OXFORD

Human Molecular Genetics, 2024, Vol. 33, 6, 501–509
https://doi.org/10.1093/hmg/ddad198
Advance access publication date 17 November 2023
Original Article

Epigenomic profiling of the infrapatellar fat pad in osteoarthritis

Peter Kreitmaier[1,2,3], Young-Chan Park[3], Diane Swift[4], Arthur Gilly[3], J. Mark Wilkinson[4,*,†], Eleftheria Zeggini[1,3,*,†]

[1]Technical University of Munich (TUM) and Klinikum Rechts der Isar, TUM School of Medicine and Health, Ismaninger Str. 22, Munich 81675, Germany
[2]Graduate School of Experimental Medicine, TUM School of Medicine and Health, Technical University of Munich, Ismaninger Str. 22, Munich 81675, Germany
[3]Institute of Translational Genomics, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstaedter Landstr. 1, Neuherberg 85764, Germany
[4]Department of Oncology and Metabolism, The University of Sheffield, Beech Hill Rd, Sheffield S10 2RX, United Kingdom
*Corresponding author. Institute of Translational Genomics, Helmholtz Zentrum München Deutsches Forschungszentrum für Gesundheit und Umwelt, Neuherberg 85764, Germany. E-mail: eleftheria.zeggini@helmholtz-munich.de and Metabolic Bone Unit, Sorby Wing Northern General Hospital Sheffield, S5 7AU United Kingdom. E-mail: j.m.wilkinson@sheffield.ac.uk
†J. Mark Wilkinson and Eleftheria Zeggini joint corresponding author.

Abstract

Osteoarthritis is a prevalent, complex disease of the joints, and affects multiple intra-articular tissues. Here, we have examined genome-wide DNA methylation profiles of primary infrapatellar fat pad and matched blood samples from 70 osteoarthritis patients undergoing total knee replacement surgery. Comparing the DNA methylation profiles between these tissues reveal widespread epigenetic differences. We produce the first genome-wide methylation quantitative trait locus (mQTL) map of fat pad, and make the resource available to the wider community. Using two-sample Mendelian randomization and colocalization analyses, we resolve osteoarthritis GWAS signals and provide insights into the molecular mechanisms underpinning disease aetiopathology. Our findings provide the first view of the epigenetic landscape of infrapatellar fat pad primary tissue in osteoarthritis.

Keywords: osteoarthritis; DNA methylation; infrapatellar fat pad; methylation QTL; EWAS

Introduction

Osteoarthritis is a complex joint disease that affects more than 300 million people [1]. In the face of aging populations, the impact of osteoarthritis on public health systems is estimated to increase further [1]. Current treatment methods are limited to pain management and total joint replacement, highlighting the need to develop novel, personalised treatment strategies. Therefore, it is important to enhance our understanding of the genetic and genomic basis of osteoarthritis.

To date, genome-wide association analyses (GWAS) have identified more than 150 genetic risk loci [2] of osteoarthritis, thus improving our understanding of its polygenic basis. Large-scale molecular datasets of relevant, primary cell types of osteoarthritis patients can reveal molecular mechanisms underlying disease and provide insights beyond genetic studies. Combining results from genetic and molecular studies can help pinpoint molecular mechanisms of disease development and progression, specifically the likely effector genes through which genetic risk variants exert their effect on osteoarthritis development in affected tissues.

Whilst a number of studies have investigated genome-wide molecular profiles of osteoarthritis-affected primary joint tissues [3, 4] the majority have focused on cartilage [5]. Osteoarthritis affects all joint tissues, and a small number of genome-wide molecular studies have extended molecular profiling to other primary joint tissues, such as the synovium [6, 7] or subchondral bone [8].

The infrapatellar fat pad, an adipocyte-rich tissue located inferior to the patella in the anterior part of the knee joint [9], has not been deeply studied in osteoarthritis to date. The fat pad is located among other joint tissues and protects knee components (by stabilising the patella) when exposed to mechanical stress, e.g. during exercise. In osteoarthritis-affected knees, the infrapatellar fat pad undergoes disease-related alterations, including fibrosis, inflammation and vascularization. Furthermore, it is traversed by nerves and therefore constitutes a source of knee osteoarthritis-related pain.

The fat pad may also interact with other joint tissues during osteoarthritis development and progression [9]. For example, it is proposed that the fat pad secretes pro-inflammatory and catabolic factors that promote cartilage degeneration and inhibit repair mechanisms [10]. Studies using chondrocyte cultures and fat pad-derived fat-conditioned media have provided some first insights into the potential cross-talk between the fat pad and cartilage [9].

Furthermore, the fat pad lies adjacent to the synovium, a connective tissue that lines the joint capsule. Both tissues undergo similar osteoarthritis-related changes, e.g. develop a similar immune cell profile [11]. Studies *in vitro* and in mouse

Received: July 19, 2023. **Revised:** October 13, 2023. **Accepted:** November 7, 2023

models suggest interactions between these tissues [12–14]. For example, Bastiaansen-Jenniskens *et al.* cultured fibroblast-like synoviocytes in fat-conditioned medium from fat pad samples of knee osteoarthritis patients, and suggest that fat pad induces fibrotic changes in synoviocytes by stimulating collagen synthesis as well as cell proliferation and migration [14].

