1 and PC2) principal components are shown (A). Network-based NMF or standard NMF was used to cluster the patients with OA into subgroups with variable numbers of predefined clusters ( $k$ ). Patient co-clustering matrices are coloured to show patient pairwise co-occurrence in the same cluster across the 500 clustering runs (B). Silhouette scores were used to assess the optimal number of clusters (C) and the patient fit to a subgroup (D). NMF, negative matrix factorisation; OA, osteoarthritis; PLC, posterior lateral condyle.

multiple testing with Benjamini-Hochberg correction. Differentially expressed genes (DEGs) are reported with an absolute fold change of  $\geq 1.5$  and a false discovery rate of  $\leq 10\%$ . Unsupervised clustering was carried out (online supplementary methods) with cluster stability and patient fit to each subgroup assessed by silhouette score. DEGs were analysed using Reactome pathway and Cytoscape PhenomeScape app (online supplementary methods).

#### Creation of RT-qPCR classifier panel

Shrunken centroid clustering was used (online supplementary methods) to identify the optimum gene set to create a classifier able to distinguish the groups by RT-qPCR.

#### Data availability

Fastq files are available to download from ArrayExpress E-MTAB-6266. Code to reproduce the analysis at <https://github.com/soulj/OAstratification>.

## RESULTS

### Genome-wide expression analysis of intact OA cartilage

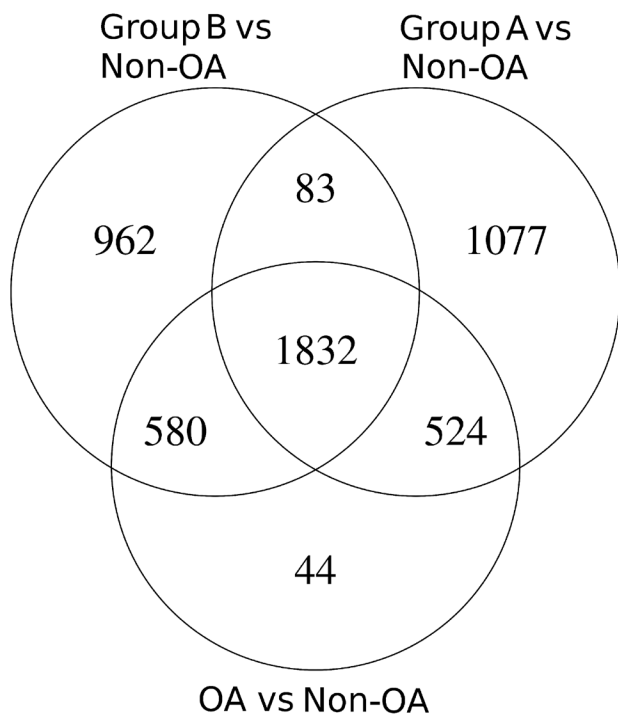
Genome-wide expression by RNA-Seq analysis was determined on intact cartilage tissue sampled at TKR for each of 44 patients with OA. All samples were visually intact, with a mean histological score of less than 10 (modified Mankin) as previously reported.<sup>8</sup> Control age-matched, non-OA cartilage from the same PLC site showed similar modified-Mankin scores, but had slightly higher mean proteoglycan and slightly lower mean DNA

content (online supplementary figure 1). Although the non-OA donors were 9 male/1 female, the female was not an outlier in any of the parameters measured. The gene expression determined (RNA-Seq) from patients with OA and non-OA was thus from age-matched full depth cartilage.

The RNA-Seq data were normalised and batch corrected using standard bioinformatic tools (see 'Methods' section). After normalisation, principal component analysis of the samples showed some heterogeneity (figure 1A), but the non-OA samples were distinct from the patients with OA and comparison of the results from the 44 patients with OA with the 10 non-OA controls identified 2692 DEGs (online supplementary table 1). The intact OA cartilage from the 44 patients thus showed a broad range of changes in gene expression compared with control non-OA cartilage. These changes also showed major overlap with previous microarray comparisons of expression in OA and non-OA cartilage.<sup>13–15</sup>

### Stratification of gene expression

The gene expression results from the 44 patients with OA were analysed to search for possible subsets that shared selective changes in expression. Using an unbiased network-assisted, non-negative matrix factorisation (Network NMF) algorithm, which has been demonstrated to yield biologically informative disease clusters, the most stable clustering identified two subgroups (figure 1B and D).<sup>16 17</sup> Measure of the degree of fit (Silhouette score, see online supplementary methods) revealed strong identity for most patients within each cluster (figure 1C).



**Figure 2** Differentially expressed genes in OA Group A and Group B. Differentially expressed genes ( $\geq 1.5$  FC and  $\leq 0.1$  adjusted P values) were determined for each OA Group A (n=32) and Group B (n=26) compared with patients with non-OA (n=10). OA, osteoarthritis.

Only two patients showed no concordance with either group and these were not included in further analysis. The RNA-Seq results thus identified that patients were in two main groups: Group A 55% (n=24) and Group B 41% (n=18) and these showed no association with histological grade, sex, age, body mass index or pain scores (VAS, KOOS) (online supplementary figure 2).

