Osteoarthritis and Cartilage xxx (2016) 1–10

Osteoarthritis and Cartilage

OARSI OSTEOARTHRITIS RESEARCH SOCIETY INTERNATIONAL

Gene expression changes in damaged osteoarthritic cartilage identify a signature of non-chondrogenic and mechanical responses

S.L. Dunn †, J. Soul †, S. Anand ‡, J.-M. Schwartz †, R.P. Boot-Handford † **, T.E. Hardingham † *

† *Wellcome Trust Centre for Cell-Matrix Research, University of Manchester, UK*
‡ *Stockport NHS Foundation Trust, Manchester, UK*

ARTICLE INFO

Article history:
Received 16 September 2015
Accepted 4 March 2016

Keywords:
Cartilage
Osteoarthritis
RNA-seq
Systems biology
PhenomeExpress

SUMMARY

Objectives: Joint degeneration in osteoarthritis (OA) is characterised by damage and loss of articular cartilage. The pattern of loss is consistent with damage occurring only where the mechanical loading is high. We have investigated using RNA-sequencing (RNA-seq) and systems analyses the changes that occur in damaged OA cartilage by comparing it with intact cartilage from the same joint.
Methods: Cartilage was obtained from eight OA patients undergoing total knee replacement. RNA was extracted from cartilage on the damaged distal medial condyle (DMC) and the intact posterior lateral condyle (PLC). RNA-seq was performed to identify differentially expressed genes (DEGs) and systems analyses applied to identify dysregulated pathways.
Results: In the damaged OA cartilage, there was decreased expression of chondrogenic genes SOX9, SOX6, COL11A2, COL9A1/2/3, ACAN and HAPLN1; increases in non-chondrogenic genes COL1A1, COMP and FN1; an altered pattern of secreted proteinase expression; but no expression of major inflammatory cytokines. Systems analyses by PhenomeExpress revealed significant sub-networks of DEGs including mitotic cell cycle, Wnt signalling, apoptosis and matrix organisation that were influenced by a core of altered transcription factors (TFs), FOSL1, AHR, E2F1 and FOXM1.
Conclusions: Gene expression changes in damaged cartilage suggested a signature non-chondrogenic response of altered matrix protein and secreted proteinase expression. There was evidence of a damage response in this late OA cartilage, which surprisingly showed features detected experimentally in the early response of cartilage to mechanical overload. PhenomeExpress analysis identified a hub of DEGs linked by a core of four differentially regulated TFs.

Introduction

Osteoarthritis (OA) is a major cause of pain and disability in older adults, constituting a significant economic burden which is increasing with an ageing population. Current treatments mainly address symptoms and it is important to develop more effective interventions which reduce disease progression[1]. OA is a heterogeneous disease in which there are multiple mechanisms contributing to joint failure, including, misalignment/deformity, muscle weakness, ligament laxity, subchondral bone sclerosis/cysts and osteophyte formation, but one common outcome is cartilage damage and loss[2].

In OA, the pattern of cartilage damage is typically on the most loaded tibial and femoral surfaces, whilst other less loaded areas remain intact. Previous studies have identified changes in gene expression in OA cartilage and provided evidence that all OA cartilage, including intact cartilage, differs greatly from cartilage on a healthy joint[3]. The changes that accompany OA thus affect all joint cartilage and these changes presumably weaken the tissue such that it becomes damaged and lost at the sites exposed to greatest mechanical load. We therefore set up a study to compare cartilage

* Address correspondence and reprint requests to: T.E. Hardingham, Wellcome Trust Centre for Cell-Matrix Research, Faculty of Life Sciences, Michael Smith Building, University of Manchester, Manchester M13 9PT, UK. Tel: 44-01612755511.
** Address correspondence and reprint requests to: R.P. Boot-Handford, Wellcome Trust Centre for Cell-Matrix Research, Faculty of Life Sciences, Michael Smith Building, University of Manchester, Manchester M13 9PT, UK. Tel: 44-01612755097.
E-mail addresses: sara.dunn@manchester.ac.uk (S.L. Dunn), jamie.soul@postgrad.manchester.ac.uk (J. Soul), sanjayanand@doctors.org.uk (S. Anand), jean-marc.Schwartz@manchester.ac.uk (J.-M. Schwartz), ray.boothandford@manchester.ac.uk (R.P. Boot-Handford), tim.hardingham@manchester.ac.uk (T.E. Hardingham).

http://dx.doi.org/10.1016/j.joca.2016.03.007
1063-4584/

from intact and damaged sites within the same joint to identify changes in gene expression which may contribute to intact cartilage becoming damaged. This paired analysis of samples from the same joint minimises confounding variables in patient age and genetics and increases the power of the study[4,5]. Identifying genes and regulatory pathways involved as OA cartilage becomes damaged may provide new targets for treatment to delay or reverse the damage.

This study of specific sites of OA knee cartilage was carried out using RNA-sequencing (RNA-seq). Compared to microarray technology this provides a greater dynamic range of analysis with increased sensitivity and specificity to provide enhanced identification of differentially expressed genes (DEGs). We have also applied PhenomeExpress, which incorporates cross-species gene-phenotype associations, to identify dysregulated pathways in damaged OA cartilage.

