Multi-Tissue Epigenetic and Gene Expression Analysis Combined With Epigenome Modulation Identifies RWDD2B as a Target of Osteoarthritis Susceptibility

Eleanor Parker,1 Ines M. J. Hofer,1 Sarah J. Rice,1 Lucy Earl,1 Sami A. Anjum,1 David J. Deehan,2 and John Loughlin1

Objective. Osteoarthritis (OA) is polygenic, with more than 90 risk loci currently mapped, including at the single-nucleotide polymorphism rs6516886. Previous analysis of OA cartilage DNA identified 6 CpG dinucleotides with methylation levels that correlated with the rs6516886 genotype, forming methylation quantitative trait loci (mQTLs). We undertook this study to investigate these mQTLs and to map expression quantitative trait loci (eQTLs) across joint tissues in order to identify a particular gene as a target of the rs6516886 association effect.

Methods. Nucleic acids were extracted from the cartilage, fat pad, synovium, and peripheral blood from OA patients. Methylation of CpGs and allelic expression imbalance of potential target genes were assessed by pyrosequencing. A chondrocyte cell line expressing deactivated Cas9 (dCas9)–TET1 was used to directly alter CpG methylation levels, with effects on gene expression quantified by polymerase chain reaction.

Results. Multiple mQTLs were detected, with effects strongest in joint tissues and methylation at CpG cg20220242 correlating most significantly with the rs6516886 genotype. CpG cg20220242 is located upstream of RWDD2B. Significant methylation levels at CG dinucleotides (CpGs) in RWDD2B codes for a protein about 28 kDa. Modulating CpG methylation reverses the impact of the risk allele. RWDD2B codes for a protein about which little is currently known. Further analysis of RWDD2B as a target of OA genetic risk will provide novel insight into this complex disease.

INTRODUCTION

Osteoarthritis (OA) is the most common musculoskeletal disorder, affecting 250 million people worldwide (1). OA is characterized by loss of articular cartilage in addition to other joint disruptions, including osteophyte formation, inflammation of the synovium, meniscal damage, and bone remodeling. These structural changes lead to chronic pain, decreased quality of life, and comorbidities, including type 2 diabetes mellitus, cardiovascular disease, and premature death (2–4). OA is polygenic, and multiple genome-wide association studies (GWAS) have been conducted to identify osteoarthritis risk loci, with >90 independent signals identified so far in Europeans (5–11). The overwhelming majority of these loci are intergenic or intronic and are known or predicted to mediate their effect via altered target gene expression, acting as expression quantitative trait loci (eQTLs) (12). Additionally, OA-associated polymorphisms have been shown to correlate with methylation levels at CG dinucleotides (CpGs) in cis,

1Eleanor Parker, BSc, MSc, PhD, Ines M. J. Hofer, BSc, MRes, Sarah J. Rice, MRPharmS, MRes, PhD, Lucy Earl, BSc, MRes, Sami A. Anjum, MBBS, MRes, MRCS, John Loughlin, BSc, PhD: Newcastle University, Newcastle-upon-Tyne, UK; 2David J. Deehan, MBChB, MSc, MD, FRCS (Eng), DSc, FRCS (Tr&Orth): Newcastle University Teaching Hospitals NHS Trust, Freeman Hospital, Newcastle-upon-Tyne, UK. No potential conflicts of interest relevant to this article were reported. Address correspondence to Eleanor Parker, BSc, MSc, PhD, or John Loughlin, BSc, PhD, Newcastle University, Biosciences Institute, Newcastle-upon-Tyne NE1 3BZ, UK. Email: eleanorlouiseparker@gmail.com or john.loughlin@ncl.ac.uk. Submitted for publication April 27, 2020; accepted in revised form July 28, 2020.
a mechanism known to regulate gene expression, thereby acting as methylation QTLs (mQTLs) (13). Indeed, several of these OA loci have been shown to correlate with expression in addition to methylation, known as methylation–expression QTLs (meQTLs), where methylation regulates expression (14). It is therefore clear that DNA methylation drives a significant share of the functional consequences of OA genetic risk loci, and identification and investigation of eQTLs, mQTLs, and meQTLs at disease loci can prioritize genes and gene-regulatory elements, for subsequent study, as targets of the association effect.