### Classification of new patient samples

We next sought to establish if we could identify each group by RT-qPCR analysis of the expression of a small panel of genes. Using a shrunken centroid classifier approach, we identified an optimal set of 10 genes, whose expression discriminated between the groups. The gene selection predicted from RNA-Seq was validated by RT-qPCR analysis (online supplementary figure 3A,B) and used to train a SVM classifier with fivefold cross-validation using the known group identities. The expression pattern of these 10 genes was conserved in a validation cohort of 16 patients (online supplementary figure 3C) and the RT-qPCR classifier showed an area under the curve of 0.92 in predicting the patient groups compared with the full RNA-Seq analysis.

### Pathways activated in Group A and Group B relative to non-OA

In the full gene expression analysis from 60 patients, there were 2980 genes that were differentially expressed (DEGs) compared with non-OA (online supplementary table 2). However, when the results were divided into Group A and Group B and each compared separately to non-OA, there were an additional 2122 DEGs detected, with 1077 of those distinctive to Group A and 962 distinctive to Group B (figure 2 and online supplementary table 3). Splitting the patients into two groups thus identified many more gene changes specific to OA.

For the initial analysis, we used an unbiased Reactome pathway approach to compare DEGs in each group with

**Table 1** Reactome pathway analysis

Pathway	Adj. P value
<b>Group A vs non-OA</b>	
Eukaryotic translation elongation	6.2e-63
Extracellular matrix organisation	7.3e-13
Chemokine receptors bind chemokines	0.00615
MyD88 deficiency (TLR2/4)	0.04693
Glycosaminoglycan metabolism	0.00584
Inflammasomes	0.01242
<b>Group B vs non-OA</b>	
Eukaryotic translation elongation	1.6e-69
Extracellular matrix organisation	4.3e-12
Negative regulation of TCF-dependent signalling by WNT ligand antagonists	0.00707
Regulation of complement cascade	0.00758
Eicosanoid ligand-binding receptors	0.00758
Syndecan interactions	0.01795
<b>Group B vs Group A</b>	
Extracellular matrix organisation	9.2e-11
Chemokine receptors bind chemokines	8.9e-08
Complement cascade	0.00012
G alpha (i) signalling events	0.00036
Adaptive immune system	0.00068

A selection of highly significant differentially regulated Reactome pathways between OA Group A (n=32) and Group B (n=26) and non-OA (n=10). OA, osteoarthritis.

non-OA cartilage (online supplementary table 4). This identified in Group A changes in chemokine signalling, inflammasome activation, changes in glycosaminoglycan synthesis, toll receptor activation and innate immune responses, whereas Group B included complement regulation, Wnt signalling, eicosanoid receptor signalling and syndecan interactions (table 1).

Reactome pathways also showed that a major feature common to both groups was multiple changes in ribosomal protein expression (table 1 and online supplementary table 4). This may have two core origins: first, reflecting a large increase in protein expression in OA chondrocytes as part of a disease response<sup>16</sup> and second, a change in chondrocyte differentiation that results in quite distinct patterns of altered gene expression.<sup>18</sup>

The second major change was in matrix protein genes that were highly expressed. In the 100 most expressed genes in non-OA cartilage, there were 27 matrix genes, but this increased to 37 matrix genes in Group A and 39 in Group B. The results showed a large increase in matrix protein gene expression in both patient groups, but as detailed below each group showed a different pattern of changes. Several growth factors with anabolic effects (IGF2, FGF1, 2 and 18, TGF $\beta$ 1, 3 and GDF5, 10 and BMP6) were increased in both groups, potentially supporting the increased and altered pattern of protein expression (online supplementary table 3). A list of the major pathways common to both groups compared with non-OA is shown in online supplementary table 4.

### Group A and Group B show altered patterns of matrix protein gene expression

Group B showed changes reflecting a switch to less chondrogenic genes with reduced expression of the major chondrocyte transcription factor, SOX9 and SIRT1, which is a supporter of chondrogenesis and reduced expression of the matrix genes, aggrecan (ACAN) and type IX collagen (COL9A2/3)<sup>19,20</sup> (online

supplementary table 3 and 5). In contrast the non-chondrogenic genes versican (VCAN) and Type I collagen (COL1A1/2) were increased together with the pericellular matrix collagens, (COL6A1, 2, 3), laminins (LAMA2, A4, B1, C1) and perlecan (HSPG2) (supplementary table 3 and 5).<sup>21</sup> An altered phenotype in Group B is also suggested by increased expression of transcription factors associated with an osteogenic phenotype, including osterix and RUNX2 (online supplementary table 5). Together these results suggest the chondrocytes in Group B patients are active in a less chondrogenic pathway of remodelling. In contrast, Group A showed increased expression of major cartilage collagens Type II, Type V and Type IX and Type XI and less expression of Type I (online supplementary table 3).