Materials and methods

Study design

Cartilage was obtained under Ethics Committee approval with written informed consent from eight patients with symptomatic OA at total knee replacement ($n = 8$, age range 65–79 years, mean age 70.3). Cartilage from paired osteochondral samples were isolated from the intact posterior lateral condyle (PLC) and the damaged distal medial condyle (DMC) for RNA-seq analysis (Group A). An additional group of paired OA samples were used to validate RNA-seq analysis ($n = 8$, age range 64–76 years, mean age 69.9) (Group B). Cartilage was transferred to RNA later for extraction and RNA-seq and/or reverse transcription quantitative polymerase chain reaction (RT-qPCR) analysis. Full depth osteochondral blocks were taken from adjacent sites, fixed in 10% neutral buffered formalin (Sigma–Aldrich) and decalcified in 20% ethylenediaminetetraacetic acid (EDTA).

Histology

De-calcified osteochondral samples were dehydrated in graded ethanol (Fisher Scientific) and immersed in xylene (Sigma–Aldrich). Samples embedded in paraffin wax were cut into sections (5 μm thickness) and stained with 0.1% safranin O-fast green for histological grading using a modified Mankin score[6]. Significant differences were determined using one way analysis of variance (ANOVA) non-parametric Friedman test.

Sulphated glycosaminoglycan assay (sGAG)

The sGAG content of cartilage tissue from the PLC and DMC was determined after overnight digestion in papain at 60°C using the dimethylmethylene blue (DMMB) assay with absorbance read at 570 nm[7].

RNA extraction

Total RNA was extracted from 200 to 400 mg of cartilage using TRIzol (LifeTechnologies) reagent and homogenisation (Braun Mikrodismembrator) following freezing in liquid nitrogen. The RNA was purified using RNeasy Qiagen clean-up columns (Qiagen) and for sequencing had a RIN score of >6 (2200 TapeStation, Agilent Technologies).

RT-qPCR

cDNA was synthesised from 0.5 to 1 μg of total RNA using MLV reverse transcriptase and random hexamers (Life Technologies). For RT-qPCR analysis primer sequences are listed in Supplementary Table 1. Gene expression was normalised to an average of glyceraldehyde 3 phosphate dehydrogenase (GAPDH) and beta actin (ACTB), which were in the bottom 1% of gene variability in the RNA-seq results. Relative gene expression levels were determined using the $2^{-\Delta\Delta Ct}$ analysis method[8]. Differences in expressed genes were identified using the non-parametric Wilcoxon signed-ranked test where P-values ≤ 0.05 were considered significant. Statistical analysis was with GraphPad Prism version 6.04.

RNA-seq

Strand specific RNA-seq libraries were generated from 0.5 to 1 μg RNA using the TruSeq® Stranded mRNA Sample Preparation Kit (Illumina, Inc.) and 101 bp paired-end reads were generated, yielding at least 39 million reads per sample. The fastq files generated by HiSeq Illumina 2000 platform were analysed with FastQC and scanned against other genomes for possible contamination. Low quality reads, contaminated barcodes and primers were further trimmed with Trimmomatic[9,10]. All libraries were aligned to hg19 assembly of human genome using Tophat-2 with the best score matches reported for each read[11]. The mapped reads were counted by genes with HTSeq against gencode v16 to reflect gene abundance[11,12]. Inter gene expression comparisons were based on calculated fragments per kilobase of transcript per million mapped reads (FPKM). Within the 16 datasets reads from 33,960 (60%) of 56,562 human genes in gencode v16 were detected. Following removal of those with lowest reads, to optimise detection of DEGs, the analysis was on 17,160 genes.

A standard method for estimation of fold change and dispersion for RNA-seq data (DESeq2) was used to initially identify DEGs[13]. The false discovery rate for the analysis (10%) was selected to provide the maximum number of DEG (1575 DEG) with a reasonable level of confidence to best inform the subsequent analysis. For comparison, a lower false discovery rate 5%, gave 1375 DEG (identified in red in Supplementary Table 2, Sheet 2). The 5000 genes with most significant changes by P-value were removed and the remaining genes used as *in silico* negative controls for batch effect factor calculation with RUVg[14]. DESeq2 was then used with batch correction to identify DEGs. The resulting P-values were adjusted for multiple testing with Benjamini–Hochberg (BH) correction. Data access to R code to reproduce the bioinformatics analysis is at https://github.com/soulj/Dunnetal2015. The RNA-seq data is available from ArrayExpress (E-MTAB-4304).

Comparison with previous microarray studies

To compare the results with two microarray studies of damaged and intact OA cartilage[4,5], the dataset GSE57218 was downloaded from Gene Expression Omnibus and Snelling *et al.* provided the raw data from their study[4,5,15]. Both array datasets were analysed as previously reported[4,5]. DEGs in all datasets were defined with 1.5 fold change and an adjusted P-value of ≤ 0.1, which are thresholds used commonly for transcriptomic analysis[16,17]. Hyper-geometric overlap statistics were used to calculate probability of the observed overlap of DEGs.

PhenomeExpress sub-network identification

PhenomeExpress was used with protein–protein, phenotype–phenotype and protein–phenotype networks created; to identify groups of interacting DEGs related to OA phenotypes[21]. With a maximum initial sub-network size of 7, an empirical P-value threshold of 0.05 was used to filter sub-networks by random sampling (10,000 sub-networks) of the filtered PPI network.

Please cite this article in press as: Dunn SL, et al., Gene expression changes in damaged osteoarthritic cartilage identify a signature of non-chondrogenic and mechanical responses, Osteoarthritis and Cartilage (2016), http://dx.doi.org/10.1016/j.joca.2016.03.007