One such identified OA risk locus is marked by single-nucleotide polymorphism (SNP) rs6516886 (T>A; minor allele frequency [MAF] of 0.29), which is located on chromosome 21q21.2 and where the major allele (T) was found to be associated with both knee and hip OA in Europeans at genome-wide significance (8). Previously, we have reported that the genotype at rs6516886 correlated with methylation levels at 6 CpG sites, using an Illumina Infinium HumanMethylation450 genome-wide CpG array and cartilage DNA samples from 87 patients (15). The CpGs were located at an interval of <90 kb and mapped upstream and downstream of rs6516888, with the T risk allele associated with increased methylation at 5 CpGs, (cg00065302, cg05468028, cg18001427, cg20220242, and cg24751378) and with decreased methylation at the most distal CpG, cg16140273 (15). The associated region encompasses several genes, including the following: LTN1, which codes for E3 ubiquitin-protein ligase (listerin); RWDD2B, which codes for RWD domain–containing protein 2B; USP16, which codes for ubiquitin carboxyl-terminal hydrolase 16; CCT8, which codes for chaperonin containing TCP1 subunit 8; and MAP3K7CL, which codes for MAP3K7 C-terminal–like protein.

Although cartilage is the central tissue involved in OA, there is joint-wide involvement of other tissues in the disease etiology (16). Therefore, to determine the potential functional role of rs6516886 in OA, we aimed to investigate whether the identified mQTLs are active across different joint tissues and, further, to identify potential gene targets of the association signal to understand how this locus increases OA risk.

**MATERIALS AND METHODS**

**Patient samples.** Joint tissue samples, including cartilage, synovium, and infrapatellar fat pad, were obtained from 348 patients who had primary hip or knee OA and who had undergone joint replacement surgery at the Newcastle-upon-Tyne NHS Foundation Trust hospitals. The Newcastle and North Tyneside Research Ethics Committee granted ethical approval for the collection, with each donor providing verbal and written informed consent (REC reference no. 14/NE/1212). For 55 patients, peripheral blood samples were also collected prior to surgery, using EDTA vacutainers for DNA extraction and Tempus tubes for RNA extraction (ThermoFisher Scientific). Patient details are available in Supplementary Table 1, available on the Arthritis & Rheumatology website at http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract.

**Nucleic acid extraction from tissue samples.** Joint tissue samples were frozen and ground to a powder using a mixer mill (Retsch Limited) under liquid nitrogen, prior to nucleic acid extraction. Cartilage DNA and RNA were extracted using an EZNA Tissue DNA isolation kit (WWR; Omega Biotek) and TRIzol (Life Technologies), respectively. Nucleic acids from synovium and fat pad were extracted using an EZNA DNA/RNA isolation kit and from blood using a QIAamp DNA blood mini kit (Qiagen) and a Tempus Spin RNA isolation kit (ThermoFisher Scientific).

**Genotyping.** SNPs were genotyped using pyrosequencing, as previously described (12,15). Briefly, pyrosequencing assays were designed using PyroMark assay design SW 2.0 (Qiagen), polymerase chain reactions (PCRs) were performed using a G-Storm GS4 Q4 Quad Block Thermal Cycler (Somerton Biotechnology Centre), and sequencing was performed using the PyroMark Q24 Advanced platform and reagents kit, according to the instructions of the manufacturer (Qiagen). Primer sequences are listed in Supplementary Table 2 (http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract).