### Pathway analysis comparing Group A with Group B

To investigate further the differences between the two groups, we applied a network-based PhenomeExpress analysis incorporating data from published disease-related studies.<sup>22 23</sup> This technique identifies gene-sets that have known direct molecular interactions and scores them based on the total differential expression of the gene-set. Activities distinguishing the two groups included oxidative stress, innate-immune responses, Wnt signalling, chemokine signalling, apoptotic threshold and calcium regulation (figure 3; online supplementary figure 3 and supplementary table 6). The heatmaps of this analysis shows downregulation of genes in many pathways in Group A compared with non-OA. Where such pathways are protective to cartilage homeostasis, Group A could be more susceptible to other adverse changes that may be metabolic and/or mechanical and caused by extrinsic/environmental or intrinsic/genetic factors.

### Changes in circadian rhythm and TGF $\beta$ signalling

Circadian clock mechanisms have recently been identified as an important pathway in regulating cartilage matrix homeostasis. As the circadian regulation of genes controls a 24 hours cycle of coordinated matrix protein synthesis and degradation,<sup>24</sup> any change in amplitude or phase may contribute to a loss of matrix integrity. Circadian disruption by genetic deletion of the regulator BMAL1 in mouse cartilage results in an early onset joint degeneration.<sup>20 24</sup> In our expression results, BMAL1 (ARNTL) and its binding partner NPAS2 were reduced in both Groups compared with non-OA, whereas in contrast, CRY1/2 (members of the circadian oscillator complex) were increased (table 2A). This change in ratio suggests an imbalance in clock control and although the timing of cartilage sampling was not strictly controlled, it was generally early in the same quarter (09:00–15:00 hours) of the 24 hours cycle and the ratio is less variable than the absolute levels of expression (24). There was also reduced expression of other genes involved in circadian regulation (table 2A) including the transcriptional repressor Rev-ErbA (NR1D1).<sup>25</sup> Consequences of the loss of circadian rhythm involve the downregulation of transcriptional regulators such as NFATC1/2.<sup>26 27</sup> In our data, the downstream targets NFATC1/2 were increased in expression in both groups, which suggests their expression is maintained independent of the circadian changes. Clock disruption may act through other major pathways and TGF $\beta$  signalling was shown to be affected by BMAL or NR1D1 knockdown in cultured human articular chondrocytes, which resulted in altered SMAD signalling and increased TGF $\beta$ 3 and the receptor TGFBR1 expression.<sup>26 27</sup> As these were increased in expression in both patient groups, upregulated activity in TGF $\beta$  signalling pathways may account for the increased expression

of target genes, including cartilaginous and non-cartilaginous matrix proteins (table 2).<sup>28</sup>

### Mechanical load-related changes

Joint cartilage has a prime load bearing role. Repetitive loading is necessary for maintaining cartilage function and whereas excess overload can cause damage, joint disuse also leads to atrophic changes. Chondrocytes therefore have mechanosensitive homeostatic mechanisms that regulate the assembly and maintenance of cartilage matrix.<sup>29</sup>