Only a small number of studies have examined the profile of infrapatellar fat pad in osteoarthritis patients. Gandhi *et al.* characterised microarray-based gene expression profiles of the infrapatellar fat pad in 34 (29 and five in late and early stage knee osteoarthritis, respectively) individuals [15]. Other studies have investigated the molecular characteristics of osteoarthritis fat pad in genomic regions of osteoarthritis risk signals [16–18] or focused on cytokines and extracellular matrix genes [19].

In this study, we focus on DNA methylation, an epigenetic mark that refers to the covalent addition of a methyl-group to the DNA. Methylation is dynamic, tissue-specific, and plays a regulatory role in gene expression. In general, promoter methylation is negatively correlated with gene expression, whereas methylation in other parts of the genome, such as the gene body, remain less understood.

We examine the genome-wide DNA methylation profile of infrapatellar fat pad adipocytes of osteoarthritis-affected knees. We (1) compare fat pad and blood methylation profiles matched from the same patients, (2) generate a genome-wide methylation quantitative trait loci (mQTL) map in fat pad and (3) resolve osteoarthritis GWAS signals by integrating omics with genetic association data.

Results

Distinct epigenetic profiles in blood and fat pad adipocytes

We investigated global differences in the epigenetic profile between fat pad and peripheral blood samples for the first time. We performed PCA integrating infrapatellar fat pad samples from knee osteoarthritis patients (n = 70) and matched blood samples from a subset of these individuals (n = 58). We identified a separation of fat pad and blood samples along the first principal component, which was associated with tissue type (logistic regression p value: 2.7×10^−7, beta: −0.013, SE: 0.0026). This underlines the tissue-specificity of the epigenetic profile on a global level (Fig. 1A).

To characterise tissue-specificity on the methylation site level, we performed an epigenome-wide association study (EWAS) of matched fat pad and blood samples from the same patient (n = 58) and identified 84 973 (of 780 177 tested sites, 10.89%) strongly differentially methylated sites (DMS) between fat pad and whole blood samples ($P < 6.4\times10^{-8}$, beta > 2, Table S1). Of these, 33 391 and 51 582 showed hyper- and hypomethylation in fat pad tissue, respectively (Fig. 1B). Together, these results highlight extensive differences in the epigenetic profile of fat pad and peripheral blood.

Genome-wide mQTL map in fat pad adipocytes

We performed cis-mQTL analysis to estimate genetic variants that are associated with methylation levels of nearby methylation sites (<= 1 Mb). We identified 35 948 mQTL-targeted methylation sites (Fig. 2A, Methods), including cg20673407 (Fig. 2B) and cg14016568 (Fig. 2C). Together, this constitutes the first genome-wide mQTL map of infrapatellar fat pad adipocytes in knee osteoarthritis. The full summary statistics are publicly available in the Musculoskeletal Knowledge Portal (http://mskkp.org).

Osteoarthritis GWAS signal resolution

Next, we integrated the newly-generated fat pad mQTL map with GWAS results of two osteoarthritis traits, namely knee osteoarthritis and total knee replacement [2], to determine methylation sites with a putative causal role in osteoarthritis.

We applied colocalisation to estimate a probability for methylation mediating the osteoarthritis-promoting effect of risk variants. In total, we identified 16 methylation sites for which mQTL signals colocalised with 11 (of 25 tested, 44%) GWAS signals (Posterior probability for colocalisation > 80%) (Tables 1 and S2). For knee osteoarthritis, we resolved 9 (of 24 tested, 37.5%) GWAS signals that colocalised with mQTL of 13 methylation sites (Fig. 3A). Analogously, colocalising mQTL with GWAS results for total knee replacement resolved 5 (of 10 tested, 50%) GWAS signals and revealed 7 methylation sites with a potential causal role in osteoarthritis (Fig. 3B).

Next, we performed causal inference analysis by applying two-sample Mendelian randomization (MR) to estimate the putative causal effect of methylation on osteoarthritis. In these MR models, we used mQTL as instruments as well as mQTL-targeted methylation sites and osteoarthritis as exposure and outcome, respectively (Method). Here, we detected 36 methylation sites with a putative role ($P < 7.70\times10^{-07}$) in osteoarthritis (Fig. 3C), in total (Fig. S1, Table S3). For knee osteoarthritis, we identified 32 methylation sites, of which 15 and 17 revealed a link of hyper- and hypomethylation with osteoarthritis, respectively. For total knee replacement, we identified 15 methylation sites with a putative causal role (9 and 6 showing hyper- and hypomethylation in osteoarthritis, respectively). Eleven methylation sites were identified in both osteoarthritis-relevant traits, for which the direction of effect was concordant.

MR and colocalisation identified 37 putative causal methylation sites, in total. Of these, 15 were identified in both approaches, thus providing two lines of evidence for their respective causal involvement (Tables 1 and S4). Together, colocalisation and MR results suggest that these methylation marks mediate the regulation of genetic risk variants on effector genes in fat pad.

