Published in final edited form as:

Gene. 2016 January 15; 576(0): 312–318. doi:10.1016/j.gene.2015.10.037.

Identification of Differentially Methylated Regions in New Genes Associated with Knee Osteoarthritis

Carolina A. Bonin¹, Eric A. Lewallen¹, Saurabh Baheti², Elizabeth W. Bradley¹, Michael J. Stuart¹, Daniel J. Berry¹, Andre J. van Wijnen^{1,3}, and Jennifer J. Westendorf^{1,3,*}

¹Department of Orthopedic Surgery, Mayo Clinic, Rochester, MN

²Department of Biomedical Statistics and Informatics, Rochester, MN

³Department of Biochemistry & Molecular Biology, Mayo Clinic, Rochester, MN

Abstract

Epigenetic changes in articular chondrocytes are associated with osteoarthritis (OA) disease progression. Numerous studies have identified differentially methylated cytosines in OA tissues; however, the consequences of altered CpG methylation at single nucleotides on gene expression and phenotypes are difficult to predict. With the objective of detecting novel genes relevant to OA, we conducted a genome-wide assessment of differentially methylated sites (DMSs) and differentially methylated regions (DMRs). DNA was extracted from visually damaged and normal appearing, non-damaged human knee articular cartilage from the same joint and then subjected to reduced representation bisulfite sequencing. DMRs were identified using a genome-wide systematic bioinformatics approach. A sliding-window of 500bp was used for screening the genome for regions with clusters of DMSs. Gene expression levels were assessed and cell culture demethylation experiments were performed to further examine top candidate genes associated with damaged articular cartilage. More than 1,000 DMRs were detected in damaged osteoarthritic cartilage. Nineteen of these contained five or more DMSs and were located in gene promoters or first introns and exons. Gene expression assessment revealed that hypermethylated DMRs in damaged samples were more consistently associated with gene repression than hypomethylated DMRs were with gene activation. Accordingly, a demethylation agent induced expression of most hypermethylated genes in chondrocytes. Our study revealed the utility of a systematic DMR search as an alternative to focusing on single nucleotide data. In particular, this approach uncovered promising candidates for functional studies such as the hypermethylated protein-coding genes *FOXP4* and *SHROOM1*, which appear to be linked to OA pathology in humans and warrant further investigation.

*Corresponding Author: Jennifer J. Westendorf, Ph.D., Mayo Clinic, 200 First Street SW, Rochester, MN 55905, Phone: 507-538-5651, Fax: 507-284-5075, westendorf.jennifer@mayo.edu.

Author contributions

Writing: CAB, EAL, JJW; Study design: JJW, MJS, DJB; Sample collection: MJS, DJB; Experiments: CAB, EAL, EWB; Bioinformatics analyses: SB; Results interpretation: CAB, EAL, SB, EWB, AJW, JJW; Funding: EWB, AVW, JJW.

Conflicts of interest

None

Keywords

Articular cartilage; DNA methylation; differentially methylated regions; *FOXP4*; *SHROOM1*

1. Introduction

Osteoarthritis (OA) is a progressive multi-factorial disease that commonly affects joints of the hand, hip and knee. Nearly 10% of men and 20% of women > 60 years of age worldwide present OA symptoms [1] translating into a tremendous global socio-economic burden. In the United States, the average total direct cost associated with OA from 2008–2011 added up to U\$339.7 billion, a trend that will likely continue as the population ages [2].

It is established that mechanical loading of joints directly affects chondrocyte metabolism, and acute injury or persistently altered mechanical stimuli can lead to OA [3, 4]. Early stage disease is characterized by increased bone remodeling and invasion of deeper cartilage layers by blood vessels, allowing contact with chondrolytic enzymes. This is often followed by secondary inflammation and disruption of bone homeostasis with a shift toward tissue sclerosis and thickening of the calcified cartilage layer [5, 6]. Microscopically, an initial repair response causes chondrocytes to proliferate and increase cartilage matrix production, while at a later stage they become hypertrophic and increase synthesis of catabolic cytokines and matrix degrading proteases. This cascade of events that degrades collagens and proteoglycans within the cartilage extracellular matrix (ECM) in OA has been well documented [7]. Although phenotypic changes are well characterized, the identification of genes and molecular pathways involved are still underway.

Epigenetic regulation of gene expression in chondrocytes is a key mechanism behind the phenotypic changes observed in OA [8]. In particular, DNA methylation surveys, which measure presence of a methyl group on cytosine residues (5meC) at CpG di-nucleotides, have been a focus of OA studies because they can associate epigenetic states of genes and pathways with disease progression. CpG methylation is typically a repressive mark for gene expression, especially when sites are located in the gene promoter region, but can be increased in actively transcribed genes if located within intragenic regions [9]. Thus, the significance of CpG methylation in regulating gene expression is highly variable when observed in the context of CpG locus density [10]. An extensive study of more than 42 whole-genome methylation datasets and 30 human cell and tissue types revealed that only ~20% of CpGs are dynamically regulated. Most of these differentially methylated regions in the genome span 1kb and are often distal to gene transcription start sites (TSS) [10].

