



Published in final edited form as:

*Arthritis Rheumatol.* 2015 May ; 67(8): 2129–2140. doi:10.1002/art.39179.

## Genome-wide mapping of DNA hydroxymethylation in osteoarthritic chondrocytes

Sarah E. B. Taylor, PhD<sup>1</sup>, Ye Henry Li, BA<sup>2</sup>, Wing H. Wong, PhD<sup>3</sup>, and Nidhi Bhutani, PhD<sup>1</sup>

<sup>1</sup>Department of Orthopaedic Surgery, Stanford University School of Medicine, Stanford, CA 94305, USA

<sup>2</sup>Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305, USA

<sup>3</sup>Department of Statistics, Stanford University School of Medicine, Stanford, CA 94305, USA

### Abstract

**Objective**—To examine genome-wide 5hmC distribution in osteoarthritic (OA) and normal chondrocytes to investigate the effect on OA-specific gene expression.

**Methods**—Cartilage was obtained from OA patients undergoing total knee arthroplasty or control patients undergoing anterior cruciate ligament reconstruction. Genome-wide sequencing of 5hmC-enriched DNA (5hmC-seq) was performed for a small cohort of normal and OA chondrocytes to identify differentially hydroxymethylated regions (DhMRs) in OA chondrocytes. 5hmC-seq data was intersected with global OA gene expression data to define subsets of genes and pathways potentially affected by increased 5hmC levels in OA chondrocytes.

**Results**—70591 DhMRs were identified in OA chondrocytes compared to normal chondrocytes, 44288 (63%) of which were increased in OA chondrocytes. The majority of DhMRs (66%) were gained in gene bodies. Increased DhMRs were observed in ~50% of genes previously implicated in OA pathology including MMP3, LRP5, GDF5 and COL11A1. Furthermore, analyses of gene expression data revealed gene body gain of 5hmC appears to be preferentially associated with activated but not repressed genes in OA chondrocytes.

**Conclusion**—This study provides the first genome-wide profiling of 5hmC distribution in OA chondrocytes. We had previously reported a global increase in 5hmC levels in OA chondrocytes. Gain of 5hmC in the gene body is found to be characteristic of activated genes in OA chondrocytes, highlighting the influence of 5hmC as an epigenetic mark in OA. In addition, this study identifies multiple OA-associated genes that are potentially regulated either singularly by gain of DNA hydroxymethylation or in combination with loss of DNA methylation.

Osteoarthritis (OA) is a multifactorial disease that affects as much as 40% of the elderly population (1), yet a clear understanding of OA etiology and disease remain elusive. Many

---

**Corresponding author:** Nidhi Bhutani 300, Pasteur Drive Edwards Bldg., R164 Stanford, CA 94305-5341 Tel: 650 721 2234 Fax: 650 723 9370 [nbhutani@stanford.edu](mailto:nbhutani@stanford.edu).

Accession numbers

The microarray expression and 5hmC-sequencing data in this report have been deposited to GEO in a SuperSeries with accession number GSE64395. This SuperSeries is composed of the SubSeries GSE64393 and GSE64394.

factors contribute to the onset and progression of OA including (but not limited to): obesity, trauma, aging and genetic predisposition (2, 3). Genome-wide association studies (GWAS), candidate gene approaches and linkage analyses (reviewed in 4, 5) have revealed particular genes or specific chromosomal loci associated with the disease. However, it has become apparent that the majority of individual risk alleles have only a modest contribution to disease inheritance (6). The focus has therefore shifted to epigenetic changes associated with OA, their potential impact on OA-associated gene expression and disease pathology (7). Reversible epigenetic changes can be attractive therapeutic targets, as demonstrated for various cancers (reviewed in 8), hence an increased understanding of the nature and extent of epigenetic alterations in OA is needed.

The role of modified cytosines, especially DNA methylation of cytosine residues in promoter regions, has been well studied and is associated with gene repression (9). There is increasing evidence for DNA methylation and demethylation to be regulatory in OA-associated gene expression (7, 10). Initial investigations focused on genes that had previously been implicated in disease susceptibility and progression (including MMP3, 9, 13, ADAMTS4 and IL1 $\beta$ ) and identified a loss of methylation in the promoters of these genes (11-14). Recently, several genome scale studies have characterized the methylome of cartilage DNA from OA patients. The first published analyses of genome wide DNA methylation in OA identified 91 differentially methylated loci between normal and OA knee cartilage, and further identified a cluster of OA patients where an increased, OA-related, inflammatory response may be regulated by DNA methylation (15). High-density methylation arrays have been used to compare methylation patterns between knee and hip OA chondrocytes (16) revealing not only that OA cartilage has a unique methylation profile when compared to healthy controls, but also that there are site specific differences in methylation between hip OA and knee OA chondrocytes. Another recent report examined DNA methylation changes in different histopathological grades of OA hip cartilage, and found differential methylation at several CpG sites was highly correlated with disease score, indicating that some epigenetic marks are acquired with disease progression (17).

In the past few years, huge advancements in the understanding of DNA demethylation dynamics have been made (18). Since the initial discovery that oxidation of methylated cytosine (5mC) by the TET family of enzymes results in the formation of hydroxymethylated cytosine (5hmC) (19, 20), it has become widely accepted that 5hmC is an intermediate in the active DNA demethylation pathway. However, a growing body of evidence suggests that 5hmC not only exists as an intermediate of DNA demethylation but also as a stand alone, stable epigenetic mark that can affect gene expression (21, 22). Genome-wide analyses of 5hmC distribution in embryonic stem cells (ESCs) and neurons have found a positive correlation between 5hmC accumulation in gene bodies and gene transcription (23, 24). Accumulation of 5hmC has also been found at the start sites of genes with bivalent promoters (containing both activating (H3K4me3) and repressive (H3K27me3) histone marks) in ESCs, suggesting that it contributes to a 'poised' epigenetic state (25). Our own studies into the role of 5hmC in cartilage differentiation have found that gain of 5hmC in the gene bodies of cartilage specific genes is linked to transcriptional upregulation in differentiated chondrocytes (manuscript in review).

We recently reported a remarkable dysregulation of 5hmC homeostasis in patients with OA, with increases in global 5hmC levels observed in OA chondrocytes when compared to normal chondrocytes (26). In addition, we demonstrated that 5hmC gain at specific sites in the promoters of key OA genes was associated with increased expression. Based on these findings, the goal of the present study was to map the precise distribution of altered global patterns of 5hmC in OA and to define their functional effect on gene expression. The aim of these analyses is to define subsets of genes and pathways affected by increased 5hmC levels in OA in order to provide new mechanistic insights into the epigenetic regulation of OA.

## Materials and Methods

### Chondrocyte isolation and culture

Normal articular chondrocytes (donor ages = 24 weeks, 6 months, 34 years) were purchased from Lonza or obtained from patients with no history of OA symptoms undergoing surgical procedures (donor age = 18 months, 27 years). Articular chondrocytes were harvested from OA cartilage samples obtained during total knee arthroplasty (donor ages = 61, 64, 65, 73 and 74 years). Cartilage was dissected and the chondrocytes cultured in high-density monolayers for limited passages (2-3) over 5-7 days, as described previously (27). All samples were obtained under approved Human Subjects IRB protocols.