**Targeted CpG methylation analysis.** DNA samples were bisulfite-converted using an EZ DNA methylation kit according to the instructions of the manufacturer (Zymo Research). CpG site methylation analysis was then performed by pyrosequencing as described above (primer sequences are listed in Supplementary Table 2, http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract). Methylation analysis was performed in duplicate for each sample, with a 5% variance between replicates used as a quality control threshold; samples exceeding this threshold were excluded from analysis. For CpG sites with a global average methylation level of ≤15%, a quality control threshold of 1% was used.

**Chromatin interactions.** The UCSC Genome Browser (http://genome.ucsc.edu/) (17) was mined to identify long-range chromatin interactions stemming from the CpG sites within the locus and the associated SNP. All publicly available Hi-C and long-range chromatin interaction data sets were utilized.

**Quantitative gene expression.** Complementary DNA (cDNA) was synthesized with a SuperScript First-Strand synthesis kit and random hexamers, according to the manufacturer’s instructions (Invitrogen), and quantitative PCR (qPCR) was subsequently performed using predesigned TaqMan assays (Integrated DNA Technologies), as previously described (12). Investigated genes were normalized to the housekeeping genes HPRT1, 18S, and GAPDH, and the relative expression of each gene was calculated using the 2−ΔΔCt method (12).
Allelic expression imbalance (AEI). AEI at transcript SNPs located in the exons of investigated genes was quantified by pyrosequencing (primer sequences are listed in Supplementary Table 2, http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract), as previously described (18). Analysis was performed in triplicate on DNA and cDNA for each sample, and a quality control threshold of 5% variance across replicates was applied. Allelic expression for each sample was produced using PyroMark Advanced software, and the output for cDNA was normalized to that of its respective DNA.

Generation of a stable chondrocyte cell line with inducible deactivated Cas9 (dCas9)–TET1 expression. The PiggyBac 138-dCas9-TET1 vector (19) and modified 137-transposase “helper” vector (20) were kind gifts from the laboratory of R. Jaenisch, MIT, Boston, MA. The PiggyBac vector was digested with NsiI-HF and NdeI enzymes (New England Biolabs) to remove the 2-kb sequence encoding the Neo’/kanamycin resistance gene. Following this, a 1.7-kb sequence, including the puromycin resistance gene and complementary overhangs to the vector backbone, was ligated into the plasmid. The sequence of the resulting construct was confirmed by Sanger sequencing (Source BioScience). The modified dCas9-TET1 and the transposase constructs were transfected into Tc28a2 immortalized chondrocytes by nucleofection (Amaxa 4D; Lonza) at a PiggyBac:helper ratio of 3:1. Post-nucleofection cells were cultured with 1 μg/ml puromycin for 2 weeks. Doxycycline-inducible expression of dCas9-TET1 protein was confirmed and optimized by immunoblotting using a Cas9 antibody (7A9-3A3; Cell Signaling Technology).

Clustered regularly interspaced short palindromic repeat (CRISPR)/dCas9. Guide RNA (gRNA) sequences were designed using the Integrated DNA Technologies design tool (https://eu.idtdna.com/site/order/designtool/index/CRISPR_CUSTOM) to target the region encompassing and flanking cg20220242 (Supplementary Figure 1, http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract). Identified gRNA sequences were assessed for off-target risk and on-target potential using the design tool. Low scoring gRNAs were rejected, resulting in 2 gRNAs being selected for testing. The gRNAs were then purchased as synthetic Alt-R CRISPR-Cas9 CRISPR RNAs (crRNAs) alongside Alt-R CRISPR-Cas9 trans-activating CRISPR RNA (tracrRNA) (Integrated DNA Technologies). A nontargeting Alt-R CRISPR-Cas9 negative control crRNA (Integrated DNA Technologies) was used to exclude TET1 off-target effects.