Annotated genes of the identified 37 methylation sites have been previously linked to osteoarthritis using causal approaches on genome-wide mQTL maps of cartilage or synovium. This includes *WWP2* (annotated to fat pad relevant methylation site cg04703221), a chondrocyte regulator [20] for which methylation has been causally linked to osteoarthritis in low disease-grade cartilage and synovium [7]. *ALDH1A2* (cg12031962, cg12031962 and cg08668585) has also previously been linked to osteoarthritis at the methylation [7] (in low-and high-grade osteoarthritis cartilage as well as synovium) as well as expression [6] (low-grade osteoarthritis cartilage) levels. Furthermore, we identified osteoarthritis-linked methylation in the collagen type *COL27A1* (cg21771125).

We also identified likely effector genes that were not previously resolved in molecular QTL maps of primary osteoarthritis cartilage and synovium [6, 7] including *USP8* (cg01701297 and cg05456662; involved in cell proliferation), *TSKU* (cg17107561; encodes development-linked extracellular matrix protein) and *FER1L4* (cg14387502 cg05220160; involved in plasma membrane organization) which can be linked to osteoarthritis-relevant mechanisms (Discussion). Together, integrating the fat pad mQTL profile with osteoarthritis GWAS results using colocalisation and MR identified 37 methylation sites with a potential causal involvement in osteoarthritis in fat pad tissue.

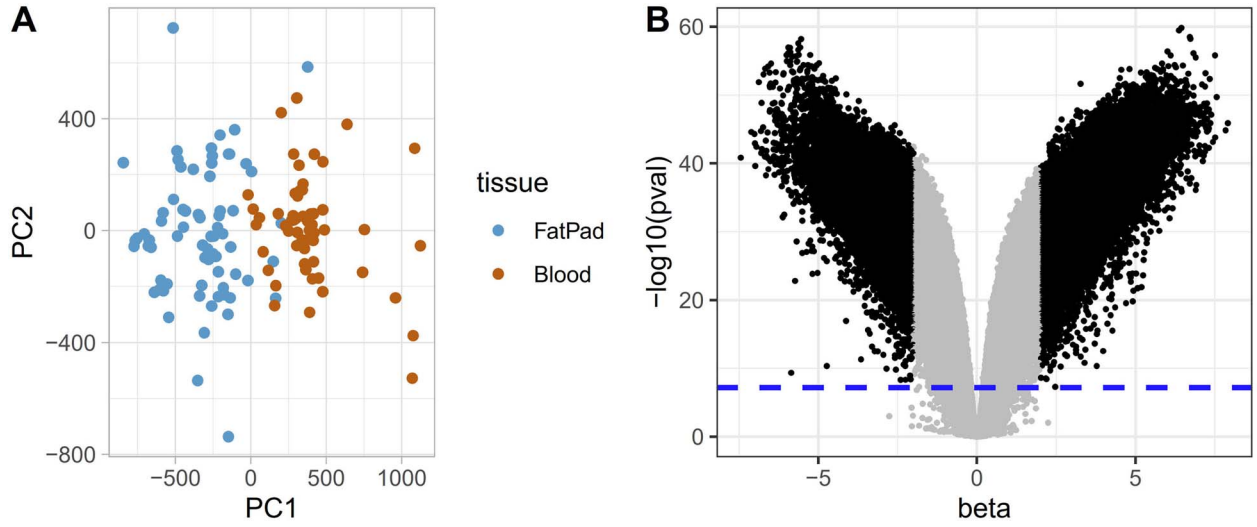

Figure 1. Distinct methylation profiles in blood and fat pad adipocytes. We investigated differences in the methylation profile between fat pad and blood. (A) On a global level, principal component analysis separates fat pad and blood samples along the first principal component. (B) On the methylation site level, a volcano plot demonstrates the multitude of differentially methylated sites. Sites with strong, differential methylation levels (beta > 2) exceeding the Bonferroni significance threshold ($P < 6.41 \times 10^{-8}$, dashed line) are shown in black, otherwise in grey.