### Profiling and analysis of hydroxymethylated DNA

Total DNA was extracted from 4 normal and 4 OA patient chondrocytes and was enriched for 5hmC using a biotin-based streptavidin pull down technique (Hydroxymethyl Collector, Active Motif) originally described by Song et al. (28), that has since been utilized for multiple studies (24, 29, 30). Libraries were prepared using 500ng of 5hmC enriched DNA using the NEBNext kit (NEB) and were sequenced on an Illumina HiSeq 2000 with 1×50bp reads.

An iterative quality check and filtering procedure was performed to obtain good quality reads after the trimming of the initial 6 bases. The Burrows-Wheeler Aligner, v0.7.5a r405 (31), was used with default parameters to align the filtered reads to the human reference genome (hg19) and unique alignments with a MAPQ score >5 were used for downstream analyses. The total number of reads with a MAPQ score >5 (and overall total reads) were as follows: 24 weeks = 58817869 (78797560), 6 months = 25354632 (40706975), 27 years = 27261745 (38573842), 34 years = 29616025 (47588394), 61 years = 29256816 (34634218), 65 years = 21372649 (25125432), 73 years = 23543180 (26754181), 74 years = 37563182 (51553357).

To identify regions that gained or lost 5hmC at specific DhMRs, the normal samples were set as the control group and the OA samples as the treated group, and diffReps (32) was used with default settings. DiffReps normalizes each sample by removing regions of low read counts and then calculates a normalization ratio for each sample based on the remaining reads. The medians of the ratios are then used as normalization factors. To assess differential sites, diffReps uses a negative binomial test on sliding windows and the significant windows are selected by a predefined cutoff ( $p < 1e-4$ ). The significant windows that overlap with each

other are then merged, and the differential sites are used to perform the statistical tests again. The p-value for each differential site and the best p-value for the sliding window within each differential site are reported. Subsequently, DhMRs are classified by diffReps into genomic locations and annotated using the human (hg19) reference genome.

The distribution of 5hmC was visualized using ngs.plot (33) with default settings, and with the UCSC genome browser.

### Gene expression analyses

Total RNA was extracted from 5 normal and 2 OA samples and was run on Human Gene 1.0 ST Arrays (Affymetrix). Data analysis was performed using dChip (34) with default settings. Pathway analysis of differentially expressed genes was performed using MetaCore version 6.21 (Thomson Reuters). Gene co-expression analysis was performed using GeneFriends (<http://genefriends.org>).

We obtained lists of genes that have been identified as having an association with knee OA by GWAS studies from the publically available HuGE Navigator, and of genes reported to display increased gene expression in late stage OA (35). These lists were intersected with the DhMRs we identified in the OA patients to identify subsets of genes where 5hmC may be playing a functional role.

## Results

### A genome wide increase in 5hmC levels is observed in osteoarthritic chondrocytes

We recently reported a global increase in 5hmC levels in OA chondrocytes compared to normal chondrocytes (32 OA and 17 normal samples) (26) using multiple approaches. To build upon these initial studies, we interrogated the global distribution of 5hmC in a small cohort of normal (n = 4) and OA (n = 4) chondrocytes with the goal of analyzing key genes and pathways that gain or lose 5hmC in OA chondrocytes when compared to normal chondrocytes. Following chemical conjugation and affinity purification of 5hmC-enriched sequences (28), we performed high-throughput DNA sequencing on DNA enriched for 5hmC from normal and OA chondrocytes, to determine the exact location and distribution of 5hmC. 5hmC enrichment was confirmed by including 5hmC positive control DNA that was recovered at the recommended efficiency (Hydroxymethyl Collector, Active Motif). Analysis of the genome wide 5hmC distribution identified dramatic, global 5hmC increases in OA chondrocytes compared to normal chondrocytes, consistent with our previous findings. A high degree of correlation was observed between samples, with Pearson correlation coefficients of gene body 5hmC levels ranging from 0.93 - 0.99 for the normal samples and from 0.57 - 0.91 for the OA samples (Table S1).

Upon examining the average profiles of 5hmC in normal and OA chondrocytes, we identified a total of 70591 DhMRs using diffReps (32), with 44288 DhMRs increased and 26303 DhMRs decreased in the OA chondrocytes, compared to the normal chondrocytes (Tables S2 and S3). To verify the DhMRs identified by the peak calling software we independently examined the 5hmC profiles of the samples using the UCSC genome browser (Figure 1A). Differential peaks were observed in the promoters of MMPs 1 and 3 that we

previously reported using an independent locus-specific technique based on glycosylation and restriction enzyme digestion that can accurately distinguish 5mC from 5hmC (26) which further validated our sequencing data.

### **5hmC gain is preferentially associated with gene bodies**

Analysis of the genomic locations of DhMRs in OA chondrocytes revealed that the greatest gain of 5hmC occurred in gene bodies, other intergenic regions and within the promoter region (Figure 1B and 1C). Previous reports of 5hmC distribution in ESCs, neurons and blood lineages (22, 24, 25, 28, 36, 37) have similarly demonstrated that the majority of the stable 5hmC peaks reside in the gene bodies. While the majority of DhMRs lost were also from gene bodies and intergenic regions, DhMRs gained (29,110) were significantly greater than the DhMRs lost (10,006) in gene bodies. Conversely, DhMRs lost or gained in intergenic regions were comparable (15,252 and 11,386 respectively) (Figure 1C). Although smaller in number, the DhMRs gained within 3kb of the regulatory promoter regions were 3.7-fold higher than the DhMRs lost.

To further analyze the DhMRs in OA chondrocytes, we performed pathway analyses on the DhMRs with 2-fold increase in 5hmC in the gene body or within 3kb of the promoter in the OA chondrocytes when compared to the normal chondrocytes. Increased DhMRs in the promoter regions were found in Wnt signaling and bone remodeling pathways (Figure 1D), processes that have been implicated in OA pathology. Since the number DhMRs increased in gene bodies was very high (20,766 DhMRs) we focused on the top 10% of these DhMR-containing genes, which represented pathways related to Wnt signaling, cell growth regulation and transcriptional regulation (Figure 1E). Interestingly, Wnt signaling was observed to be a key pathway with DhMRs gained in both gene bodies and promoter regions. On the other hand, DhMRs lost in OA chondrocytes were in genes associated with cell adhesion, skeletal muscle development and cytoskeletal rearrangement (Figure 1F and G).