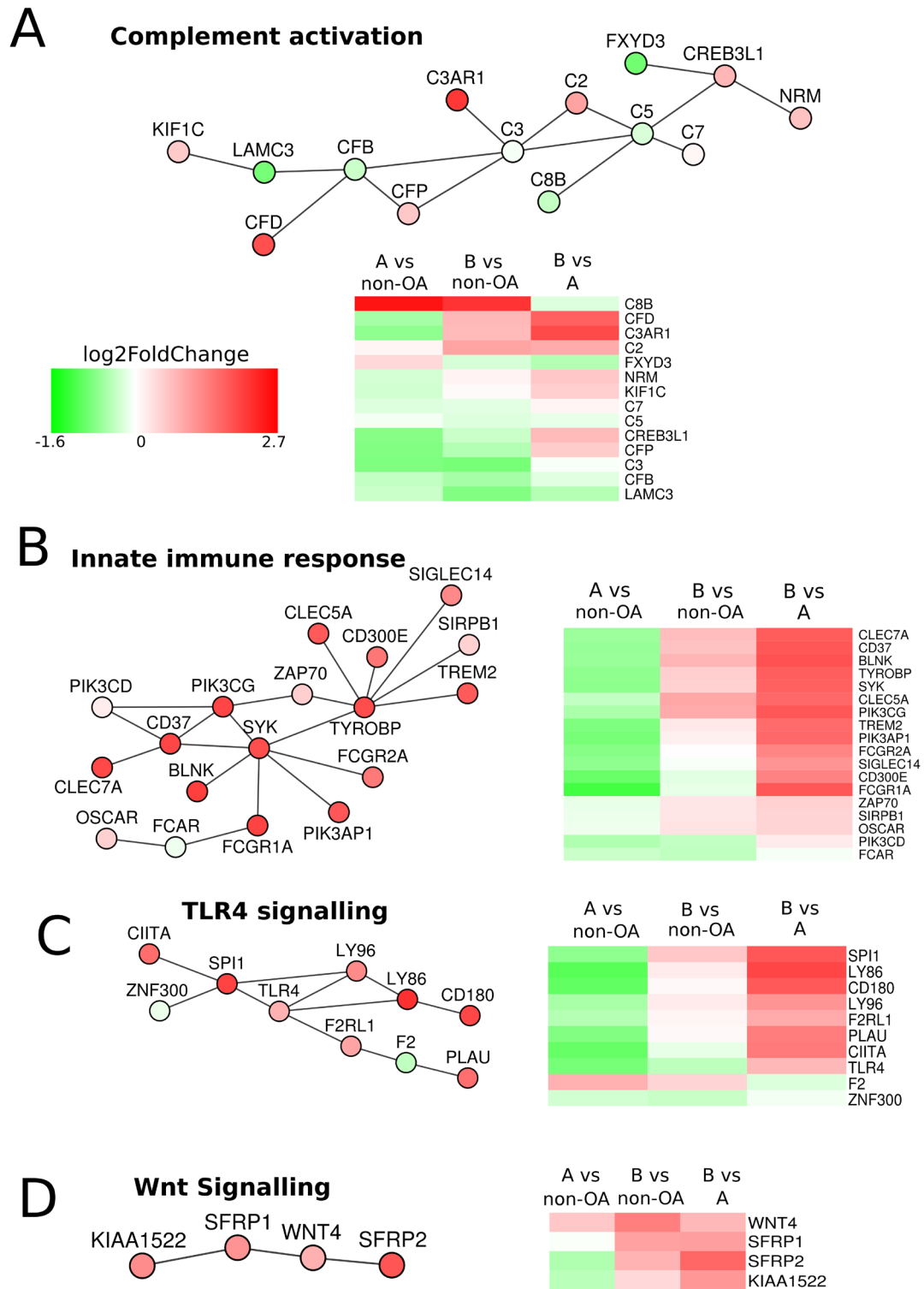
The increases in less chondrogenic genes in Group B may reflect a response to a changed pattern of mechanical stress across the joint causing altered loading on chondrocytes.<sup>30 31</sup> Mechanosensors, ion channels, calcium control, water balance and cytoskeletal tensing are all implicit in chondrocyte load responses,<sup>29 31 32</sup> and these show a spectrum of changes, which are more pronounced in Group B. The water channel AQP1 and the main calcium regulated ion channel in chondrocytes TRPV4 are strongly increased in Group B.<sup>32 33</sup> Also increased in Group B are other calcium-activated potassium channels including KCNN3, 4,<sup>34</sup> the voltage sensitive calcium activated chloride channel ANO1 and potassium channels KCNQ5, KCNS3 and KCTD12, KCNMB4 (table 2B). Cytoskeletal proteins in Group B also showed increased alpha smooth muscle actin (ACTA2) and increases in contractile proteins, scaffold and membrane proteins, which together suggest chondrocytes are exposed to greater shear stress and deformation (table 2B).<sup>35</sup> Group A also showed increased expression of TRPV4, but low or no increase in these other ion channels and cytoskeletal proteins.

### Inflammation and proteases

Inflammation has been proposed as a contributor to OA progression.<sup>36–38</sup> Although the gene expression reported here is from cartilage chondrocytes, the pattern of expression would include responses to inflammatory mediators from surrounding tissues and SF that penetrate cartilage. However, in the intact OA cartilage sampled in both groups exhibited very low expression and no differential expression of IL1 $\alpha$ ,  $\beta$ , TNF, IL6 and OSM and also of CCL2 (MCP-1), CXCL9 (MIG) and CXCL10 (IP10) (online supplementary table 3). Furthermore, the alarmins S100A8/9 (MRP8/14), which are linked to activation of the inflammatory cascade,<sup>39</sup> were also less expressed in both OA groups. The expression of NF $\kappa$ B complex proteins, such as the kinase activator (IKK-  $\beta$ ) and the inhibitory binding partner I $\kappa$ B- $\beta$  were less expressed in both groups and IKK- $\gamma$  (NF $\kappa$ B essential modulator) was little changed compared with non-OA. There was thus little evidence of classic inflammatory pathway activation in the intact OA cartilage sampled or of activation by factors from SF.

A range of inflammatory cytokines have been detected in OA SF, although only at low concentration (0–20 pg/mL).<sup>40 41</sup> As several growth factors are increased in expression in both OA groups (see above) with the potential to reverse the response of chondrocytes to inflammatory cytokines,<sup>42 43</sup> it appears likely that any potential effect of cytokines at the concentrations present in SF on chondrocyte gene expression may be overcome by endogenous growth factors. As knee OA is a slowly progressive condition, more evidence is needed from earlier disease stages to show when inflammatory mediators are expressed in chondrocytes and/or are present in SF at high enough concentrations to contribute to cartilage pathology.

The expression of matrix proteinases showed that a mixed pattern of responses with gelatinase (MMP9) and collagenase



**Figure 3** PhenoExpress analysis of OA Group A and Group B. Differential gene expression data between Group A (n=32) and Group B (n=26) were analysed with the PhenoExpress algorithm to identify dysregulated subnetworks using known disease gene associations. Nodes are coloured by fold change. Heatmaps show the gene expression between Group A, Group B and non-OA for each subnetwork.