Array-based surveys have revealed significant differences in DNA methylation between OA and non-OA bone and cartilage in the human hip and knee at hundreds to thousands of CpG sites throughout the genome [11–16]. In some of these studies, differences among sample groups were most pronounced in genes involved in inflammation and immune-responses [12, 14], while others reported an over-representation of genes involved in developmental pathways [15, 16]. These studies corroborated previous findings for some identifiable genes associated with the onset and progression of OA. Among them are matrix metalloproteinases (*MMP-3*, *-9*, *-13*, *-14*) and aggrecanases (e.g., a disintegrin and metalloproteinase with

thrombospondin motifs; *ADAMTS-4*), whose expression are both directly [17] and indirectly (e.g., via leptin expression) [18] associated with the demethylation of CpG sites in promoter regions. Demethylation has also been reported for genes related to the inflammatory process of OA, such as *IL1B* [19, 20] and *NF- κ B* [21]. Other epigenetic changes have been detected in *GDF5* [22] and *SOX9* [23].

The inherent complexity in interpreting CpG methylation often produces weak correlations between CpG methylation and gene expression data [11]. Therefore, in addition to focusing on individual CpG sites, the genome can be scanned for clusters of differentially methylated CpG sites spanning a short region [16]. Indeed, differentially methylated regions (DMRs) have been invaluable for tissue characterization [24], with broad applications in developmental [25] and aging studies [26]. Importantly, DMRs have been used as markers for cancer diagnosis and progression [27]. Yet, the majority of genome-wide OA methylation studies have focused on single-nucleotide CpG sites, rather than their clusters.

Here, reduced representation bisulfite sequencing technique (RRBS) was used to search for novel differentially methylated CpG sites and DMRs between visually damaged and non-damaged knee cartilage. An advantage of this method over array-based methods is that it extracts methylation information directly from the converted DNA sequence rather than read counts. Direct genome sequencing also provides more complete and unbiased genomic coverage with higher accuracy, even in comparison to the most advanced high-density gene arrays such as the Infinium 450 Beadchip array (Illumina), which detects only 1.5% of CpGs in the human genome [28]. The expression of genes associated to the identified DMRs was further evaluated by qPCR and cell culture experiments. The overarching goal of this work was to generate and prioritize hypotheses that can be directly investigated by mechanistic and functional manipulations.

2. Methods

2.1 Sample collection

Articular cartilage samples were collected during total knee arthroplasties prompted by late-stage OA. Paired samples from 10 patients were collected from sections of visually damaged ("OA damaged (D)") and visually intact ("non-damaged (ND)") cartilage on femoral condyles. In addition, five patients were sampled for data validation purposes. Tissues were frozen and stored at -80°C until DNA/RNA isolation. All surgeries were conducted at Mayo Clinic (Rochester, MN) and all tissues were de-identified according to approvals granted by the Mayo Clinic Institutional Review Board.

2.2 DNA isolation and preparation for sequencing

DNA was isolated from mortar-and-pestle-ground frozen specimens with the DNeasy Kit (Qiagen). DNA (250ng) was digested with *MspI* (New England Biolabs, Catalog Number: R0106M) and purified using Qiaquick Nucleotide Removal Kit (Qiagen, Catalog Number: 28004). End-repair A tailing was performed (New England Biolabs, Catalog Number: M0212L) and TruSeq methylated indexed adaptors (Illumina, Catalog Number: 15025064) were ligated with T4 DNA ligase (New England Biolabs, Catalog Number: M0202L). Size

selection was performed with Agencourt AMPure XP beads (Beckman Coulter, Catalog Number: A63882). Bisulfite conversion was performed using EZ-DNA Methylation Kit (Zymo Research, Catalog Number: D5001) as recommended by the manufacturer with the exception that incubation was performed using 55 cycles of 95 °C for 30 sec and 50°C for 15 min. Following bisulfite treatment, the DNA was purified as directed and amplified using Pfu Turbo C Hotstart DNA Polymerase (Agilent Technologies, Catalog Number: 600414). Library quantification was performed using Qubits dsDNA HS Assay Kit (Life Technologies, Catalog Number: Q32854) and the Bioanalyzer DNA 1000 Kit (Agilent Technologies, Catalog Number: 5067-1504). The final libraries from RRBS were prepared for sequencing as per the manufacturer's instructions in the Illumina cBot and HiSeq Paired end cluster kit v.3. The samples were placed onto seven lanes of a paired-end flow cell at concentrations of 7–8 pM and the control sample, PhiX, was placed in the eighth lane to allow the sequencer to account for the unbalanced representation of cytosine bases. The flow cell was then loaded into the Illumina cBot for generation of cluster densities. After cluster generation, flow cells were sequenced as 51×2 paired end reads using Illumina HiSeq 2000 with TruSeq SBS sequencing kit v.3. Data was collected using HiSeq data collection v. 1.5.15.1 software, and the bases were called using Illumina's RTA v.1.13.48.