### **Genes activated (but not repressed) in OA demonstrate 5hmC gain in gene bodies**

To examine the gene expression changes between the normal and OA chondrocytes we initially performed microarrays on a small cohort of 5 normal and 2 OA samples. The samples showed good correlation (Figure 2A) and the identified gene expression changes agreed with the literature (35, 38). The array results were validated by an independent real-time PCR based estimation of the expression of key OA genes; MMPs 3 and 13 showed much higher expression in OA samples compared to normal samples (data not shown). The top 30 DhMRs within the gene body related to increased gene expression are shown in Table 1. Pathway analyses of the genes with 1.5-fold upregulation in the OA chondrocytes revealed OA related pathways including inflammatory responses and Notch signaling (Figure 2B). Genes that were 1.5-fold downregulated in OA chondrocytes when compared to normal chondrocytes were related to cell cycle pathways and neural signaling (Figure 2C). To assess the role 5hmC might be playing at a transcriptional level we used ngs.plot (39) to visualize the 5hmC distribution in these subsets of genes. The average profile for the genes with 1.5-fold increase in expression in OA chondrocytes showed increased 5hmC in the promoter region and throughout the gene body compared to the normal chondrocytes

(Figure 2D), whereas the average profile for the genes with 1.5-fold decrease in expression in OA chondrocytes showed a sharp increase in 5hmC in the gene promoter region but similar 5hmC throughout the gene body in both OA and normal chondrocytes (Figure 2E). The presence of 5hmC in gene bodies therefore appears to be specific to genes that are activated (but not repressed) in OA and thus, has a strong positive correlation with gene activation. No such correlation was apparent between the DhMRs within the promoter region (within 2kb of the transcriptional start site) for activated or repressed genes (Figure 2D and E).

Since our gene expression data was from a very small cohort, drawing definitive conclusions from such a small data set (especially in the case of a heterogeneous disease such as OA) is challenging. Therefore, we extended the analyses to a published dataset in which gene expression profiling was performed on a larger cohort with chondrocytes isolated from the knee cartilage of 78 normal and diseased patients (35). In this study, Aigner and colleagues identified 56 genes that were upregulated in late-stage OA. Upon examining the 5hmC profile for these genes in our normal and OA chondrocytes, increased 5hmC in the promoter and gene body regions are consistent with the observations of activated genes in our expression dataset (Figure 2F and Table S4). Genes identified by the Aigner study that have decreased expression in OA also had 5hmC profiles consistent with our results (Figure 2G). These analyses from a larger cohort validated the strong co-relation between 5hmC gain in gene bodies and gene activation in OA. However, 5hmC gain is not the sole epigenetic mark associated with gene activation. Out of the genes associated with the top 10% of DhMRs gained in gene bodies, only 16.6% are activated (1.2-fold increased expression in OA), suggesting that other activating marks besides 5hmC are required for gene activation.

### 5hmC distribution in OA susceptibility genes

To identify the role 5hmC might be playing in OA disease onset and progression, we examined the 5hmC status of genes that have previously been implicated in OA pathology by GWAS studies (see Materials and Methods). Of the 103 genes in the HuGE Navigator OA gene list, 51 genes had increased DhMRs in OA chondrocytes, whereas only 9 genes showed DhMRs that were decreased in OA and 43 genes showed no change in 5hmC (Figure 3A, B and Table S5). When we examined these trends in combination, we identified increased DhMRs in approximately 50% of genes that have previously been implicated in OA pathology, while 40% of genes show no change in 5hmC status and only 10% show decreased DhMRs. To calculate the statistical significance of the overlap between the presence of 5hmC and changes in gene expression we performed a two-sample test for equality of proportion, which gave a highly significant p-value of  $2.2e-16$ .

### 5hmC distribution in genes previously identified as being differentially methylated in OA

Several recent reports have explored genome-wide methylation patterns (16, 17, 40) and have identified differential methylation in a subset of genes in OA. However, these studies used sodium bisulfite to assess methylated cytosines, which cannot be used to distinguish methylation from hydroxymethylation (41). As such, the presence of hypermethylation could be due to a mixture of 5mC and 5hmC. We therefore investigated 5hmC distribution in the genes to gain a further understanding of the distribution of 5mC and 5hmC in OA. We

also examined several genes identified as having differential methylation by candidate gene approaches where the role of 5hmC has not been considered (11-14, 42).

Of the genes reported as being hypermethylated in OA, FGFR2, FTO, GDF5, NFATC1 and TGFB2 (but not BMP1R and SNCAIP) (Figure 3C) showed increased DhMRs as per our data, suggesting that the level of methylation may have been overestimated by the studies employing bisulfite-sequencing techniques. ADAMTS10, PBRM1, RUNX2 and SUPT3H display hypomethylation in OA (Figure 3C) but variable 5hmC status (gain and loss of DhMRs). IL1 $\beta$  and IL4R displayed hypomethylation and showed concomitant decreases in DhMRs, confirming the demethylation of these genes in OA. Interestingly, ADAMTS17, ADAMTS5, COL1A2, DLX5, FILIP1, HTRA1, MMP3, RUNX1 and TGFB1 are hypomethylated in OA but also exhibit increases in hydroxymethylation, suggesting that loss of 5mC and gain of 5hmC may act as ‘dual’ activating marks in these genes. Consistent with this observation, FURIN, HTRA1, MMP3 and RUNX1, clearly showed hypomethylation outside of the gene body (i.e. in the 5'UTR or promoter region) whereas the gain in hydroxymethylation was observed within the gene body. Conversely for DLX5, hypomethylation was observed in the gene body whereas 5hmC gain was identified in the promoter region. Overall, these genes represent a subset of key OA genes regulated by the DNA methylation-hydroxymethylation-demethylation dynamics in OA.

### Genes and pathways associated with increased DhMRs and increased gene expression

Next, we examined which gene pathways were upregulated as a result of the aberrant 5hmC gain in OA (Figure 4A). The top 30 DhMRs within the gene body that are related to increased gene expression (1.5-fold) are shown in Table 1. These genes include known chondrogenic factors like Sox5 as well as genes with unknown functions in cartilage biology or OA like PLXDC2, recently identified as a mitogen in neural progenitors (43) and NEBL, an actin-binding protein (44). To elucidate the potential roles of PLXDC2 and NEBL in OA, we performed gene co-expression analyses (see Materials and Methods). While PLXDC2 did not appear to be co-expressed with any genes known to be involved in cartilage biology, Sox9 (the master transcriptional regulator of chondrogenesis) was found to be co-expressed with NEBL ( $p = 0.626$ ) (Tables S6 and S7).

We relaxed our stringency of fold change in gene expression to 1.2-fold to include a greater number of genes and examined the cross-section of these genes with the 2-fold change in DhMRs in OA (Table S8). We performed pathway analyses on the subset of genes that had 2-fold increase in DhMRs within 3k of the promoter region or in the top 10% of DhMRs in the gene body (and had an associated increase in gene expression (1.2-fold)). Both analyses revealed similar pathways: Notch signaling, hedgehog signaling, ossification and bone remodeling, Wnt signaling and cartilage development (Figure 4B and C). We also examined the genes with decreased DhMRs and decreased gene expression in OA, and identified pathways related to cell cycling and skeletal muscle development (Figure 4D, E and F). Since aberrant Notch and Wnt signaling are known to play a role in the onset of OA, we explored these pathways further. Of the 236 proteins present in the Notch signaling pathway, we identified 17 genes with elevated 5hmC and increased gene expression in OA (Figure 5A and B). Similarly, of the 177 proteins present in the Wnt signaling pathways, 16 were

identified as having elevated 5hmC and increased gene expression in OA. To calculate the statistical significance of these overlaps we performed a two-sample test for equality of proportion, which gave highly significant p-values of 0.003766 and 0.005744 for the co-occurrence of DhMRs and elevated expression in the Wnt and Notch pathways, respectively.