(MMP13) both increased and more strongly so in Group B, whereas stromelysin (MMP3) and the aggrecanases (ADAMTS4, 5) were downregulated in both groups. The only matrix

proteinase highly expressed by chondrocytes was the serine proteinase HTRA1 which was increased in both groups (online supplementary table 3).

**Table 2** Expression changes in circadian clock related genes in patient Groups A and B; 2B—expression changes in mechanosensitive ion channels and cytoskeletal proteins in patient Groups A and B

Name	Group A fold change	Group B fold change	
<b>(A) Gene</b>			
ARNTL	Bmal1 primary clock oscillator	-1.67	-1.88
NPAS2	Transcription factor activator interacts with Bmal1	-1.64	-1.85
CRY1	Inhibitor of Bmal1/NPAS2 as part of clock oscillator	+1.33	+1.37
PER1	Inhibitor of the core clock oscillator	-1.50	-1.87
NR1D1	Rev-Erb $\alpha$ nuclear hormone receptor, negative regulator of Bmal1 transcription	-1.72	-1.90
BHLHE40	DEC1, Basic helix-loop-helix Target of clock	NS	-1.35
RORC	Nuclear hormone receptor transcr. factor, positive regulator of Bmal1 transcription	-1.40	-2.15
DBP	Transcription factor activating clock target genes via D-box cis elements	+2.39	+1.95
NFIL3	Transcription factor suppressing clock target genes via D-box cis elements	-2.29	-2.67
NFATC1	Component of DNA binding transcription complex	+1.49	+1.30
NFATC2	Component of DNA binding transcription complex	+3.02	+3.33
<b>(B) Ion channels</b>			
TRPV4	Mechanosensitive Ca <sup>++</sup> channel	+2.28	+1.59
AQP1	Aquaporin, water channel	NS	+3.28
ANO1	Voltage gated Ca <sup>++</sup> activated Cl <sup>-</sup> channel	+2.27	+3.15
KCNA1	K <sup>+</sup> voltage gated channel	+1.61	+1.63
KCNQ5	K <sup>+</sup> voltage gated channel	+1.53	+1.57
KCNS3	K <sup>+</sup> voltage gated channel modifier	NS	+2.53
KCNK6	K <sup>+</sup> channel	NS	+1.31
KCNN3	Ca <sup>++</sup> activated K <sup>+</sup> channel	+1.71	+2.64
KCNN4	Ca <sup>++</sup> activated K <sup>+</sup> channel	NS	+2.86
KCNMB4	Ca <sup>++</sup> activated K <sup>+</sup> channel subunit 4	NS	+1.58
KCTD12	K <sup>+</sup> channel tetramerisation domain containing 12	NS	+1.96
SCNNIA	Non-gated Na <sup>+</sup> channel	-3.46	-4.94
<b>Cytoskeleton</b>			
ACTA2	Alpha smooth muscle actin	-2.86	+1.81
VIN	Vinculin cytoskeletal protein	NS	+1.37
TMSB4X	Thymosin B4 cytoskeletal protein	NS	+1.87
EZR	Ezrin-Plasma mem-cytoskeletal link	NS	+2.08
TPM4	Tropomyosin 4 actin binding	NS	+1.36
MYH9/11	Smooth muscle myosin heavy chain, contractile protein	9—NS 11—NS	+1.32 +2.30

GWAS, genome-wide association study.

### Integration with other studies

In a previous knee OA study, Fernández-Tajes *et al* reported DNA methylation analysis on cartilage from 25 patients and identified a small subgroup that differed from the rest. Microarray analysis on a separate group of patients identified 47 significantly altered genes that distinguished the small group and were directly correlated with the methylation analysis.<sup>44</sup> Selecting 9 of the 47 genes, they showed by RT-qPCR these could identify the small subgroup. Comparison of the expression results of these 47 genes with our RNA-Seq analysis showed major overlap (online supplementary table 7) and showed that their panel of nine genes were also all differentially expressed between our Group A and Group B (see supplementary figure 4). These results suggest that the Group B identified in our study corresponds to the small group identified using an entirely different approach by Fernández-Tajes *et al*.<sup>44</sup>

From their results, Fernández-Tajes *et al* concluded that this small subset showed evidence of inflammation and reduced matrix protein synthesis. A further analysis based on GO annotations of gene expression identified inflammatory response, leucocyte activation, cytokine production and chemokine activity, as factors more enhanced in the subset. Our conclusions

from Reactome pathway analysis detected some similar changes comparing Group B with Group A (table 1); however, from the more detailed analysis possible with genome-wide expression data across 60 patients, we find strong evidence of increased matrix protein expression in both our patient groups and little evidence of classical inflammatory activation. Our analysis thus agrees with the group identification of Fernández-Tajes *et al*,<sup>44</sup> but differs in the conclusions from the pathway analysis.