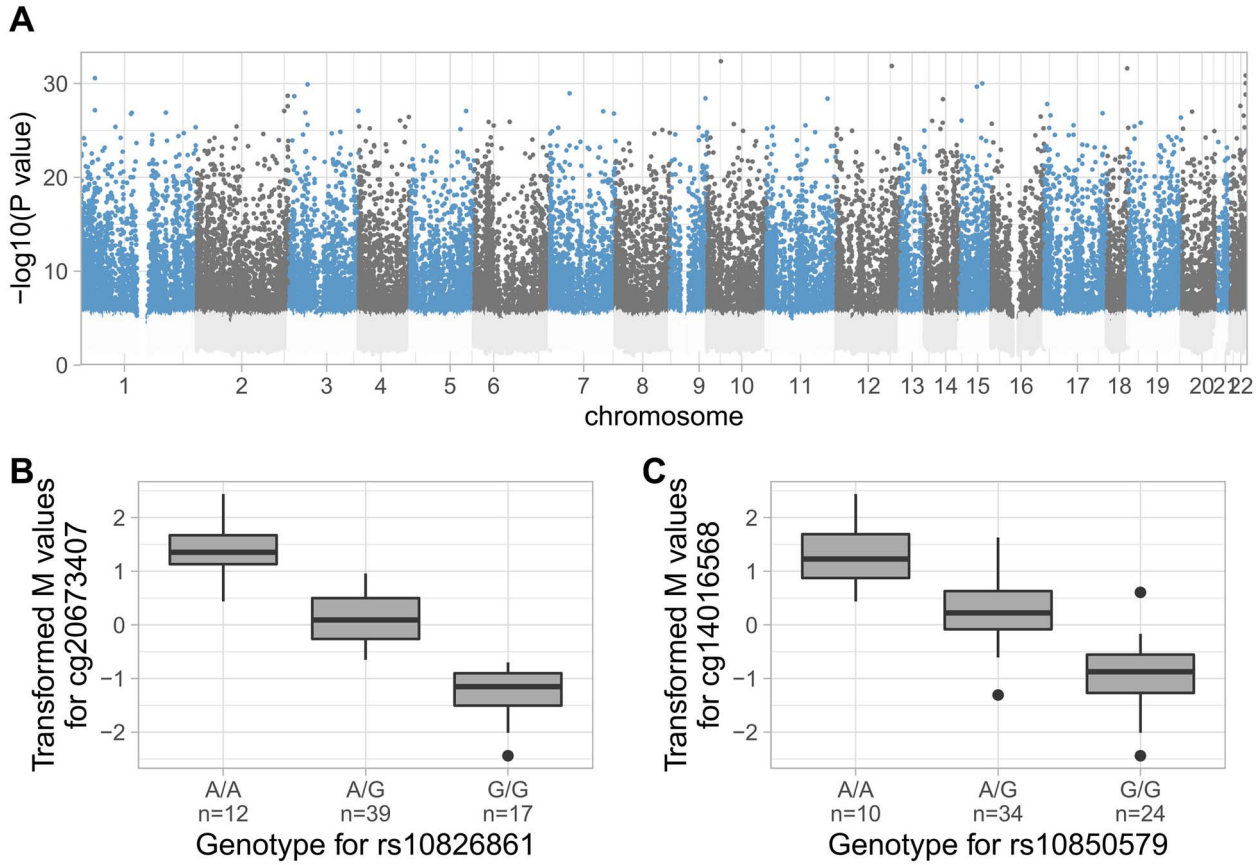

Figure 2. The mQTL map in fat pad adipocytes (A) Manhattan plots depicting the negative log of the *P* value of the most significant association per methylation site across all variants. QTL targeted methylation sites are shown in blue or dark grey, otherwise in light grey. As examples, the boxplots illustrate the effect (B) of rs10826861 on cg20673407 (beta = −1.40, SE = 0.05, $P = 4.15 \times 10^{-33}$) and (C) of rs10850579 on cg14016568 (beta = −1.20, SE = 0.05, $P = 1.10 \times 10^{-28}$). The boxplots represent 25th, 50th, and 75th percentiles, and whiskers extend to 1.5 times the interquartile range.

Discussion

Osteoarthritis is a common joint disorder with a polygenic architecture. Genome-wide molecular profiles of affected primary tissues remain understudied and excluded from large molecular data resources, such as GTEx [21], ENCODE [22] and RoadMap [23]. In this study, we characterised the first epigenome-wide profile of osteoarthritis-affected infrapatellar fat pad. We identify extensive differences from the epigenetic profile of peripheral blood, generate the first genome-wide mQTL map in fat pad, and identify methylation sites with a likely causal role in osteoarthritis development and progression.

Table 1. Overview of colocalisation signals. Overview of 16 methylation sites for which fat pad mQTL colocalise with an osteoarthritis GWAS risk signal for koa and/or tkr (indicated by the column "Coloc GWAS trait"). For 15 of these methylation sites, we also identified a putative causal effect (column "MR effect": Positive and negative effects indicate links of hyper- and hypomethylation with osteoarthritis, respectively) on osteoarthritis using MR ($P < 7.70\times10^{-07}$). Abbreviations: Chr, chromosome (hg38); Pos, position (hg 38); Msite, methylation site; MR, Mendelian randomization; Coloc PP, posterior probability for colocalization; koa, knee osteoarthritis; tkr, total knee replacement.