## Discussion

The importance of epigenetic regulation in the gene expression dynamics in OA has become increasingly evident in recent years (5, 7). Genome-wide and loci-specific interrogation of DNA methylation patterns in OA cartilage have identified conserved differentially methylated regions (DMRs) in key OA-specific genes (15-17). Cytosine methylation (5mC) is now known to undergo further modifications by the TET enzymes to 5hmC, 5fC and 5caC eventually leading to DNA demethylation (18-20), and the effect of these stable modifications on OA gene expression is of great interest. Our previous studies had demonstrated a global increase in DNA hydroxymethylation in OA chondrocytes as compared to normal chondrocytes, with higher 5hmC levels associated with MMP1 and 3 promoters (26). In the present study, we report genome-wide analyses of 5hmC distribution and gene expression changes in normal and OA chondrocytes to decipher the effect of the 5hmC gain on OA-specific gene expression.

Recent reports have suggested that 5hmC is an activating epigenetic mark, with enrichment of 5hmC in the gene body showing a strong correlation with increased gene transcription in multiple cell types (24, 25, 37). Several studies have suggested a functional role for 5hmC in neurological diseases (reviewed in 45) and multiple cancers (46). These findings indicate that the correlation between elevated 5hmC in gene bodies and increased transcription is robust and that 5hmC plays a role in gene activation.

Our data demonstrates that the most striking 5hmC gain is observed in the gene bodies of activated genes in OA, highlighting that 5hmC plays a regulatory role in OA gene expression. This finding adds more weight to the notion that the presence of 5hmC in gene bodies has positive effects on transcription and that 5hmC dynamics may play a key role in the transcriptional shift that occurs with disease onset and progression. Analysis of DhMRs in genes that have been previously implicated in OA (35) and in our own gene expression data, revealed approximately 50% of genes associated with OA have increased 5hmC in the disease state. It is, however, important to note that while the genes activated in OA show increased 5hmC in gene bodies, not all genes that gain 5hmC in gene bodies or elsewhere are activated. As such, while enrichment of 5hmC does appear to be playing a regulatory role in a subset of genes involved in the OA pathology, there are clearly other contributing factors (and other epigenetic changes) that will be important to delineate.

### Age-associated 5hmC dynamics and the effect on OA

It is difficult to ascertain in end-stage OA chondrocytes whether the 5hmC gain is associated only with the disease and is not simply due to aging. In our previous report, we had addressed this question by comparing global 5hmC levels in normal and OA chondrocytes from similar age groups, with an abrupt increase in 5hmC levels observed in OA patients with no gradual increase with age (26). In the present studies, we have examined this



question in an alternate manner. Since the 5hmC gain is clearly observed in genes activated specifically in OA, we reasoned that if 5hmC gain were age-associated, we would observe 5hmC gain in genes activated in aged cartilage. Although not in human subjects, a recent study (47) identified genes that displayed transcriptional changes in aged equine cartilage when compared to young cartilage. Intersection of genes with increased DhMRs in OA with genes showing changed expression with aging identified 6% of the activated genes and 9% of repressed genes to contain a notable number of DhMRs (>10). Interestingly, no common genes were identified between genes activated with aging in equine cartilage (47) and genes activated in human OA (from Aigner et al. 2006 (35)) (data not shown). Instead, 14 genes were identified that decrease in expression during normal equine aging but increase in expression in human OA, including multiple collagen genes (COL1A1, COL1A2, COL2A1, COL3A1, COL4A1, COL5A1, COL11A1). Therefore, no direct effect of 5hmC could be identified on age-associated gene expression changes within these analyses.

### **OA genes and pathways potentially regulated by DNA hydroxymethylation**

Our studies have identified various key OA-associated genes and pathways that appear to be differentially hydroxymethylated in OA cartilage, suggesting a potential regulation by DNA hydroxymethylation. These include the Notch, Wnt and hedgehog signaling pathways, which have all been previously implicated in OA onset or disease severity (48-50). Of the 103 known OA susceptibility genes that were examined, 51 genes showed DhMRs that were increased in OA chondrocytes, which included the key OA genes MMP3, COL2A1, GDF5 and COMP. In addition, some of the genes exhibiting differential hydroxymethylation were also identified previously in genome-wide DNA methylation studies between normal and OA patients. A caveat in these studies was the use of sodium bisulfite that cannot distinguish between 5mC and 5hmC. Indeed some of the genes reported to be hypermethylated, FGFR2, FTO, GDF5, NFATC1 and TGFB2 showed increased DhMRs as per our data, suggesting that the level of methylation may have been overestimated in these genes. Nevertheless, these overlaps also validate our analyses. Additionally, we observe 5hmC gain in the gene bodies of ADAMTS17, ADAMTS5, COL1A2, DLX5, FILIP1, HTRA1, MMP3, RUNX1 and TGFB1 (which are hypomethylated in OA), suggesting that demethylation and 5hmC gain may act as 'dual' activating marks in these genes.

In conclusion, we demonstrate that there are dynamic changes in the chondrocyte hydroxymethylome in OA, allowing it to be very clearly distinguished from normal chondrocytes. Our findings, which identify enrichment of 5hmC in many key genes and pathways in OA, suggest that the increased activation of these pathways could, at least in part, be a result of the 5hmC-associated transcriptional activation. A greater understanding of the regulatory role of DNA methylation/hydroxymethylation-demethylation dynamics in the regulation of OA-specific gene expression will provide further insights into the onset and progression of this widespread disease. In addition, it would be interesting to test in future studies whether modulation of 5hmC levels by inhibition of TET activity could inhibit OA-specific gene expression and hence be therapeutic in OA.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

We are grateful to Drs. Stuart Goodman and Jason Dragoo for their kind assistance in procurement of cartilage samples from patients undergoing surgery. We would like to thank Gary Mantalas of the Stanford University School of Medicine Stem Cell Institute Genome Center for performing the Illumina HiSeq runs and Natalia Kosovilka of the Stanford University School of Medicine Protein and Nucleic Acid Facility for running the microarrays. We are also thankful to Dr. Mandy Peffers of Liverpool University for sharing her data.