2.3 Sequence alignment and quality control

The raw sequence FASTQ files obtained from the sequencing instrument were processed using a Streamlined Analysis and Annotation Pipeline for Reduced Representation Bisulfite Sequencing (SAAP-RRBS) [29]. Initial data processing included the removal of adaptor sequences from sequencing files, which were then mapped to a reference human genome (version Hg19) using the bisulfite sequence mapping program BSMAP [30]. All samples had > 50% aligned reads and only the status of CpG sites with > 5x coverage were considered for downstream analyses. All samples had similar methylation patterns, presenting a bimodal distribution (Supplementary Fig. 1S). In addition, CpG call rates were calculated for all samples to verify that they surpassed an 80% call rate threshold (Supplementary Fig. 2S).

2.4 Differential methylation

CpGs with consistent methylation ratios of either 0 or 1 across all samples were removed from the datasets. Differentially methylated CpG sites (DMSs) between OA and non-OA tissues were detected using paired t statistics (p values adjusted after Bonferroni correction). Both differentially methylated individual CpG sites (DMS) and differentially methylated regions (DMRs) in the genome were identified. To identify DMRs, we applied a sliding-window approach with a window size of 500 bp and a variable slide step (depending on next CpG location) across identified significant DMS ($p < 0.05$) between the sample groups. Windows that contained more than two CpG sites and were supported with at least five reads were kept and merged to define a DMR. The p values were averaged across the CpG sites within a region and were used to weight the significance of DMRs. The resulting gene lists associated with DMRs were analyzed using the gene over-representation analysis tool in PANTHER v 9.0 [31], which is continuously updated for gene ontology (GO) term annotations directly curated by the Gene Ontology Consortium [32].

2.5 Real-time PCR

Total genomic RNA was isolated from each tissue using a cartilage-specific RNA extraction kit (Biochain) following the manufacturer's protocol. The isolated RNA was reverse transcribed into cDNA using the SuperScript III First-Strand Synthesis System (Invitrogen) and gene expression was quantified using semi-quantitative real-time reverse transcriptase PCR (qRT-PCR). Each reaction was performed with 12.5 ng cDNA in 10 μ l using the QuantiTect Sybr Green PCR Kit (Qiagen) and the CFX384 Real-Time System (BioRad). RNA extractions and qRT-PCR reactions were both performed in triplicate. Primer sequences are listed in Supplementary Table S1. Transcript levels of mRNA were normalized to the housekeeping gene *GAPDH* and quantified using the 2^{-C_t} method.

2.6 DNA methylation inhibitor experiments

T/C28a2 human chondrocytes were maintained in DMEM supplemented with 10% FBS and 1% antibiotic/antimycotic. For experimental conditions, cells were seeded at a density of 2.7×10^4 cells/cm² and incubated overnight. To inhibit DNA methylation, 1 μ M 5-azacytidine (5-AzaC) or vehicle (PBS) was added on days 1 and 3 of culture. Cells were collected on day 4 for RNA extraction, which was conducted using TRIzol (Life Technologies) following the manufacturer's protocol. Gene expression was evaluated as described above, in triplicate, with averages $\pm$ SEM presented. T/C28a2 cells were kindly provided by M.B. Goldring.

3. Results

3.1 Differential methylation between damaged and non-damaged OA cartilage

Bisulfite sequencing revealed the methylation status of approximately 1.9 million CpGs throughout the genomes of damaged (D) and visually non-damaged (ND) articular cartilage from 10 OA patients. A total of 39,322 DMSs were identified between D and ND sample groups ($p < 0.05$) and in their majority, they were located in intergenic and intron regions (Figs. 1A, B). Notably, DMSs in OA cartilage were less frequently found in 5'UTR and promoter regions (0 and 5% respectively; Fig. 1B) than would be expected based on the overall distribution of CpGs in the genome (5 and 16% respectively; Fig. 1A). As well, only 14% of DMSs in OA cartilage were found in CpG islands, which was lower than the overall frequency of CpG sites throughout the genome (43%). No difference in the distribution of DMSs regarding genomic location was observed when comparing hypermethylated to hypomethylated DMSs (Supplementary Table S2).

To assess data at the level of individual CpG loci, we selected DMSs with at least 10% difference between D and ND datasets and an adjusted p value cutoff of $p < 0.01$ to match paired t test results. A total of 3,951 differentially methylated CpG sites (DMS) were detected (Supplementary Tables S3a and b). Some genes with well-established roles in ECM degradation were found to contain DMSs. For example, *ADAMTS-4* had hypomethylated CpG sites in damaged tissues (e.g., chr1: 1611680773; 21% difference between D and ND datasets; $p = 0.004$).

Hierarchical clustering of samples based on DMSs ($n = 3,951$) revealed two clusters consistent with D (red) and ND (blue) cartilage specimens (Fig. 2). Sub-clustering was observed within each group and consisted of two patients (#4 and 5, light red and blue) who each presented a distinct methylation profile. In addition, two independent sub-clusters of D and ND samples containing two and four samples were also detected. Sampling limitations precluded us from specifically testing for a statistical pattern regarding these patient sub-clusters. Overall, DMSs clustered specimens by sample condition (D vs. ND) rather than by patient.