Msite	Chr	Pos (Msite)	Gene (Msite)	GWAS lead snp	Coloc GWAS trait	Coloc PP	MR effect	MR pval	MR GWAS trait
cg01030629	5	142 425 831	*SPRY4-AS1*	rs10038860	koa	0.98	−0.04	1.9×10^−08	koa
cg01100316	4	1 744 409	*TACC3*	rs7680647	koa	0.96	0.06	1.5×10^−08	koa
				rs4865462	tkr	0.86	0.10	4.9×10^−09	tkr
cg01150736	1	219 476 017		rs2791549	koa	0.81			
cg01701297	15	50 462 696	*RNA5SP395*	rs4380013	koa	0.90	−0.05	3.3×10^−09	koa
			USP8						
cg04703221	16	69 933 160	*MIR140*	rs34195470	koa	0.94	−0.08	5.7×10^−07	tkr
			WWP2				−0.06	1.0×10^−11	koa
cg04878480	12	48 012 099	*AC004801.4*	rs7967762	koa	0.92	0.04	5.8×10^−09	koa
			AC004801.5	rs7967762	tkr	0.92	0.07	2.2×10^−09	tkr
cg05456662	15	50 424 073	*USP8*	rs4380013	koa	0.93	0.06	3.9×10^−09	koa
cg08668585	15	57 955 405	*ALDH1A2*	rs4144005	tkr	0.86	0.08	3.6×10^−10	tkr
cg10169515	12	123 222 989	*MPHOSPH9*	rs753350451	koa	0.86	−0.03	2.2×10^−08	koa
cg10239804	12	48 104 587	*PFKM*	rs7967762	koa	0.92	−0.05	5.8×10^−09	koa
			SENP1	rs7967762	tkr	0.92	−0.09	2.2×10^−09	tkr
cg12031962	15	58 061 651	*ALDH1A2*	rs4144005	tkr	0.90	0.07	3.6×10^−10	tkr
cg15373332	9	114 173 564	*COL27A1*	rs72760655	koa	0.99	0.06	4.4×10^−10	koa
				rs7023177	tkr	0.99	0.10	5.5×10^−10	tkr
cg15672022	5	142 426 207	*SPRY4-AS1*	rs10038860	koa	0.98	−0.08	1.9×10^−08	koa
cg16740022	5	142 426 441	*SPRY4-AS1*	rs10038860	koa	0.98	−0.04	1.9×10^−08	koa
cg17669802	20	35 387 551	*UQCC1*	rs143384	tkr	0.90	−0.08	9.41×10^−21	koa
							−0.11	2.2×10^−14	tkr
cg17729365	19	10 643 944	*SLC44A2*	rs2163832	koa	0.96	−0.06	1.0×10^−08	koa

Comparing fat pad and blood methylation profiles reveals abundant epigenetic differences underlining the epigenetic tissue-specificity of blood and joint tissues, thus highlighting the necessity to investigate disease-affected tissues.

We present the first genome-wide mQTL map for osteoarthritis-affected infrapatellar fat pad. Colocalising this mQTL map with osteoarthritis GWAS results resolved eleven genetic osteoarthritis risk signals, thus providing evidence for methylation mediating the genetic effect of these GWAS signals on osteoarthritis in fat pad.

We supplemented these causal insights using MR and, together with colocalisation, identified 37 methylation sites with a putative causal role in osteoarthritis in fat pad. Some methylation sites were close to genes (such as *WWP2*, *ALDH1A2*, and *COL27A1*) that have been previously causally linked to osteoarthritis using genome-wide molecular QTL maps of other primary joint tissues [6, 7], suggesting a disease-relevant role across joint tissues.

We also identify genes that have not been previously resolved in molecular QTL maps of primary osteoarthritis tissues [6, 7] such as *USP8, TSKU* and *FER1L4*. USP8 is involved in epidermal growth factor receptor regulation [24], a receptor linked to angiogenic and inflammatory mechanisms. *TSKU* is an inhibitor of Wnt signaling, a pathway which has been consistently linked to osteoarthritis across tissues, e.g. in cartilage, synovium and subchondral bone [25]. *FER1L4* regulates inflammatory factor IL-6 in osteoarthritis-affected cartilage [26] and is linked to VEGF, an osteoarthritis-linked angiogenic factor [27].

These signals can be related to signalling pathways that may contribute to osteoarthritis development in the infrapatellar fat pad and its interaction with other joint tissues.

For example, methylation-mediated upregulation of cytokines may be involved. Elevated levels of IL-6 and VEGF have been previously observed in fat pad samples of osteoarthritis patients [28]. Both factors are regarded to affect surrounding joint tissues. IL-6 is linked to protective, but also inflammatory and catabolic mechanisms in the cartilage [29]. Increased fat pad mRNA expression of VEGF has been associated with higher vascularisation of the neighbouring synovium [28], suggesting interactions between these tissues. The identification of a Wnt pathway regulator (TSKU) may relate to the production of WISP2, a target of the Wnt pathway, for which increased expression levels have been identified in osteoarthritis-affected fat pad [30].

Together, we have identified genes linked to processes that are observed in osteoarthritis-affected fat pad, such as inflammation or vascularization [28], and suggest an involvement of the detected methylation sites in disease-related alterations.

In this work, we investigate blood and fat pad methylation profiles of osteoarthritis patients, which reflect disease processes that could be cause or effect. The integration of genetic data, coupled to colocalisation and causal inference analyses, were all used to deconvolute the role of methylation in osteoarthritis. Obtaining healthy joint tissue as a control for the osteoarthritis-affected joint presents a major challenge, as removal of healthy structural human joint tissue is precluded on ethical and acceptability grounds. The fat pad mQTL map provides insights into genetic effects on infrapatellar fat pad methylation for the first time. Larger sample sizes will be required to achieve 80% power to detect mQTLs across the allele frequency spectrum (Fig. S2).