Grants/financial support:

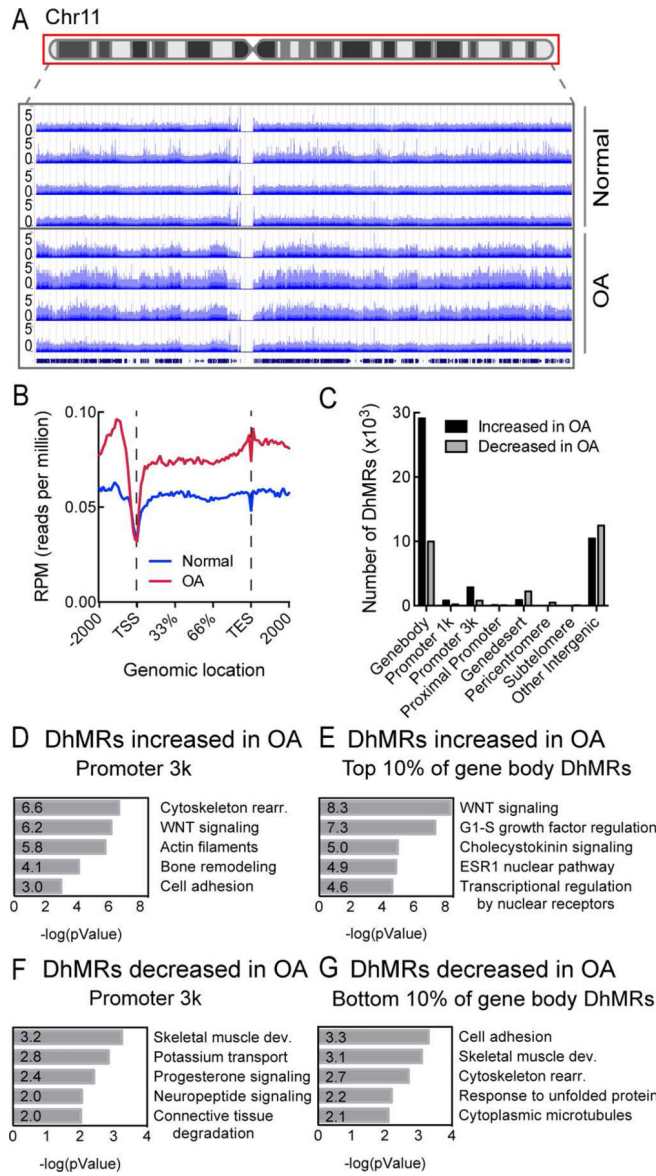
No financial support or other benefits were obtained from any commercial sources for this study and the authors declare that they have no competing financial interests. This work was supported by an Innovative Research Grant from Arthritis Foundation (to N.B.). Y.L. and W.W. are partially supported by NIH grants R01 GM10983601 and R01 HG007834.

## References

1. Aigner T, Richter W. OA in 2011: Age-related OA--a concept emerging from infancy? *Nature reviews Rheumatology*. 2012; 8(2):70–2. [PubMed: 22231238]
2. Blagojevic M, Jinks C, Jeffery A, Jordan KP. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. *Osteoarthritis and Cartilage*. 2010; 18(1):24–33. [PubMed: 19751691]
3. Spector T. Risk factors for osteoarthritis: genetics. *Osteoarthritis and Cartilage*. 2004; 12:39–44.
4. Valdes AM, Spector TD. The genetic epidemiology of osteoarthritis. *Current Opinion in Rheumatology*. 2010; 22(2):139–143. [PubMed: 20090528]
5. Gonzalez A. Osteoarthritis year 2013 in review: genetics and genomics. *Osteoarthritis and Cartilage*. 2013; 21(10):1443–1451. [PubMed: 23845519]
6. Reynard LN, Loughlin J. The genetics and functional analysis of primary osteoarthritis susceptibility. *Expert reviews in molecular medicine*. 2013:15.
7. Blanco FJ, Rego-Pérez I. Editorial: Is It Time for Epigenetics in Osteoarthritis?: Editorial. *Arthritis & Rheumatology*. 2014; 66(9):2324–2327. [PubMed: 24838530]
8. Dawson Mark A, Kouzarides T. Cancer epigenetics: from mechanism to therapy. *Cell*. 2012; 150(1):12–27. [PubMed: 22770212]
9. Jones PA. The role of DNA methylation in mammalian epigenetics. *Science*. 2001; 293(5532):1068–1070. [PubMed: 11498573]
10. Young DA. Editorial: More evidence for a role of CpG methylation in the pathogenesis of osteoarthritis. *Arthritis & Rheumatism*. 2013; 65(3):555–558. [PubMed: 23233319]
11. Hashimoto K, Oreffo ROC, 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 & Rheumatism*. 2009; 60(11):3303–3313. [PubMed: 19877066]
12. 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 & Rheumatism*. 2005; 52(10):3110–3124. [PubMed: 16200590]
13. Cheung K, Hashimoto K, Yamada N, Roach H. Expression of ADAMTS-4 by chondrocytes in the surface zone of human osteoarthritic cartilage is regulated by epigenetic DNA de-methylation. *Rheumatology International*. 2009; 29(5):525–534. [PubMed: 18941754]
14. 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-1 (IL1B) genes in chondrocytes depends on methylation of specific proximal promoter CpG sites. *Journal of Biological Chemistry*. 2013; 288(14):10061–10072. [PubMed: 23417678]

15. 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]
16. 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: methylation profile of OA cartilage. *Arthritis & Rheumatology*. 2014; 66(9):2450–2460. [PubMed: 24838673]
17. 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: genome-wide methylation in OA. *Arthritis & Rheumatology*. 2014; 66(10):2804–2815. [PubMed: 24980887]
18. Bhutani N, Burns DM, Blau HM. DNA demethylation dynamics. *Cell*. 2011; 146(6):866–72. [PubMed: 21925312]
19. Kriaucionis S, Heintz N. The nuclear DNA base, 5-hydroxymethylcytosine is present in brain and enriched in Purkinje neurons. *Science*. 2009; 324(5929):929–930. [PubMed: 19372393]
20. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*. 2009; 324(5929):930–935. [PubMed: 19372391]
21. Bachman M, Uribe-Lewis S, Yang X, Williams M, Murrell A, Balasubramanian S. 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nature Chemistry*. 2014; 6:1049–1055.
22. Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K, et al. Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes & Development*. 2011; 25(7):679–684. [PubMed: 21460036]
23. Stroud H, Feng S, Morey Kinney S, Pradhan S, Jacobsen SE. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biology*. 2011; 12(6):R54–62. [PubMed: 21689397]
24. Hahn Maria A, Qiu R, Wu X, Li Arthur X, Zhang H, Wang J, et al. Dynamics of 5-Hydroxymethylcytosine and chromatin marks in mammalian neurogenesis. *Cell Reports*. 2013; 3(2):291–300. [PubMed: 23403289]
25. Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, et al. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature*. 2011; 473(7347):394–397. [PubMed: 21552279]
26. Taylor SEB, Smeriglio P, Dhulipala L, Rath M, Bhutani N. A global increase in 5-hydroxymethylcytosine levels marks osteoarthritic chondrocytes. *Arthritis & Rheumatology*. 2014; 66(1):90–100. [PubMed: 24449578]
27. Smith RL, Lindsey DP, Dhulipala L, Harris AH, Goodman SB, Maloney WJ. Effects of intermittent hydrostatic pressure and BMP-2 on osteoarthritic human chondrocyte metabolism in vitro. *J Orthop Res*. 2011; 29(3):361–8. [PubMed: 20882590]
28. Song C-X, Szulwach KE, Fu Y, Dai Q, Yi C, Li X, et al. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nature Biotechnology*. 2011; 29(1):68–72.
29. Thomson JP, Hunter JM, Nestor CE, Dunican DS, Terranova R, Moggs JG, et al. Comparative analysis of affinity-based 5-hydroxymethylation enrichment techniques. *Nucleic Acids Research*. 2013; 41(22):e206–e206. [PubMed: 24214958]
30. Ivanov M, Kals M, Kacevska M, Barragan I, Kasuga K, Rane A, et al. Ontogeny, distribution and potential roles of 5-hydroxymethylcytosine in human liver function. *Genome Biol*. 2013; 14(8)
31. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009; 25(14):1754–1760. [PubMed: 19451168]
32. Shen L, Shao N-Y, Liu X, Maze I, Feng J, Nestler EJ. diffReps: Detecting differential chromatin modification sites from ChIP-seq data with biological replicates. *PLoS ONE*. 2013; 8(6)
33. Shen L, Shao N, Liu X, Nestler E. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics*. 2014; 15(1):284. [PubMed: 24735413]

34. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proceedings of the National Academy of Sciences*. 2001; 98(1):31–36.
35. Aigner T, Fundel K, Saas J, Gebhard PM, Haag J, Weiss T, et al. Large-scale gene expression profiling reveals major pathogenetic pathways of cartilage degeneration in osteoarthritis. *Arthritis & Rheumatism*. 2006; 54(11):3533–3544. [PubMed: 17075858]
36. Ficiz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature*. 2011; 473(7347):398–402. [PubMed: 21460836]
37. Madzo J, Liu H, Rodriguez A, Vasanthakumar A, Sundaravel S, Caces Donne Bennett D, et al. Hydroxymethylation at gene regulatory regions directs stem/early progenitor cell commitment during erythropoiesis. *Cell Reports*. 2014; 6(1):231–244. [PubMed: 24373966]
38. Chou CH, Lee CH, Lu LS, Song IW, Chuang HP, Kuo SY, et al. Direct assessment of articular cartilage and underlying subchondral bone reveals a progressive gene expression change in human osteoarthritic knees. *Osteoarthritis and Cartilage*. 2013; 21(3):450–461. [PubMed: 23220557]
39. Shen L, Shao N, Liu X, Nestler E. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics*. 2014; 15(1)
40. den Hollander W, Ramos YFM, Bos SD, Bomer N, van der Breggen R, Lakenberg N, et al. Knee and hip articular cartilage have distinct epigenomic landscapes: implications for future cartilage regeneration approaches. *Annals of the Rheumatic Diseases*. 2014; 0:1–5.
41. Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS ONE*. 2010; 5(1)
42. Reynard LN, Bui C, Canty-Laird EG, Young DA, Loughlin J. Expression of the osteoarthritis-associated gene GDF5 is modulated epigenetically by DNA methylation. *Human Molecular Genetics*. 2011; 20(17):3450–3460. [PubMed: 21642387]
43. Miller-Delaney SFC, Lieberam I, Murphy P, Mitchell KJ. Plxdc2 is a mitogen for neural progenitors. *PLoS one*. 2011; 6(1)
44. Cósér VM, Meyer C, Basegio R, Menezes J, Marschalek R, Pombo-de-Oliveira MS. Nebulette is the second member of the nebulin family fused to the MLL gene in infant leukemia. *Cancer Genetics and Cytogenetics*. 2010; 198(2):151–154. [PubMed: 20362230]
45. Sun W, Zang L, Shu Q, Li X. From development to diseases: The role of 5hmC in brain. *Genomics*. 2014; 104(5):347–351. [PubMed: 25205306]
46. Rampal R, Alkalin A, Madzo J, Vasanthakumar A, Pronier E, Patel J, et al. DNA hydroxymethylation profiling reveals that WT1 mutations result in loss of TET2 function in acute myeloid leukemia. *Cell Reports*. 2014; 9(5):1841–1855. [PubMed: 25482556]
47. Peffer MJ, Liu X, Clegg PD. Transcriptomic signatures in cartilage ageing. *Arthritis research & therapy*. 2013; 15(4)
48. Hosaka Y, Saito T, Sugita S, Hikata T, Kobayashi H, Fukai A, et al. Notch signaling in chondrocytes modulates endochondral ossification and osteoarthritis development. *Proceedings of the National Academy of Sciences*. 2013; 110(5):1875–1880.
49. Loughlin J, Dowling B, Chapman K, Marcelline L, Mustafa Z, Southam L, et al. Functional variants within the secreted frizzled-related protein 3 gene are associated with hip osteoarthritis in females. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101(26):9757–9762. [PubMed: 15210948]
50. Lin AC, Seeto BL, Bartoszko JM, Khoury MA, Whetstone H, Ho L, et al. Modulating hedgehog signaling can attenuate the severity of osteoarthritis. *Nature Medicine*. 2009; 15(12):1421–1425.



**Figure 1. Genome-wide increases in 5hmC are associated with osteoarthritic (OA) chondrocytes**  
 A. Visualization of the 5hmC distribution across a representative chromosome in normal (donor ages = 24 weeks, 6 months, 27 and 34 years) and OA chondrocytes (donor ages = 61, 65, 73 and 74 years). B. Composite profile of the 5hmC distribution averaged across all genes in representative normal and OA chondrocyte samples. TSS = transcriptional start site, TES = transcriptional end site. C. The number and genomic location of the DhMRs that are increased and decreased in OA chondrocytes when compared to normal chondrocytes. ‘Other Intergenic’ = DhMR identified was not located in any of the other genomic locations. D. E. Pathway analysis of the genes with DhMRs with 2-fold increase within 3kb of the transcriptional start site (Promoter 3k) (D) and in the top 10% of DhMRs with a 2-fold increase in the gene body (E) in OA when compared to normal chondrocytes. F. G. Pathway analysis of the genes with DhMRs with 2-fold decrease within 3kb of the transcriptional start site (Promoter 3k) (F) and in the bottom 10% of DhMRs with 2-fold decrease in the