3.2 Differentially methylated regions

More than 1,000 DMRs, with at least two significant DMSs, were detected between D and ND datasets. Of these, 442 were hypermethylated while 560 were hypomethylated in D tissues (Supplementary Tables S4a and b). GO term over-representation analyses were performed separately for the genes associated with these hypermethylated and hypomethylated DMRs. A significant over-representation of genes involved in limb and skeletal development was detected in the hypermethylated DMRs (Table 1). In contrast, the hypomethylated DMRs were most closely associated with pathways that control organ development and morphogenesis. We proceeded to focus our investigations on DMRs with a minimum of five DMSs. This search resulted in 48 hypermethylated and 28 hypomethylated DMRs in D datasets. Among these "top" DMRs, 19 were located within gene promoter regions or first exon/introns (Table 2).

3.3 Data validation: gene expression and T/C28a2 cell culture

Next, we assessed expression levels for the protein-coding genes associated with "top DMRs" (Table 2). Reduced gene expression was observed in five out of the six hypermethylated genes in damaged tissues: *ALX4*, *HOXC4*, *FOXP4*, *SHROOM1* and *TBX5*. However, the protein-coding genes *ALX*, *HOXC4* and *TBX5* were not consistently detected in all samples and had low transcript levels (< 0.5 ; Supplementary Table S5). *FOXP4* and *SHROOM1*, which were both detected in all samples, had marginally lower expression values in damaged tissues (Student's *t* test, $p = 0.06$ and $p = 0.07$, respectively, Figs. 3A, 4A). Increased gene expression in damaged tissues was only observed for one out of the six hypomethylated genes (*FANK1*; Student's *t* test, $p = 0.007$), which had consistent low expression levels and high variability across samples (Supplementary Table S5).

The DNA demethylase inhibitor, 5-AzaC, significantly increased gene transcripts of *FOXP4*, *SHROOM1* (Figs. 3B and 4B), as well as several other hypermethylated genes (*SLC43A2*, *ANKRD20A8P*, *PSIMCT-1*; Supplementary Fig. S4) in T/C28a2 chondrocytes; however, 5-AzaC decreased *HOXC4* expression. 5-AzaC also increased the expression of the hypomethylated genes *EFNB1*, *NHS*, *E2FB* and *LINC00221*, but this result was only statistically significant for *EFNB1* (Supplementary Fig. S4).

3.4 Detailed analyses of FOXP4 and SHROOM1 DMRs

Using gene expression assessments as a guide, we re-examined the methylation status of the *FOXP4* and *SHROOM1* DMRs in more detail for each patient. For *FOXP4*, we found that most patients (except #3, 5, 8) had a differential methylation status when comparing

damaged and non-damaged samples; a pattern most pronounced for the CpG sites at chr 6: 41528776; 41528781 (Figs. 3C, D). Similarly, all patients (except #1) had marked differences in methylation status (average = 35%) for all *SHROOM1* DMSs between damaged and non-damaged tissues (Figs. 4C, D).

The patient-level variability for *FOXP4* and *SHROOM1* DMRs was additionally explored by clustering and heat maps of all samples driven by methylation status. Although inter-patient variability was detected, DMRs tended to cluster samples by disease state rather than patient, suggesting that a single DMR can be informative for characterizing OA tissues (Supplementary Fig. S3).

4. Discussion

Previous methylation array analyses of human cartilage have revealed significant differences in the genomic methylation profiles of tissues from osteoarthritic joints and identified key roles for epigenetic events in disease progression, in particular those involved in cartilage degradation and immune response. However, the detection of novel genes and molecular pathways has been limited by the low coverage of array assays and difficulty in interpreting the functional significance of single nucleotide methylation status. With the objective of detecting novel genes relevant to gross tissue damage in OA joints, we employed reduced bisulfite-converted DNA sequencing technology and conducted a systematic search for DMRs between visually damaged and normal appearing, non-damaged articular cartilage from the same human knee joint. Our data revealed more than 1,000 OA DMRs; 19 of which had 5 or more DMSs located within promoters or first introns/exons. By using paired samples from the same joint, we showed that these DMRs can be used to inform on OA disease pathology and better understand epigenetic mechanisms by typically clustering samples by disease state rather than patient. Gene expression assessment of DMR-associated genes and demethylation experiments demonstrated that hypermethylated DMRs are inversely correlated with gene expression, while hypomethylated DMRs are unpredictable. Importantly, our work revealed novel OA-related DMRs in the *FOXP4* and *SHROOM1* genes.

Most of the CpGs with variable methylation status observed in knee cartilage tissues were found in introns and intergenic regions, with only a small fraction of them occurring within coding sequences or promoters. Two other recent OA epigenetic studies reported similar findings [13, 14]. These results emphasize the potential of bisulfite sequencing technologies to uncover important epigenetic changes related to OA that occur outside of gene promoters. However, this information can only be productively harnessed after detailed mapping of distal regulatory elements is completed [33]. For the purposes of this work we focused on DMRs found within promoter or first intron/exon regions.