### Potential biomarkers in synovial fluid

From our gene expression analysis, it was possible to predict 478 secreted proteins (online supplementary table 8) differentially expressed between the two groups that may be released from the articular cartilage and therefore detectable in SF. Many of the 478 secreted proteins are not specific to cartilage and may enter SF from other joint tissues and from the circulation. The differences in their gene expression may also not correlate well with their concentration in SF. However, among these proteins, some may form potential biomarkers that could enable the identity of the two groups to be determined by SF analysis and leading candidates include several proteins already detected in SF and

**Table 3** Chondrocyte secreted proteins differentially expressed in Group A and Group B

Gene	Protein	Group A fold increase	Group B fold increase
WIF1	Wnt inhibitor factor 1	1.82	
S100A1	S100 Ca <sup>++</sup> binding protein	1.77	
S100B	S100 Ca <sup>++</sup> binding protein	1.72	
C4BPA	C4 binding protein A	1.73	
GPX3	Glutathione peroxidase 3	1.62	
ADAMTS2	AdamTS-like 2, microfibril assembly	1.55	
SCRG1	Stimulator of chondrogenesis	1.48	
CHAD	Chondroadherin	1.47	
MGP	Matrix Gla protein	1.47	
SERPINF1	Pigment epithelium derived factor		3.60
RNASE1	Pancreatic style secretory RNase		3.56
SPARCL1	Sparc-like protein 1 (hevin)		3.11
POSTN	Periostin		2.79
PLTP	Phospholipid transfer protein		2.26
IGFBP4	IGF binding protein 4		2.23
SFRP1	Secreted Frizzled related protein 1		2.19
CLEC3B	Tetranectin		2.07
TGFBI	TGF beta induced protein		2.03
DKK3	Dickkopf 3, Wnt inhibitor		2.00
TNC	Tenascin C		1.80
TIMP3	Tissue inhibitor of metalloproteinases 3		1.67
CRTAC1	Cartilage acidic protein 1		1.55
ASPN	Asporin (serine leucine-rich matrix protein)		1.47

cartilage explant culture media (table 3). Large scale proteomic analysis of OA SF correlated with the differential expression in patients' cartilage of the classifier set of 10 genes may enable a select group of proteins to be identified whose combination of concentrations in SF distinguishes patients in the two groups. This would make more accessible the use of these results to stratify patients with OA.

## DISCUSSION

A major finding of this study was that the pattern of gene expression in knee OA cartilage falls into two clear groups. The cartilage focused on is from an undamaged site. The results do not therefore reflect different degrees of tissue damage, but are intrinsic to OA cartilage in the two groups. The design of this study also offers an advance over previous microarray comparisons of OA and non-OA cartilage, by using RNA-seq with more complete genome wide coverage and greater dynamic range and by carrying out the study on a scale large enough to assess heterogeneity among patients. The results showed little evidence of induced inflammatory responses in each of 60 patients, which suggests that at this late stage of disease, the effect of local inflammation is not a common or persistent factor. This is largely in agreement with other studies that have noted little evidence for inflammatory cytokine activity in late OA.<sup>40 45</sup>

The common changes in circadian regulators in both OA groups suggest that a decline in clock regulation may be an important factor that adversely affects the fundamental mechanisms of cartilage homeostasis and, in the long term, this may compromise tissue integrity. More detailed studies are needed to substantiate this and to understand the circadian impact on other major biological and mechanical signalling pathways and

how they relate to the major increases and different patterns of matrix protein gene expression. A possible cause for the matrix protein response in Group B being less chondrogenic is suggested by a whole set of gene expression changes linked to altered mechanical demands, such as increased tissue shear. Thus, there was increased expression of key mechanoreceptors, in calcium signalling, in associated ion channels, in cytoskeletal organisers, in cell matrix interface and pericellular proteins, which all have links to mechanotransduction. The altered mechanical environment may result in a switch in chondrocyte phenotype, which results in a less chondrogenic response. The different pattern of matrix expression in the two groups would predict altered cartilage matrix properties, but it will require measurements of tensile and compressive stiffness to determine the different biomechanical changes in the two groups and to show if these changes predict different rates of progression of cartilage damage.<sup>30 46</sup>

Many failed clinical trials in knee OA have possibly benefited some patients, but had poor success overall and have not delivered clear benefit above placebo within the time frame of the trial.<sup>47–49</sup> The results in this study identify a metabolic basis for stratifying patients into two groups, which are based on major differences in gene expression. While these changes are specific to chondrocytes in cartilage they may also reflect responses to mechanical signals and biological mediators from across the joint. This new route to OA stratification is potentially important as it suggests that the different patterns of gene expression in the two groups reflect different metabolic activities and raises the possibility that the responses of the two groups may be susceptible to different modes of intervention.<sup>47</sup> To realise any practical benefit from this analysis requires further validation and testing. The identification of the classifier panel of genes and potential biomarkers in SF provide tools to facilitate progress, but new investigations are needed to show any predictive value and establish if stratifying patients on this molecular basis can help distinguish responders and non-responders and enable preselection of patients that respond more consistently to new targeted treatments.

**Acknowledgements** We thank Qing-Jun Meng (Wellcome Trust Centre for Cell-Matrix Research, University of Manchester) for helpful discussion on the circadian regulation of genes in cartilage. We thank Professor Sally Roberts, Robert Hunt and Agnes Jones Orthopaedic Hospital, Oswestry, UK for providing samples of tissue from a non-OA knee under local ethical approval. We would like to acknowledge the University of Manchester Genomic Technologies Core Facility for RNA-sequencing and Ping Wang (Bioinformatics Core Facility, The University of Manchester) for the initial RNA-seq data quality assurance.