TC28a2-dCas9-TET1 cells were plated at subconfluence (100,000 cells/well) in 6-well plates (VWR) and allowed to adhere for 24 hours. Deactivated Cas9–TET1 expression was induced by incubation with 2 μM doxycycline hyclate (Sigma-Aldrich) in 2 ml complete medium for a further 24 hours. CRISPR RNAs were annealed to tracrRNA in a 1:1:2 ratio for paired gRNA and at a 1:1 ratio for single guides at 95°C for 5 minutes, while the nontargeting negative control crRNA was annealed to tracrRNA. Complexes of crRNA/tracrRNA were separately combined with Dharmafect1 (Dharmacon) in serum-free medium (PCS-500-030; ATCC). Next, 1.2 ml of the doxycycline-containing medium was aspirated from each well and replaced with 200 μl transfection mix for 24 hours, after which the medium was changed to complete growth medium, and the cells expanded until confluence in T25 flasks (VWR) was reached (between 2 and 5 days). Guide RNA transfection efficiency was previously confirmed using green fluorescent protein–labeled gRNAs (Parker E et al: unpublished observations). Cells were collected and pelleted by trypsinization, and the cell pellets were frozen. RNA was extracted from pellets using a NucleoSpin TriPrep Kit (Macherey-Nagel, supplied by ThermoFisher Scientific), and qPCR was undertaken as described above. DNA was extracted using a PureLink Genomic DNA Purification Kit (Invitrogen, supplied by Thermo Fisher Scientific), and methyl-ation analysis of cg20220242 (and of upstream and downstream flanking CpGs) was performed as described above and using the

Figure 1. In silico investigation of the rs6516886 locus visualized using the UCSC Genome Browser (http://genome.ucsc.edu). A and B, The positions of the 6 CpGs and rs6516886 (A) and Roadmap ChIP-seq data in adipose-derived mesenchymal stem cell–cultured cells and mesenchymal stem cell–derived chondrocyte–cultured cells, displaying peaks in this region (B). Yellow = enhancer; green = strong transcription; dark green = weak transcription; red = active transcription start site (TSS); orange-red = flanking active TSS; yellow-green = genic enhancers. Data were interpreted using the 15-state model. C, H3K27Ac marks, an indicator of active regulatory elements. D, DNase I hypersensitivity clusters, a mark of open chromatin. Gray boxes indicate the area of the hypersensitive region. The darkness of the box is proportional to the maximum signal strength observed. In C and D, data were obtained from the ENCODE project. E, UCSC reference genes. Genomic location and scale are denoted at the top.
primers listed in Supplementary Table 2 (http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract).

**Statistical analysis.** Kruskal-Wallis tests were used to assess the significance of the association between genotype and methylation of CpGs. For AEI, P values were calculated using Wilcoxon’s matched pairs signed rank test for LTN1, RWDD2, USP16, and CCT8 and an unpaired t-test for MAP3K7CL and BACH1. For two-way analyses, P values were calculated using Mann-Whitney 2-tailed exact tests. Correlations between genotype, age, and methylation, and the significance of the correlations, were determined using a standard least squares linear regression model.

**RESULTS**

SNP rs6516886 acting as an mQTL across multiple tissues. Previous work (15) has identified 6 CpG sites that correlate with the genotype at the OA risk–conferring SNP rs6516886 (Supplementary Table 3, http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract). In silico analysis of this region revealed that cg00065302, cg05468028, cg18001427, cg20220242, and cg24751378 all reside in areas marked as promoters in relevant cell types (Figure 1). Additionally, these sites are enriched with H3K27ac, an indicator of active regulatory elements, and DNase I hypersensitivity regions, which indicate accessible chromatin regions, a mark of a transcribed region. In contrast, the most distal of the CpGs, cg16140273, resides in a region with no notable epigenetic marks of a transcriptionally active or regulatory region.