Despite some expected inter-patient variability, a particular subset of DMRs that contained five or more DMSs was broadly consistent across damaged versus visually non-damaged samples. Nonetheless, validation via RT-qPCR and demethylase inhibitor treatment supported these findings for only a few genes. Interpreting the significance of methylation status has been problematic for all OA genome-wide studies to date, and poor correlations

between methylation and gene expression at the individual gene level are often observed [11, 14]. In our study, which systematically focused on DMRs with a high number of DMSs, most of the validated genes were associated with top hypermethylated DMRs in damaged tissues. Meanwhile the validation data on gene expression of hypomethylated genes was highly variable. Not surprisingly, hypomethylation data are comparatively less informative than hypermethylation. This is due, at least in part, to the fact that DNA methylation is only one of several epigenetic mechanisms affecting gene expression; others include non-coding RNAs and histone modifications. In fact, epigenetic research in cancer cell lines has recently demonstrated that hypomethylation does not necessarily activate gene expression on its own and this is frequently only achieved when hypomethylation is coupled with a gain of repressive chromatin marks [34].

Among genes examined, *FOXP4* had highest gene expression in all cartilage tissues studied. In agreement with the *FOXP4* differential methylation status, damaged articular cartilage had reduced gene expression relative to visually non-damaged articular cartilage from the same OA joint. Further, the demethylation treatment of articular chondrocytes suggested that DNA methylation suppresses *FOXP4* expression. Notably, two hypermethylated CpG sites associated with *FOXP4* were detected in another study of OA damaged hip cartilage, although these were not the exact same sites we uncovered [13]. Forkhead box (*FOX*) genes encode transcription factors that repress transcription and a recent study suggests a fundamental role for the complex *FOXP1/2/4* in endochondral ossification. Thus, when overexpressed, this complex impedes osteoblast formation and chondrocyte hypertrophy, while its deficiency results in premature osteoblast differentiation and severe growth plate defects [35]. This report suggested that *FOXP1/2/4* may adversely affect chondrocyte hypertrophy by inhibiting *RUNX2*, a master regulator of endochondral ossification.

Our work also identified *SHROOM1* as a potential gene associated with OA. SHROOM proteins interact with actin and regulate cell morphology [36], and studies in *Xenopus* suggest that these proteins control epithelial thickening [37]. A role for SHROOM in cartilage homeostasis has not been documented and high levels of inter-patient variability hampered the validation of this gene by qPCR. Nevertheless, differential methylation was high and significant between damaged and non-damaged datasets at 15 DMSs, warranting further investigation regarding its potential role in OA pathology.

5. Conclusion

This work provides an assessment of differences in methylation profile between visually damaged and non-damaged knee cartilage from the same OA joint using reduced bisulfite sequencing technology, with a focus on DMRs where clusters of DMSs span a short genomic region. The use of paired samples reduced the complexity of data interpretation and emphasized epigenetic changes relative to OA. The combined validation of the methylation dataset via gene expression assessment and demethylation experiments revealed *FOXP4* and *SHROOM1* as promising candidates for future functional studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Xiaodong Li, Dr. Yang Lin, Lomeli Carpio, Dr. Scott Reister, Dr. Emily Camilleri and Dr. Amel Dudakovic for their assistance.

Role of funding sources

This work was supported by NIH grants: R01 AR049069 (AvW); R01 AR68103 (JJW), R01 DE020194 (JJW), K01 AR65397 (EWB); T32 AR56950 (EAL); Mayo Clinic Center for Individualized Medicine.