**Contributors** TEH and RB-H designed the study. SD carried out sample collection, characterisation and RNA prep for analysis. JS and J-MS developed the analysis of the data. SA and FS-I provided clinical support for patient interaction and tissue provision. TEH, RB-H and SL drafted the paper, which was reviewed and approved by all coauthors.

**Funding** This work was supported by grants from Arthritis Research UK (20414) and the EU (SYBIL—European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 602300); The Wellcome Trust Centre for Cell-Matrix Research, University of Manchester, is supported by core funding from the Wellcome Trust (grant number 203128/Z/16/Z).

**Competing interests** None declared.

**Ethics approval** IRAS 114697.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** All data from the study are available by Public Access within the text or in online supplementary files. All RNA-seq and metadata are available at ArrayExpress and code is available at GitHub.

**Open Access** This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use,

provided the original work is properly cited. See: <http://creativecommons.org/licenses/by/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

## REFERENCES

- Ball J, Carrington MJ, McMurray JJ, *et al.* Atrial fibrillation: profile and burden of an evolving epidemic in the 21st century. *Int J Cardiol* 2013;167:1807–24.
- Lip GYH, Brechin CM, Lane DA. The global burden of atrial fibrillation and stroke: a systematic review of the epidemiology of atrial fibrillation in regions outside North America and Europe. *Chest* 2012;142:1489–98.
- Loeser RF. Aging and osteoarthritis. *Curr Opin Rheumatol* 2011;23:492–6.
- Loeser RF, Goldring SR, Scanzello CR, *et al.* Osteoarthritis: a disease of the joint as an organ. *Arthritis Rheum* 2012;64:1697–707.
- Loitz M, Loeser RF. Effects of aging on articular cartilage homeostasis. *Bone* 2012;51:241–8.
- Goldring MB. Update on the biology of the chondrocyte and new approaches to treating cartilage diseases. *Best Pract Res Clin Rheumatol* 2006;20:1003–25.
- Waarsing JH, Bierma-Zeinstra SM, Weinans H. Distinct subtypes of knee osteoarthritis: data from the Osteoarthritis Initiative. *Rheumatology* 2015;54:1650–8.
- 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;24:1431–40.
- Hoemann CD. Molecular and biochemical assays of cartilage components. *Methods Mol Med* 2004;101:127–56.
- Kim YJ, Sah RL, Doong JY, *et al.* Fluorometric assay of DNA in cartilage explants using Hoechst 33258. *Anal Biochem* 1988;174:168–76.
- Bray NL, Pimentel H, Melsted P, *et al.* Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 2016;34:525–7.
- Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res* 2015;4:1521.
- 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;54:3533–44.
- Sato T, Konomi K, Yamasaki S, *et al.* Comparative analysis of gene expression profiles in intact and damaged regions of human osteoarthritic cartilage. *Arthritis Rheum* 2006;54:808–17.
- Karlsson C, Dehne T, Lindahl A, *et al.* Genome-wide expression profiling reveals new candidate genes associated with osteoarthritis. *Osteoarthritis Cartilage* 2010;18:581–92.
- Hofree M, Shen JP, Carter H, *et al.* Network-based stratification of tumor mutations. *Nat Methods* 2013;10:1108–15.
- Cai D, He X, Wu X, *et al.* 2008. Non-negative Matrix Factorization on Manifold. International Conference on Data Mining 63–72.
- Katsara O, Attur M, Ruoff R, *et al.* Increased activity of the chondrocyte translational apparatus accompanies osteoarthritic changes in human and rodent knee cartilage. *Arthritis Rheumatol* 2017;69:586–97.
- Brew CJ, Clegg PD, Boot-Handford RP, *et al.* Gene expression in human chondrocytes in late osteoarthritis is changed in both fibrillated and intact cartilage without evidence of generalised chondrocyte hypertrophy. *Ann Rheum Dis* 2010;69:234–40.
- Yang W, Kang X, Liu J, *et al.* Clock Gene Bmal1 Modulates Human Cartilage Gene Expression by Crosstalk With Sirt1. *Endocrinology* 2016;157:3096–107.
- Wilusz RE, Sanchez-Adams J, Guilak F. The structure and function of the pericellular matrix of articular cartilage. *Matrix Biol* 2014;39:25–32.
- Soul J, Hardingham TE, Boot-Handford RP, *et al.* Phenomeexpress: a refined network analysis of expression datasets by inclusion of known disease phenotypes. *Sci Rep* 2015;5:8117.
- Soul J, Dunn SL, Hardingham TE, *et al.* PhenomeScape: a cytoscape app to identify differentially regulated sub-networks using known disease associations. *Bioinformatics* 2016;32:3847–9.