Methylation analysis of cartilage, fat pad, synovium, and blood DNA samples from OA patients showed a correlation between methylation and the rs6516886 genotype, generating significant rs6516886 mQTLs at cg00065302, cg05468028, cg18001427, cg20220242, and cg16140273 across the tissues (Figure 2). We were unable to develop a pyrosequencing assay for cg24751378, which was therefore excluded from further investigation, whereas the cg05468028 assay captured an additional CpG site 2 bp upstream of cg05468028, which we termed cg05468028_2. This CpG was also an mQTL (Figure 2). The OA risk–conferring T allele of rs6516886 correlated with increased methylation across all tissues at cg00065302, cg05468028, cg05468028_2, cg18001427, and cg20220242,
and with decreased methylation across all tissues at cg16140273. These are the same directions of effect that we had previously observed in our array analysis of cartilage DNA (15). The most significant results obtained in the mQTL analysis were for cg20220242, with P values of <0.0001 in cartilage, fat pad, and synovium.

In summary, we replicated the mQTLs that we had previously reported in cartilage and demonstrated that they were also active across other joint tissues and blood. The most striking effect was for cg20220242.

**Methylation data independent of genotype.** When methylation data was plotted independent of genotype, we observed that there was a significantly higher methylation in blood for cg00065302, cg05468028, cg05468028_2, and cg00065302, compared to the other tissues (Supplementary Figure 2, http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract). Additionally, cg00065302 showed a clear decrease in methylation in cartilage samples. This effect can be seen in Supplementary Figure 3 (http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract), where data on patients for whom ≥2 tissue samples were available are plotted together. For the majority of patients, it can be clearly observed that at cg00065302, methylation values were the highest in blood and the lowest in cartilage. A striking example was in patient 333, who displayed a 24% difference between methylation in cartilage and blood. Analysis of methylation data stratified by joint revealed no significant differences between hip and knee cartilage (Supplementary Figure 4, http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract).

**Quantifying the effect of rs6516886 genotype on methylation.** We next determined how much of the observed interpatient variability in methylation was determined by genotype alone, with the magnitude of the genotypic effect calculated by linear regression analysis and represented in a heatmap (Supplementary Figure 5A, http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract). As with the mQTL analysis, cg20220242 displayed the greatest genotypic effect size across the joint tissues, with a 23.1% contribution in cartilage, 41.9% in fat pad, and 28.5% in synovium. Conversely, genotype had a relatively small effect on cg20220242 methylation variability in blood, with only a 6.5% contribution. Across the CpG sites, the greatest genotypic contribution to methylation was observed in the fat pad, except for cg16140273, in which the greatest contribution was observed in blood (20.2%). A linear regression analysis was also carried out to determine whether age contributed to the interpatient methylation variation, which confirmed that genotype had a much stronger role, and age had at most a minimal effect on methylation (0–3.8%) (Supplementary Figure 5B, http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract).

**Candidate genes at the association locus.** In silico analysis of chromatin interactions identified several examples of chromatin looping between areas containing the CpGs and genes at the locus (Figure 3). The cluster of cg05468028, cg05468028_2, cg18001427, and cg20220242, which are <1 kb apart, is located within the promoter/upstream region of RWDD2B but also shows interactions with LTN1, USP16, and CCT8. Both cg00065302 and cg16140273 show interactions with promoter regions of BACH1, with cg16140273 also interacting with CCT8 and MAP3K7CL. Similarly, cg00065302, which resides on the edge of the LTN1 promoter region, also shows interactions with RWDD2B, USP16, CCT8, and MAP3K7CL.

Quantification of the expression of these genes revealed that all 6 were expressed in the joint tissues and in blood (Supplementary Figure 6, http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract). Blood showed the highest level of expression.
for all investigated genes, except for RWDD2B, which was most highly expressed in fat pad ($P = 0.0004$).