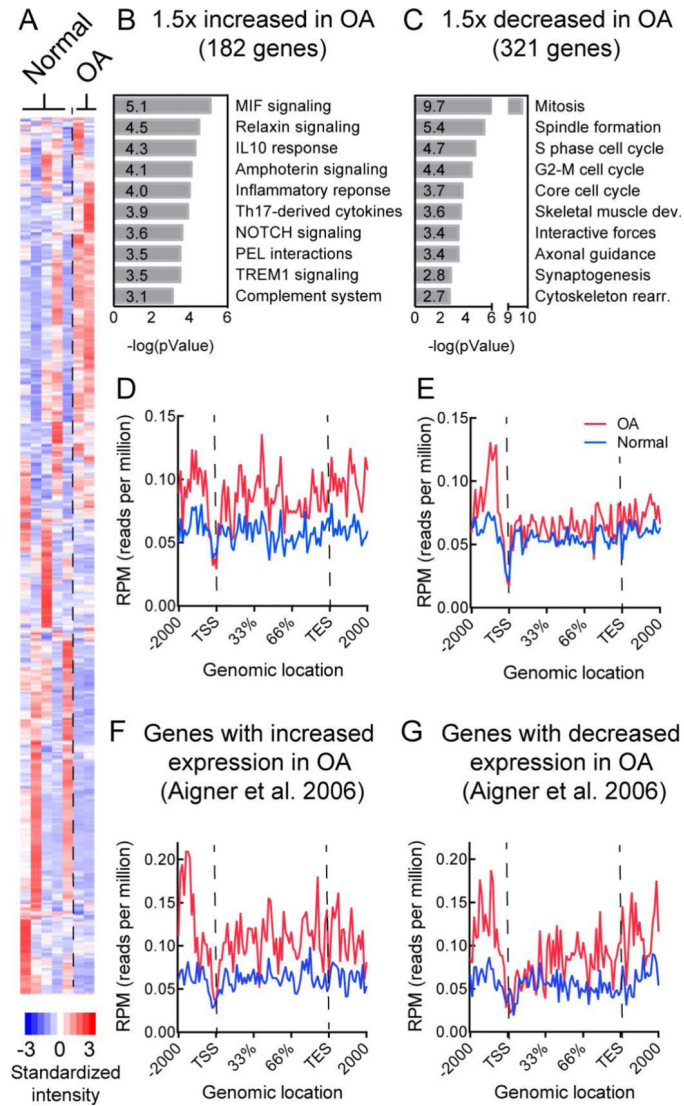
gene body in OA when compared to normal chondrocytes. Corresponding p-values are shown.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 2. 5hmC is enriched in the gene bodies of genes activated in OA**

A. Heatmap of global gene expression changes between normal (donor ages = 24 weeks, 6 months, 18 months, 27 and 34 years) and OA (donor ages = 64 and 74 years) chondrocytes. Red = higher than mean expression, blue = lower than mean expression.

B. C. Pathway analysis of the genes with 1.5-fold increases in expression (activated genes) (B) and of the genes with 1.5-fold decreases in expression (repressed genes) in OA (C).

PEL = Platelet-endothelialleucocyte, dev = development, rearr = rearrangement.

Corresponding p-values are shown. D. E. Composite profiles of the average 5hmC distributions in normal (blue) and OA (red) chondrocytes in genes with a 1.5-fold increase in expression (activated genes) (D) and in genes with a 1.5-fold decrease in expression (repressed genes) in OA when compared to normal chondrocytes.

F. G. Composite profile of the average 5hmC distribution in normal (blue) and OA (red) chondrocytes in genes identified by Aigner and colleagues, 2006 (35) as being activated in OA (2-fold increase in expression) (F) and repressed in OA (2-fold decrease in expression) (G).

TSS = transcriptional start site, TES = transcriptional end site.

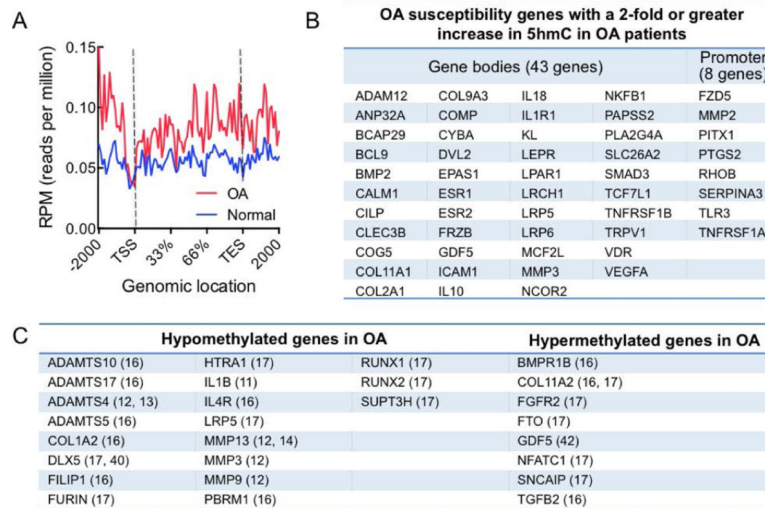
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



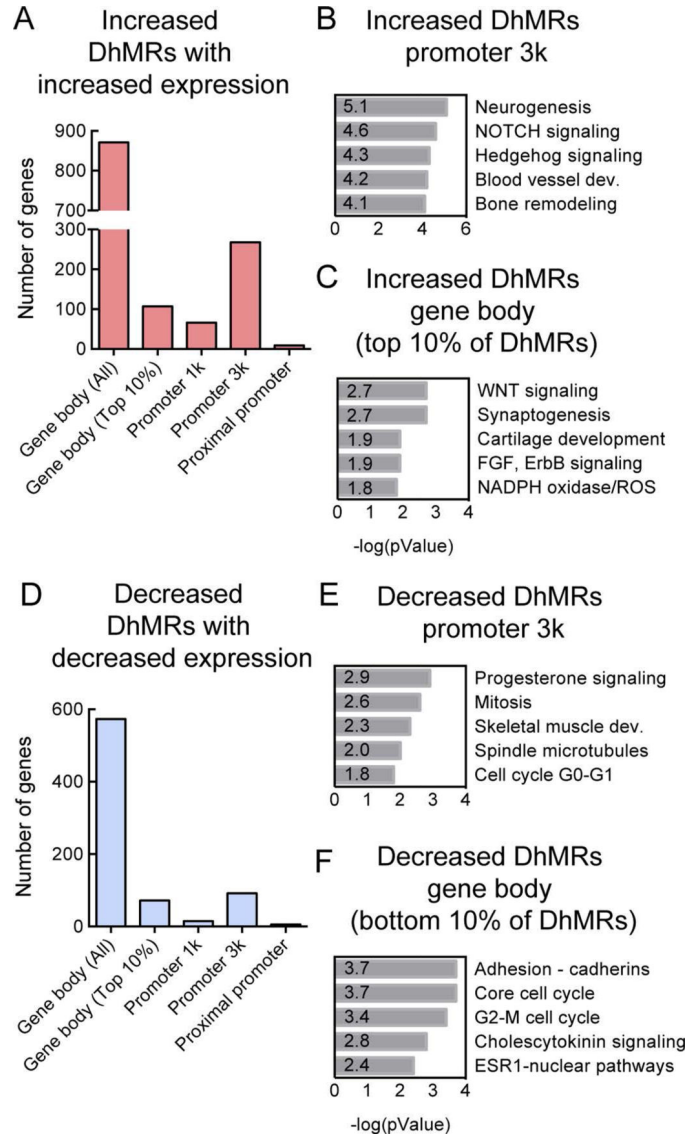


### Figure 3. OA susceptibility genes display increases in 5hmC

A. Composite profile of the average 5hmC distribution in normal (blue) and OA (red) chondrocytes in genes identified as being associated with OA by GWAS studies. TSS = transcriptional start site, TES = transcriptional end site.