References

1. Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. *Bulletin of the World Health Organization*. 2003; 81:646–656. [PubMed: 14710506]
2. Yelin, EH.; Cisternas, M.; Watkins-Castillo, SI. *Medical Care Expenditures and Earnings Losses for Select Musculoskeletal Diseases*. Third ed.. Rosemont, IL: 2014. United States Bone and Joint Initiative: The Burden of Musculoskeletal Diseases in the United States (BMUS).
3. Goldring MB, Marcu KB. Cartilage homeostasis in health and rheumatic diseases. *Arthritis research & therapy*. 2009; 11:224. [PubMed: 19519926]
4. Goldring MB, Goldring SR. Articular cartilage and subchondral bone in the pathogenesis of osteoarthritis. *Annals of the New York Academy of Sciences*. 2010; 1192:230–237. [PubMed: 20392241]
5. Burr DB, Gallant MA. Bone remodelling in osteoarthritis. *Nature reviews. Rheumatology*. 2012; 8:665–673.
6. Goldring MB, Goldring SR. Osteoarthritis. *Journal of cellular physiology*. 2007; 213:626–634. [PubMed: 17786965]
7. Goldring MB. The role of the chondrocyte in osteoarthritis. *Arthritis and rheumatism*. 2000; 43:1916–1926. [PubMed: 11014341]
8. Goldring MB, Marcu KB. Epigenomic and microRNA-mediated regulation in cartilage development, homeostasis, and osteoarthritis. *Trends in molecular medicine*. 2012; 18:109–118. [PubMed: 22178468]
9. Laird PW. Principles and challenges of genomewide DNA methylation analysis. *Nature reviews. Genetics*. 2010; 11:191–203. [PubMed: 20125086]
10. Ziller MJ, Gu H, Muller F, Donaghey J, Tsai LT, Kohlbacher O, et al. Charting a dynamic DNA methylation landscape of the human genome. *Nature*. 2013; 500:477–481. [PubMed: 23925113]
11. Delgado-Calle J, Fernandez AF, Sainz J, Zarrabeitia MT, Sanudo C, Garcia-Renedo R, et al. Genome-wide profiling of bone reveals differentially methylated regions in osteoporosis and osteoarthritis. *Arthritis and rheumatism*. 2013; 65:197–205. [PubMed: 23124911]
12. Fernandez-Tajes J, Soto-Hermida A, Vazquez-Mosquera ME, Cortes-Pereira E, Mosquera A, Fernandez-Moreno M, et al. Genome-wide DNA methylation analysis of articular chondrocytes reveals a cluster of osteoarthritic patients. *Annals of the rheumatic diseases*. 2014; 73:668–677. [PubMed: 23505229]
13. Rushton MD, Reynard LN, Barter MJ, Refaie R, Rankin KS, Young DA, et al. Characterization of the cartilage DNA methylome in knee and hip osteoarthritis. *Arthritis & rheumatology*. 2014; 66:2450–2460. [PubMed: 24838673]
14. Jeffries MA, Donica M, Baker LW, Stevenson ME, Annan AC, Humphrey MB, et al. Genome-Wide DNA Methylation Study Identifies Significant Epigenomic Changes in Osteoarthritic Cartilage. *Arthritis & rheumatology*. 2014; 66:2804–2815. [PubMed: 24980887]

15. Moazedi-Fuerst FC, Hofner M, Gruber G, Weinhaeusel A, Stradner MH, Angerer H, et al. Epigenetic Differences in Human Cartilage Between Mild and Severe OA. *Journal of Orthopaedic Research*. 2014; 32:1636–1645. [PubMed: 25212754]
16. den Hollander W, Ramos YF, Bomer N, Elzinga S, van der Breggen R, Lakenberg N, et al. Transcriptional Associations of Osteoarthritis-Mediated Loss of Epigenetic Control in Articular Cartilage. *Arthritis & rheumatology*. 2015; 67:2108–2116. [PubMed: 25892573]
17. Roach HI, Yamada N, Cheung KSC, Tilley S, Clarke NMP, Oreffo ROC, et al. Association between the abnormal expression of matrix-degrading enzymes by human osteoarthritic chondrocytes and demethylation of specific CpG sites in the promoter regions. *Arthritis and rheumatism*. 2005; 52:3110–3124. [PubMed: 16200590]
18. Iliopoulos D, Malizos KN, Tsezou A. Epigenetic regulation of leptin affects MMP-13 expression in osteoarthritic chondrocytes: possible molecular target for osteoarthritis therapeutic intervention. *Annals of the rheumatic diseases*. 2007; 66:1616–1621. [PubMed: 17502362]
19. Hashimoto K, Oreffo RO, Gibson MB, Goldring MB, Roach HI. DNA demethylation at specific CpG sites in the IL1B promoter in response to inflammatory cytokines in human articular chondrocytes. *Arthritis and rheumatism*. 2009; 60:3303–3313. [PubMed: 19877066]
20. Hashimoto K, Otero M, Imagawa K, de Andres MC, Coico JM, Roach HI, et al. Regulated transcription of human matrix metalloproteinase 13 (MMP13) and interleukin-1beta (IL1B) genes in chondrocytes depends on methylation of specific proximal promoter CpG sites. *The Journal of biological chemistry*. 2013; 288:10061–10072. [PubMed: 23417678]
21. de Andres MC, Imagawa K, Hashimoto K, Gonzalez A, Roach HI, Goldring MB, et al. Loss of methylation in CpG sites in the NF-kappaB enhancer elements of inducible nitric oxide synthase is responsible for gene induction in human articular chondrocytes. *Arthritis and rheumatism*. 2013; 65:732–742. [PubMed: 23239081]
22. Reynard LN, Bui C, Syddall CM, Loughlin J. CpG methylation regulates allelic expression of GDF5 by modulating binding of SP1 and SP3 repressor proteins to the osteoarthritis susceptibility SNP rs143383. *Human genetics*. 2014; 133:1059–1073. [PubMed: 24861163]
23. Kim KI, Park YS, Im GI. Changes in the epigenetic status of the SOX-9 promoter in human osteoarthritic cartilage. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research*. 2013; 28:1050–1060.
24. Song F, Smith JF, Kimura MT, Morrow AD, Matsuyama T, Nagase H, et al. Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102:3336–3341. [PubMed: 15728362]
25. Song F, Mahmood S, Ghosh S, Liang P, Smiraglia DJ, Nagase H, et al. Tissue specific differentially methylated regions (TDMR): Changes in DNA methylation during development. *Genomics*. 2009; 93:130–139. [PubMed: 18952162]
26. Bell JT, Tsai PC, Yang TP, Pidsley R, Nisbet J, Glass D, et al. Epigenome-Wide Scans Identify Differentially Methylated Regions for Age and Age-Related Phenotypes in a Healthy Ageing Population. *Plos Genetics*. 2012; 8:189–200.
27. Ushijima T. Innovation - Detection and interpretation of altered methylation patterns in cancer cells. *Nature Reviews Cancer*. 2005; 5:223–231.
28. Bock C. Analysing and interpreting DNA methylation data. *Nature reviews. Genetics*. 2012; 13:705–719.
29. Sun Z, Baheti S, Middha S, Kanwar R, Zhang Y, Li X, et al. SAAP-RRBS: streamlined analysis and annotation pipeline for reduced representation bisulfite sequencing. *Bioinformatics*. 2012; 28:2180–2181. [PubMed: 22689387]
30. Xi Y, Li W. BSMAP: whole genome bisulfite sequence MAPping program. *BMC bioinformatics*. 2009; 10:232. [PubMed: 19635165]
31. Mi HY, Muruganujan A, Thomas PD. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Research*. 2013; 41:D377–D386. [PubMed: 23193289]
32. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. *Nature genetics*. 2000; 25:25–29. [PubMed: 10802651]

33. Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, et al. An atlas of active enhancers across human cell types and tissues. *Nature*. 2014; 507:455–461. [PubMed: 24670763]
34. Hon GC, Hawkins RD, Caballero OL, Lo C, Lister R, Pelizzola M, et al. Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. *Genome Research*. 2012; 22:246–258. [PubMed: 22156296]
35. Zhao HX, Zhou WR, Yao ZJ, Wan Y, Cao JJ, Zhang LL, et al. Foxp1/2/4 regulate endochondral ossification as a suppresser complex. *Developmental Biology*. 2015; 398:242–254. [PubMed: 25527076]
36. Dietz ML, Bernaciak TM, Vendetti F, Kielec JM, Hildebrand JD. Differential actin-dependent localization modulates the evolutionarily conserved activity of Shroom family proteins. *Journal of Biological Chemistry*. 2006; 281:20542–20554. [PubMed: 16684770]
37. Lee C, Le MP, Wallingford JB. The Shroom Family Proteins Play Broad Roles in the Morphogenesis of Thickened Epithelial Sheets. *Developmental Dynamics*. 2009; 238:1480–1491. [PubMed: 19384856]

Fig. 1.

Location of differentially methylated CpG sites (DMSs) detected in osteoarthritis damaged and non-damaged knee tissues (n= 20 knee joints from 10 patients) relative to genomic features. a) ALL = all CpGs in the dataset (n=1,974,223); b) DMSs = differentially methylated CpG sites, identified using t paired tests ($p < 0.05$; n= 39,323). All comparisons significant, Fisher exact tests ($p < 0.0001$).

Fig. 2. Unsupervised heatmap/clustering of D and ND samples (n= 10 per group) based on differentially methylated CpG sites detected between groups ($p < 0.01$; difference between groups of 10%; n= 3,951).

Fig. 3.

Hypermethylated region within the *FOXP4* gene detected in visually damaged (D) versus normal appearing, non-damaged (ND) knee articular cartilage tissues. A) Boxplot of gene expression for *FOXP4* (n= 10 knee sections from 5 patients, $p = 0.06$). Gene expression was normalized to the housekeeping gene *GAPDH*; sample group differences were tested using Student t tests. B) Gene expression ($\pm$ SEM) in human chondrocytes treated with 5-AzaC. Data are shown as fold change relative to controls treated with PBS ($p = 0.01$). C) Methylation status at each DMS across patients (#1-10). The methylation data was divided

in quartiles (gradient from white to black corresponds to increasing methylation ratio). D)
Average methylation ratio for D and ND samples at each DMS.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

SHROOM1

Fig. 4.

Hypermethylated region within the *SHROOM1* gene detected in visually damaged (D) versus normal appearing, non-damaged (ND) knee articular cartilage tissues. A) Boxplot of gene expression for *SHROOM1* (n= 10 knee sections from 5 patients, $p = 0.07$). Transcript levels were normalized to the housekeeping gene GAPDH; sample group differences were tested using Student t tests. B) Gene expression ($\pm$ SEM) in human chondrocytes treated with 5-AzaC. Data are shown as fold change relative to controls treated with PBS ($p = 0.06$). C) Methylation status at each DMS across patients (#1-10). The methylation data was

divided in quartiles (gradient from white to black corresponds to increasing methylation ratio). D) Average methylation ratio for D and ND samples at each DMS.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE 1

Gene Ontology (GO) terms over-represented in gene lists associated with DMRs (n= 1,002) detected between D and ND datasets (n= 20, 2 specimens per patient $\times$ 10 patients). The number of genes in each GO term category (Ref. set) and DMR dataset (Exp. set) are listed. Table is sorted by *p* values (Bonferroni adjusted).