Our findings highlight differences in the epigenetic profile of fat pad tissue and blood and identify methylation sites that likely

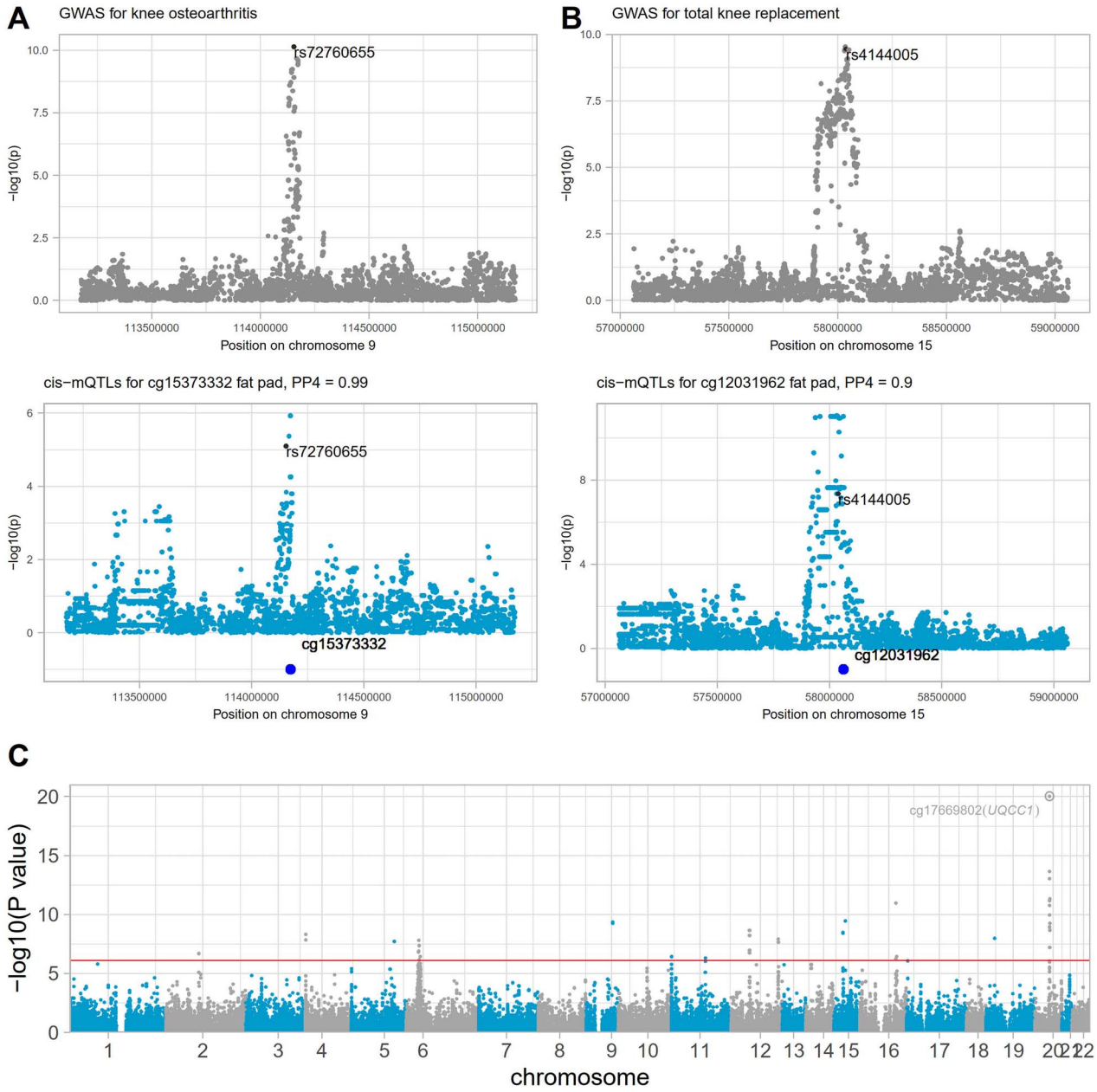

Figure 3. Osteoarthritis GWAS risk signals colocalise with mQTL two colocalisation events are exemplified in (A) and (B). In (A), we colocalised cis-mQTL for cg15373332 with a knee osteoarthritis GWAS signal (Posterior probability = 98.8%). Similarly, (B) shows cis-mQTL for cg12031962 colocalising with a total knee replacement GWAS signal (Posterior probability = 89.8%). (C) Manhattan plot depicting the Mendelian randomization *P* values to estimate the putative causal effects of methylation sites in fat pad on knee osteoarthritis or total knee replacement. The line indicates genome-wide significance applying the Bonferroni correction ($P < 8.31 \times 10^{\wedge}-07$).

exert the effect of GWAS risk signals in fat pad, shedding light on the mechanistically relevant role of fat pad methylation in osteoarthritis.