**An rs6516886 eQTL at RWDD2B.** To test for differential allelic expression, transcript SNPs were identified for the 6 genes (Supplementary Table 4, http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract), and AEI analysis was conducted. For RWDD2B, the transcript SNP was in perfect linkage disequilibrium (LD; $r^2 = 1$) with rs6516886, while for LTN1, USP16, and CCT8, the transcript SNPs were in complete or near complete LD ($D^′ = 1$). As such, the phase between the transcript SNP alleles and the rs6516886 alleles could be determined unambiguously for these 4 genes. Patients’ compound heterozygotes for the respective transcript SNP and rs6516886 were therefore used in analyses of each of these genes. For MAP3K7CL and BACH1, the transcript SNPs were not in LD with rs6516886 (Supplementary Table 4). AEI at these 2 genes was therefore assessed by comparing the variance in compound heterozygotes to that measured in patients who were homozygotic at rs6516886 (21,22). In this instance, AEI driven by genotype at rs6516886 would be evidenced by compound heterozygotes forming bidirectional clusters at ratios higher and lower than 1.

Significant AEI in RWDD2B was detected in all 3 joint tissues, with the OA risk-conferring T allele of rs6516886 correlating with a decrease in expression in the combined analysis of all patients ($P < 0.0001$; Figure 4). In blood, for which the number of patients available for analysis was much smaller, AEI approached significance ($P = 0.062$) in the same direction (Figure 4D). Pairwise analysis of the magnitude of RWDD2B AEI between the tissues revealed that cartilage showed the least AEI and blood showed the most, with risk allele expression reduced on average by 15% in cartilage and 25% in blood (Figure 4E). Fat pad and synovium also had a greater average AEI compared to cartilage, but these were not significantly different compared to blood.

For the other 5 genes (Supplementary Figures 7–11, http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract), AEI was detected in combined analyses of all patients for LTN1 in cartilage ($P = 0.01$; Supplementary Figure 7A), CCT8 in synovium ($P = 0.021$; Supplementary Figure 9C), and BACH1 in synovium ($P = 0.04$; Supplementary Figure 11C).

**Figure 4.** RWDD2B allelic expression imbalance (AEI) in cartilage (A), fat pad (B), synovium (C), and blood (D), and the comparison of the scale of AEI between tissues (E). AEI analysis was conducted using transcript single-nucleotide polymorphism rs112411829. In A–D, the risk: nonrisk allelic ratio is plotted, with a ratio <1 indicating decreased RWDD2B expression from the risk allele. For each patient, the mean of the DNA ratio (black = 3 technical repeats) and the mean of the cDNA ratio (gray = 3 technical repeats) are plotted. Numbers on the x-axis refer to the anonymized patient identification number. Box plots show the values of DNA and cDNA for all patients combined. Lines within the box represent the median, the box represents the 25th to 75th percentiles, and the whiskers represent the maximum and minimum values. $P$ values were calculated using Wilcoxon’s matched pairs signed rank test in A–D and using the Mann-Whitney 2-tailed exact test in E.
To summarize, there is compelling evidence of AEI across all joint tissues and therefore of an rs6516886 eQTL acting on RWDD2B. For other genes, any positive evidence was much less significant and was not detected in more than 1 tissue.

Observation of rs6516886 meQTLs at RWDD2B. We next plotted the RWDD2B AEI data for the heterozygotes against their individual methylation data. The most striking effects were seen at cg20220242, with significant correlations seen across all 3 joint tissues (Figure 5). Interestingly, the effects were not all in the same direction, with lower AEI levels observed with increasing methylation in cartilage (\( P = 0.02 \)) and higher levels in fat pad (\( P = 0.036 \)) and synovium (\( P = 0.034 \)). Methylation at other CpGs also correlated with RWDD2B AEI. These included cg18001427 in cartilage (\( P = 0.048 \)) and cg05468028 and cg05468028_2 in fat pad (\( P = 0.0055 \) and \( P = 0.0085 \), respectively), all with the same tissue-specific directionality as seen with cg20220242 (Figure 5).