GO term	Ref. set	Exp. set	Fold enrichment	<i>p</i> value
Hypermethylated DMRs in OA				
embryonic morphogenesis	554	39	4.64	0.0000
skeletal system development	465	32	4.53	0.0000
skeletal system morphogenesis	212	18	> 5	0.0001
limb development	169	16	> 5	0.0001
appendage development	169	16	> 5	0.0001
forelimb morphogenesis	40	9	> 5	0.0001
limb morphogenesis	150	15	> 5	0.0001
appendage morphogenesis	150	15	> 5	0.0001
cell fate commitment	236	18	> 5	0.0003
central nervous system neuron differentiation	169	15	> 5	0.0006
embryonic limb morphogenesis	130	13	> 5	0.0011
embryonic appendage morphogenesis	130	13	> 5	0.0011
odontogenesis of dentin-containing tooth	74	10	> 5	0.0021
mesenchyme development	167	14	> 5	0.0030
stem cell differentiation	282	18	4.2	0.0036
odontogenesis	103	11	> 5	0.0054
embryonic skeletal system development	127	12	> 5	0.0060
embryonic forelimb morphogenesis	34	7	> 5	0.0086
neuron fate commitment	71	9	> 5	0.0139
digestive system development	141	12	> 5	0.0174
embryonic digit morphogenesis	61	8	> 5	0.0408
Hypomethylated DMRs in OA				
multicellular organismal development	4256	107	1.72	0.0000
system development	3687	96	1.78	0.0000
single-organism developmental process	4830	113	1.6	0.0003
organ morphogenesis	853	35	2.8	0.0004
developmental process	4909	114	1.59	0.0004
single-multicellular organism process	5799	128	1.51	0.0008
anatomical structure development	4330	103	1.63	0.0008
organ development	2637	70	1.81	0.0042
multicellular organismal process	6048	129	1.46	0.0051
biological regulation	10450	195	1.28	0.0109
anatomical structure morphogenesis	2084	57	1.87	0.0244
regulation of cellular process	9560	180	1.29	0.0315

TABLE 2

Differentially methylated Regions (DMRs) within gene promoters or gene first intron/exons, where at least 5 differentially methylated CpG sites (DMSs) between D and ND datasets (n= 20, 2 specimens per patient × 10 patients) occurred within 500bp. Table is alphabetically sorted by Gene Name.

Gene Name	Chr	Start	Stop	# of DMSs	Annotation	Gene Type
Hypermethylated in OA Damaged						
<i>ALX4</i>	chr11	44325734	44326486	13	intron (NM_021926, intron 1 of 3)	protein-coding
<i>ANKRD20A8P</i>	chr2	95523148	95523200	7	promoter-TSS (NR_003366)	pseudo
<i>FOXP4</i>	chr6	41528750	41528785	6	intron (NM_001012427, intron 1 of 16)	protein-coding
<i>HOXC4</i>	chr12	54441162	54441393	5	intron (NM_014620, intron 1 of 3)	protein-coding
<i>MIR6720</i>	chr6	1393272	1393304	5	intron (NM_001452, intron 1 of 1)	ncRNA
<i>MTIJP</i>	chr16	56669318	56669590	7	promoter-TSS (NR_036677)	pseudo
<i>PSIMCT-1</i>	chr20	30135097	30135124	5	promoter-TSS (NR_003677)	pseudo
<i>SHROOM1</i>	chr5	132161329	132161839	15	exon (NM_133456, exon 1 of 7)	protein-coding
<i>SLC43A2</i>	chr17	1508270	1508522	5	promoter-TSS (NM_001284499)	protein-coding
<i>SOX9-AS1</i>	chr17	70112297	70112388	6	promoter-TSS (NR_103738)	ncRNA
<i>TBX5</i>	chr12	114846924	114847376	6	promoter-TSS (NM_000192)	protein-coding
Hypomethylated in OA						
<i>ARC</i>	chr8	143693678	143694081	5	3' UTR (NM_015193, exon 1 of 3)	protein-coding
<i>E2F6</i>	chr2	11606211	11606252	5	promoter-TSS (NM_001278278)	protein-coding
<i>EFNB1</i>	chrX	68049296	68049877	7	5' UTR (NM_004429, exon 1 of 5)	protein-coding
<i>FANK1</i>	chr10	127584737	127585425	12	promoter-TSS (NM_145235)	protein-coding
<i>LINC00221</i>	chr14	106938251	106938625	5	promoter-TSS (NR_027457)	ncRNA
<i>IRF2BPL</i>	chr14	77492413	77492731	12	exon (NM_024496, exon 1 of 1)	protein-coding
<i>MIR6165</i>	chr17	47603826	47603883	6	intron (NR_103773, intron 1 of 2)	ncRNA
<i>NHS</i>	chrX	17394218	17394270	6	exon (NM_198270, exon 1 of 8)	protein-coding