Materials and methods

Study participants

We have collected tissue samples from 210 patients undergoing total joint replacement surgery (111 women, 99 men, age 48–93 years, mean 71 years). All patients provided written, informed consent prior to participation in the study. Adipose tissue was collected from the infra-patellar fat pad by sharp dissection of the fat tissue from the surface of the patellar ligament to yield not less than $1cm^3$ of homogeneous adipose tissue. This work was approved by Oxford NHS REC C (10/H0606/20, SC/15/0132 and SC/20/0144), and samples were collected under Human Tissue Authority license 12 182, South Yorkshire and North Derbyshire Musculoskeletal Biobank, University of Sheffield, UK. We confirmed a joint replacement for osteoarthritis, with no history of significant knee surgery (apart from meniscectomy), knee infection, or fracture, and no malignancy within the previous 5 years. We further confirmed that no patient used glucocorticoid use (systemic or intra-articular) within the previous 6 months, or any other drug associated with immune modulation. We also obtained a peripheral blood sample to extract DNA from all patients.

Adipocyte and peripheral blood collection and processing

Adipose tissue samples were transported in Dulbecco's modified Eagle's medium (DMEM)/F-12 (1:1) (Life Technologies) supplemented with 2 mM glutamine (LifeTechnologies), 100 U/ml penicillin, 100 μg/ml streptomycin (Life Technologies), 2.5 μg/ml amphotericin B (Sigma-Aldrich) and 50 μg/ml ascorbic acid (Sigma-Aldrich) (serum free media). Next, the adipose tissue samples were cut into small pieces (<2mm^3) and digested in 3 mg/ml collagenase type I (Sigma-Aldrich) in serum free media for 1 h at 37°C on a flatbed shaker and resuspended in 2mls of PBS and passed through a 100 μm cell strainer (Fisher Scientific). Next, the eluent was made up to 10mls in PBS and centrifuged at 23 g for 5 min. Subsequently, the cell pellet was washed twice in PBS and centrifuged at 323 g for a further 5 min. Cells were counted using a haemocytometer and the viability checked using trypan blue exclusion (Invitrogen). The resulting cell pellet was resuspended in 650 μl of RLT buffer (Qiagen) and DTT Dithiothreitol (20ul DTT per 1 ml of RLT). The optimal cell number for spin column extraction from cells was between 4×10^6 and 1×10^7. Cells were then pelleted and homogenised. DNA extraction was carried out using Qiagen AllPrep DNA Mini Kit following the manufacturer's instructions. Samples were flash frozen in liquid nitrogen and stored at −80°C prior to assays. Peripheral blood was extracted for DNA using a Qiagen QIAamp DNA Blood Maxi kit, according to manufacturer's instructions. The whole blood DNA samples were frozen at −80°C prior to extraction.

Methylation data preprocessing

Genome-wide DNA methylation was measured using the Illumina EPIC array. We preprocessed methylation data using an R package meffil [31] based preprocessing pipeline (https://github.com/perishky/meffil/wiki).

We read and preprocessed blood DNA methylation data using the function meffil.qc, and removed ethnicity outliers, hip samples, samples with > 10% undetected (detection pvalue > 0.01) methylation values, sex outliers (> 5 ∗ sd), methylated/unmethylated signal outliers (> 3 ∗ sd) and control probe signal outlier (> 5 ∗ sd). We then applied the same procedure (same R functions and thresholds) on DNA methylation samples from fat pad samples. Finally, we normalised methylation samples of all tissues together by applying meffil function meffil.normalize.quantiles (using 16 principal components) and meffil.normalize.samples.

We removed methylation sites with more than 10% of samples low bead number (< 3) or undetected methylation values (detection $P < 0.01$), non-autosomal methylation sites, methylation sites of cross-reactive probes and in close proximity (within 10 base pairs) to common SNPs (MAF > 0.05) in European population [32–34].

We converted initially generated beta values to Mvalues (beta2m function of R package lumi) [35] which we used for downstream analyses [36]. Per tissue, we further replaced strong outliers (> 10 ∗ sd from mean) with the methylation site-specific mean value. Based on a principal component analysis, we removed two outlier samples. The resulting fat pad methylation data comprised 780 177 methylation sites for 70 patients (46 women, 24 men, age 48–93 years, mean 71 years). For 58 of 70 patients, also methylation blood samples were available.

We used publicly available annotations (https://zwdzwd.github.io/InfiniumAnnotation/EPIC.hg38.manifest.tsv.gz and EPIC.hg38.manifest.gencode.v36.tsv.gz) to map probe identifier to the genomic location (hg38) and genes.

Whole-genome sequencing data generation and preprocessing

Whole-genome sequencing (WGS) samples were available for 68 of 70 patients with matching fat pad methylation data. They were measured in two sequencing batches. Of 68 WGS samples, 60 were measured in whole blood samples in the first sequencing batch. Furthermore, eight of 68 WGS samples were measured in cartilage samples in the second batch. In both batches, DNA samples were subjected to standard Illumina paired-end DNA library construction, amplified, and subjected to DNA sequencing using the NovaSeq platform.