- Yang N, Meng QJ. Circadian clocks in articular cartilage and bone: a compass in the sea of matrices. *J Biol Rhythms* 2016;31:415–27.
- Dudek M, Gossan N, Yang N, *et al.* The chondrocyte clock gene Bmal1 controls cartilage homeostasis and integrity. *J Clin Invest* 2016;126:365–76.
- Akagi R, Akatsu Y, Fisch KM, *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.
- Greenblatt MB, Ritter SY, Wright J, *et al.* NFATc1 and NFATc2 repress spontaneous osteoarthritis. *Proc Natl Acad Sci U S A* 2013;110:19914–9.
- van der Kraan PM, Blaney Davidson EN, Blom A, *et al.* TGF-beta signaling in chondrocyte terminal differentiation and osteoarthritis: modulation and integration of signaling pathways through receptor-Smads. *Osteoarthritis Cartilage* 2009;17:1539–45.
- Sanchez-Adams J, Leddy HA, McNulty AL, *et al.* The mechanobiology of articular cartilage: bearing the burden of osteoarthritis. *Curr Rheumatol Rep* 2014;16:451.
- He B, Wu JP, Chen HH, *et al.* Elastin fibers display a versatile microfibril network in articular cartilage depending on the mechanical microenvironments. *J Orthop Res* 2013;31:1345–53.
- Clark AL, Votta BJ, Kumar S, *et al.* Chondroprotective role of the osmotically sensitive ion channel transient receptor potential vanilloid 4: age- and sex-dependent progression of osteoarthritis in Trpv4-deficient mice. *Arthritis Rheum* 2010;62:2973–83.
- O'Connor CJ, Leddy HA, Benefield HC, *et al.* TRPV4-mediated mechanotransduction regulates the metabolic response of chondrocytes to dynamic loading. *Proc Natl Acad Sci U S A* 2014;111:1316–21.
- O'Connor CJ, Ramalingam S, Zelenski NA, *et al.* Cartilage-specific knockout of the mechanosensory ion channel trpv4 decreases age-related osteoarthritis. *Sci Rep* 2016;6:29053.
- Matta C, Zákány R, Mobasheri A. Voltage-dependent calcium channels in chondrocytes: roles in health and disease. *Curr Rheumatol Rep* 2015;17:43.
- Bleuel J, Zaucke F, Brüggemann GP, *et al.* Effects of cyclic tensile strain on chondrocyte metabolism: a systematic review. *PLoS One* 2015;10:e0119816.
- Goldring MB, Otero M. Inflammation in osteoarthritis. *Curr Opin Rheumatol* 2011;23:471–8.
- Houard X, Goldring MB, Berenbaum F. Homeostatic mechanisms in articular cartilage and role of inflammation in osteoarthritis. *Curr Rheumatol Rep* 2013;15:375.
- Liu-Bryan R, Terkeltaub R. Emerging regulators of the inflammatory process in osteoarthritis. *Nat Rev Rheumatol* 2015;11:35–44.
- Vogl T, Tenbrock K, Ludwig S, *et al.* Mrp8 and Mrp14 are endogenous activators of Toll-like receptor 4, promoting lethal, endotoxin-induced shock. *Nat Med* 2007;13:1042–9.
- Tsuchida AI, Beekhuizen M, Hart MC, *et al.* Cytokine profiles in the joint depend on pathology, but are different between synovial fluid, cartilage tissue and cultured chondrocytes. *Arthritis Res Ther* 2014;16:441.
- Sohn DH, Sokolove J, Sharpe O, *et al.* Plasma proteins present in osteoarthritic synovial fluid can stimulate cytokine production via Toll-like receptor 4. *Arthritis Res Ther* 2012;14:R7.
- Tyler JA. Insulin-like growth factor 1 can decrease degradation and promote synthesis of proteoglycan in cartilage exposed to cytokines. *Biochem J* 1989;260:543–8.
- Rayan V, Hardingham T. The recovery of articular cartilage in explant culture from interleukin-1 alpha: effects on proteoglycan synthesis and degradation. *Matrix Biol* 1994;14:263–71.
- Fernández-Tajes J, Soto-Hermida A, Vázquez-Mosquera ME, *et al.* Genome-wide DNA methylation analysis of articular chondrocytes reveals a cluster of osteoarthritic patients. *Ann Rheum Dis* 2014;73:668–77.
- Sandy JD, Chan DD, Trevino RL, *et al.* Human genome-wide expression analysis reorients the study of inflammatory mediators and biomechanics in osteoarthritis. *Osteoarthritis Cartilage* 2015;23:1939–45.
- Robinson DL, Kersh ME, Walsh NC, *et al.* Mechanical properties of normal and osteoarthritic human articular cartilage. *J Mech Behav Biomed Mater* 2016;61:96–109.
- Chevalier X, Eymard F, Richette P. Biologic agents in osteoarthritis: hopes and disappointments. *Nat Rev Rheumatol* 2013;9:400–10.
- Guermazi A, Roemer FW, Felson DT, *et al.* Motion for debate: osteoarthritis clinical trials have not identified efficacious therapies because traditional imaging outcome measures are inadequate. *Arthritis Rheum* 2013;65:2748–58.
- Hardingham T. Extending the paradigm of osteoarthritis: comment on the editorial by McAlindon. *Arthritis Rheumatol* 2016;68:772–3.