B. OA susceptibility genes, identified by GWAS studies that have a 2-fold increase in 5hmC in OA patients. Out of a total of 103 genes, 52 genes displayed a 2-fold increase in 5hmC.

C. Compiled list of genes that have been previously identified as being hypomethylated or hypermethylated in OA. The related study is indicated in parentheses.



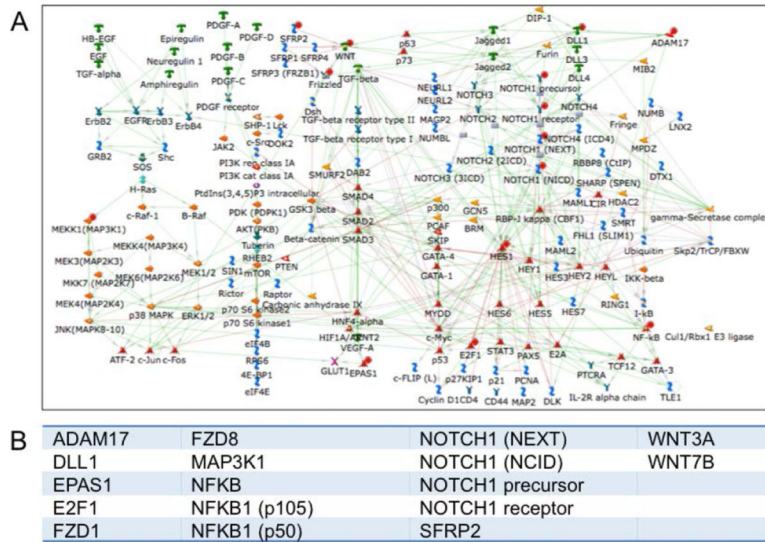
**Figure 4. Pathways and networks known to be perturbed in OA are regulated by aberrant 5hmC gain in the disease state**

A. The genomic locations of DhMRs that show a 2-fold increase in OA chondrocytes when compared to normal chondrocytes in genes that have a 1.2-fold increase in gene expression in OA. Gene body (All) = All DhMRs, Gene body (Top 10%) = Top 10% of DhMRs in gene bodies.

B. C. Pathway analysis of the genes that have increased expression and increased DhMRs within 3kb of the promoter (B) or in the top 10% of DhMRs in the gene body (C) in OA. The top 5 networks, ranked by p-value, are shown.

D. The genomic locations of DhMRs that show a 2-fold decrease in OA when compared to normal chondrocytes in genes that have a 1.2-fold decrease in gene expression in OA.

E. F. Pathway analysis of the genes that have decreased expression and decreased DhMRs within 3kb of the promoter (E) or in the bottom 10% of DhMRs that are decreased in the gene body (F) in OA. The top 5 networks, ranked by p-value, are shown.



**Figure 5. 5hmC containing genes in the Notch signaling pathway in OA**

A. The Notch signaling network. Genes with 2-fold increased 5hmC and 1.2-fold increased expression in OA are marked with a red dot.

B. A list of genes in the Notch signaling network that have 2-fold increased 5hmC and 1.2-fold increased expression in OA chondrocytes when compared to normal chondrocytes.

**Table 1**

The top 30 DhMRs within the gene body related to increased gene expression.

Gene name	Transcript name	Strand	Location	Distance to TSS	FC DhMR	p-value of DhMR	FC gene expression
SOX5	NM_152989	-	Genebody	522079.5	20.82	3.87E-08	3.11
PLXDC2	NM_032812	+	Genebody	245779.5	12.82	4.21E-08	3.18
PLXDC2	NM_032812	+	Genebody	329079.5	11.96	9.62E-08	3.18
NEBL	NM_001173484	-	Genebody	5665.5	10.34	4.95E-08	2.76
SOX5	NM_152989	-	Genebody	78779.5	8.57	1.93E-07	3.11
ST6GALNAC4	NM_175039	-	Genebody	8454.5	8.51	1.46E-07	1.67
TMEM204	NM_024600	+	Genebody	3670.5	8.51	1.47E-07	2.28
PLXDC2	NM_032812	+	Genebody	218779.5	7.06	1.14E-07	3.18
TMEM204	NM_024600	+	Genebody	8520.5	6.92	2.30E-07	2.28
PTPRG	NM_002841	+	Genebody	259758.5	6.63	1.68E-07	1.61
SOX5	NM_006940	-	Genebody	37836.5	6.50	2.74E-07	3.11
CDH23	NM_052836	+	Genebody	241510.5	6.41	1.37E-07	1.65
PLXDC2	NM_032812	+	Genebody	82429.5	6.32	2.54E-07	3.18
COL5A1	NM_000093	+	Genebody	164899.5	6.19	1.43E-07	1.84
IVNS1ABP	NM_006469	-	Genebody	10460.5	6.19	2.24E-07	2.4
DCHS1	NM_003737	-	Genebody	32723.5	6.11	1.65E-07	1.5
SOX5	NM_006940	-	Genebody	273436.5	6.11	2.56E-07	3.11
SOX5	NM_152989	-	Genebody	540179.5	6.06	2.36E-07	3.11
RALGDS	NM_006266	-	Genebody	8260.5	5.82	1.28E-07	1.53
RNF44	NM_014901	-	Genebody	5870.5	5.78	9.65E-08	1.56
SPSB1	NM_025106	+	Genebody	27010.5	5.58	9.63E-08	2.27
RASA3	NM_007368	-	Genebody	87094.5	5.54	1.24E-07	1.64
SMOC2	NM_022138	+	Genebody	15370.5	5.50	1.90E-07	3.78
METRNL	NM_001004431	+	Genebody	6834.5	5.24	1.15E-07	1.52
ROR2	NM_004560	-	Genebody	134443.5	5.24	2.25E-07	4.2
RARB	NM_000965	+	Genebody	92967.5	5.21	4.16E-07	2.64
C9orf3	NM_001193329	+	Genebody	203520.5	5.13	5.76E-08	1.91
SFMBT2	NM_001018039	-	Genebody	164302.5	5.10	1.45E-07	1.73
WTAP	NM_152857	+	Genebody	6281.5	5.10	2.16E-08	1.55

The top 30 sites with high 5hmC ( 2-fold increase) and with increased expression ( 1.5-fold increase) in the gene body in OA chondrocytes when compared to the normal chondrocytes. TSS = transcriptional start site, FC = fold change, DhMR = differentially hydroxymethylated region.