RWDD2B expression regulated by cg20220242 methylation. CpG cg20220242 was hypermethylated in chondrocytes, with median methylation levels >80% (Figure 6A). To determine whether methylation of cg20220242 is a direct regulator of RWDD2B expression, we reduced the methylation of the CpG, and of flanking CpGs, in the immortalized chondrocyte cell line Tc28a2, using CRISPR gRNAs and dCas9 fused to the DNA demethylating enzyme TET1. Using 2 different gRNAs, individually and in combination, we observed that guide 2, which localizes −15 to +5 bp across cg20220242 (Supplementary Figure 1, http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract), induced a significant decrease in methylation at cg20220242 (Figure 6A), with a reduction of 21.5% compared to control (\( P = 0.0022 \)). Guide 1 had no significant effect (\( P > 0.05 \)), and using both guides together did not reduce methylation greater than guide 2 alone (Figure 6A). We therefore focused on guide 2.

To determine the limits of the methylation modulation, we analyzed methylation at all CpGs 100 bp upstream and downstream of cg20220242 (Supplementary Figure 1, http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract). Methylation was significantly decreased (\( P < 0.05 \)) at all 5 CpGs upstream of cg20220242 and at 3 of the 8 CpGs downstream of cg20220242 (Figure 6B). The largest decrease in methylation was 36.6% at the CpG positioned −41 bp from cg20220242. There was a 3.8-fold increase in expression of RWDD2B in the demethylated cells compared to control (\( P = 0.0004 \); Figure 6C). No significant
changes in expression were observed in the flanking genes (P > 0.05 for all; Supplementary Figure 12, http://onlinelibrary.wiley.com/doi/10.1002/art.41473/abstract), indicating a specific effect on RWDD2B.

**DISCUSSION**

The number of GWAS conducted on human diseases has increased exponentially over the past 15 years. However, there is currently a huge disparity between the number of GWAS reports in the literature each year and the number of published functional follow-up studies investigating likely mechanisms of action of causal variants (23). One approach to functional analysis of signals obtained by GWAS is to study the epigenetics of the associated regions. DNA methylation is the best described and most studied epigenetic modification, and can regulate gene expression by recruiting proteins involved in gene repression or by inhibiting the binding of transcription factor(s) to DNA (24). As OA presents with pathologic changes in all of the joint tissues, it should be considered a disease of the whole joint (16), and investigation of genetic risk signals should therefore not be limited to cartilage alone. As such, when investigating the functional mechanisms of OA signals described in the GWAS literature, it is important to consider that the potential mechanisms may function differentially between tissues. The present study therefore aimed to increase understanding of the functional role of the OA risk locus marked by rs6516886, focusing on the CpGs previously identified as correlating with the genotype at this SNP in cartilage array data (15) and expanding the analysis to include multiple joint tissues and an analysis of gene expression.

Multi-tissue mQTLs operating at rs6516886 were initially discovered. The majority of these had consistently larger rs6516886 genotypic effects in the examined joint tissues compared to blood. The largest effects were seen at CpG cg20220242. However, none of the calculated genotypic effects were at or even approaching 100%, implying that rs6516886 genotype alone does not account for all of the variation in methylation that was observed. Other DNA polymorphisms or nongenetic factors presumably also influence the levels of methylation of the tested CpGs.

An rs6516886 RWDD2B eQTL was subsequently observed across the joint tissues using an AEI approach, with the OA risk–confering allele of the SNP correlating with reduced expression of the gene. Unlike for the mQTL data, this eQTL effect was as strong, if not stronger, in blood than in the joint tissues. A search of the Genotype-Tissue Expression portal (25; https://www.gtexportal.org/) revealed that there are ≥40 tissues in which RWDD2B eQTLs have been reported to correlate with rs6516886 genotype. These tissues include blood, but joint tissues were not examined. However, in each case, the OA risk–confering T allele of rs6516886 correlated with decreased expression of RWDD2B, a pattern we observed in joint tissues. While no other significant eQTL effects were observed, effects at the other genes cannot be ruled out, as the number of heterozygotes available for analysis varied between genes.

We next identified meQTLs between CpGs and RWDD2B, with cg20220242 methylation correlating significantly with RWDD2B expression across all 3 joint tissues. The direction of the meQTL effects varied between cartilage (lower AEI levels correlated with increasing methylation) and fat pad and synovium (higher AEI levels correlated with increased methylation), implying that although the meQTLs are active across the tissues, they may not necessarily operate in the same manner.

Interestingly, the RWDD2B eQTL and meQTL effects observed in our study were not restricted to cartilage but were also active in the other joint tissues. The strongest genotypic effect on methylation in fat pad was observed with rs6516886, which displayed significantly increased expression of RWDD2B compared to the other tissues. Once again, this highlights the importance of considering OA as a whole-joint disease (16) and indicates a role for fat pad in OA pathology (26,27).

All of the data generated previously have been correlative, with the OA risk–confering T allele of rs6516886 corresponding with increased methylation of cg20220242 and decreased expression of RWDD2B. To directly determine whether methylation levels of cg20220242 alter RWDD2B expression, we modulated the epigenome at and flanking the CpG using TET1-mediated

---

**Figure 6.** Methylation modulation of cg20220242 in Tc28a2–deactivated Cas9 (dCas9)–TET1 cells. A, Percentage methylation at cg20220242. Data were obtained from 3 biologic replicates. B, Percentage methylation at CpG sites 100 bp upstream and downstream of cg20220242 in control cells and cells transfected with guide 2 (G2). The position of CpGs was plotted relative to cg20220242 (dotted line) on the x-axis. Negative values are upstream, and positive values are downstream. Data were obtained from 6 biologic replicates. C, Gene expression of RWDD2B from 6 biologic replicates. Percentage methylation at CpG sites 100 bp upstream and downstream of cg20220242. Data were obtained from 3 biologic replicates.
demethylation in a chondrocyte cell line. This resulted in altered expression of the gene and, consistent with what we would have predicted from our correlative data, the decrease in methylation caused an increase in expression. We can therefore conclude that there is a direct link between a change in methylation and a subsequent change in RWDD2B expression. As far as we are aware, this is the first time that such a link has been experimentally demonstrated at an OA susceptibility locus using targeted methylome editing tools. It is also noteworthy that we have modulated the epigenome to counteract the effect of an OA risk allele by stimulating increased expression of a gene whose expression is typically reduced in carriers of the risk allele.

In conclusion, we have highlighted the expression of RWDD2B as a target of the rs6516886 OA association signal with differential methylation of cg20220242 and other CpGs in its vicinity as intermediaries in the regulation of the gene. The OA risk-conferring allele at the locus correlates with reduced expression of the gene, implying that a reduction in the level of its encoded protein is detrimental to joint health. RWDD2B codes for the RWD domain–containing protein 2B, about which very little is currently known. Proteins containing RWD domains have the capacity to bind to other cellular proteins, including ubiquitin ligases (28). The current dearth of knowledge regarding the biologic function of RWD domain–containing protein 2B means that further analysis of it and of its gene as a target of OA genetic risk will provide novel insight into this complex disease.

ACKNOWLEDGMENTS

The authors would like to acknowledge Dr Louise Reynard, Ms Melissa Kaye, and Ms Maria Tselepi for their technical advice and support.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Parker had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Parker, Loughlin.

Acquisition of data. Parker, Hofer, Rice, Earl, Anjum, Deehan.

Analysis and interpretation of data. Parker, Loughlin.

REFERENCES


ACKNOWLEDGMENTS

The authors would like to acknowledge Dr Louise Reynard, Ms Melissa Kaye, and Ms Maria Tselepi for their technical advice and support.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Parker had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Parker, Loughlin.

Acquisition of data. Parker, Hofer, Rice, Earl, Anjum, Deehan.

Analysis and interpretation of data. Parker, Loughlin.

REFERENCES