Generated CRAM files were input into samtools (samtools conda version 1.14) to create bam files. Subsequently, "bedtools bamtofastq" (bedtools conda version 2.30.0) was applied to obtain data in the fastq format. Per sequencing batch, variant calling was performed using the publicly available pipeline Sarek from nf-core (version 2.7.1, https://nf-co.re/sarek/2.7.1) with the additional options "– tools HaplotypeCaller –generate_gvcf". This uses the GATK Haplotypecaller (GATK v4.1.7.0) and generates g.vcf files. For the genome "GRCh38" was used. For the joint variant calling we adapted a publicly available pipeline (https://github.com/IARCbioinfo/gatk4-GenotypeGVCFs-nf) and used it with GATK (docker container broadinstitute/gatk:4.2.5.0). Reference files for GRCh38 were used from GATK.

For QC on the variant level, we applied Variant Quality Score Recalibration tool using a tranche threshold of 99.5% for SNPs and the recommended 99% for INDELs. For SNPs, this produces an expected false positive rate of 2.5% and an expected sensitivity of 97%.

For QC on the sample-level, we removed strong outlier het rate (two samples), and non-reference allele concordance when compared to directly typed genotype data using variants MAF > 0.01 (one sample), and one sample being a moderate outlier in sequencing depth as well as het rate. No additional sample was excluded solely based on Ti/Tv or singletons.

Furthermore, we removed one sex mismatch and two samples to avoid the inclusion of any sample pair with a relatedness > 0.2. We further excluded two ethnicity outliers identified in an ethnicity check-up using Ancestry and kinship toolkit (based on 1000G data from phase three; https://github.com/Illumina/akt/tree/master) [37]. In total, we removed nine samples.

We excluded variants with MAF < 0.01, Hardy-Weinberg equilibrium $P < 10\text{^}-5$ and call rate < = 0.99. We then selected samples of individuals with matching fat pad methylation data (n = 68) and kept bi-allelic variants with MAF > 0.05. The resulting WGS data set used for the fat pad mQTL analysis comprised 68 samples and 6 395 994 variants.

Comparing DNA methylation of blood and fat pad tissue

We integrated 70 fat pad and 58 blood samples in a principal component analysis (PCA) to investigate global epigenetic differences between these tissue types. We regressed out known technical batches (slide, row, clinical cohort) by applying Combat from the R package sva [38] and performed PCA using prcomp function.

To compare methylation profiles on methylation site level, we performed differential methylation analysis between fat pad and blood samples paired from the same patients (n = 58). We performed linear modelling using the functions lmFit and eBayes from limma [39]. We added the patient ID to ensure paired analysis design and included 19 surrogate variables (SVs) to account for technical variants. These SVs were estimated using the num.sv

function from the sva package ('be'-method) by protecting the tissue information. We highlighted methylation sites that exceed genome-wide significance threshold (Bonferroni correction with $P < 6.41 \times 10\text{^}-08$ which corresponds to 0.05/780 177 methylation sites) with strong effect size (beta > 2).

Methylation quantitative trait locus analysis

For the methylation quantitative trait locus (mQTL) analysis, we included whole-genome-sequencing data and fat pad methylation data matching from the same patients (n = 68). We included 6 395 994 bi-allelic genetic variants with a MAF > 0.05 among these 68 patients. We further normalised methylation levels using inverse-normal transformation per methylation site and estimated PEER factors [40] (R package peer, default parameter setting) to infer hidden factors which we included to correct for technical variation. We performed cis-methylation QTL analysis (cis distance: 1 Mb either side of the tested methylation site) using FastQTL (https://github.com/francois-a/fastqtl/) [41]. We first estimated nominal p values for every tested methylation site-variant pair using linear regression with the following model:

$$Methylation\ values \sim genotype + age + sex + sequencing\_batch + row + 10\ PEER\_factors$$

Here, *row* refers to the sample location on the Illumina EPIC array chip. Since *row* can influence methylation levels [42], but did not significantly (ANOVA Bonferroni $P < 0.05$) associate with any of the ten PEER factors (Table S5), we conservatively added it to the model. The variable *sequencing_batch* accounts for WGS sequencing batches. Of 68 WGS samples, 60 and 8 were extracted in the first and second WGS sequencing batch, respectively (methods section "Whole-genome sequencing data generation and preprocessing"). To optimise the number of included PEER factors, we performed mQTL analysis with five, ten and 15 PEER factors and chose the number that maximises detected mQTL targeted methylation site (5 PEER factors: 34956 mQTL targeted methylation sites, 10 PEER factors: 35948, 15 PEER factors: 35808). Secondly, we applied an adaptive permutation scheme (implemented in FastQTL, parameter –permute 1000 10 000) to estimate a q value and nominal *P*-value threshold per methylation site. Methylation sites with a q value < 5% Storey-Tibshirani FDR are regarded as mQTL targeted. For each mQTL-targeted methylation site, significant QTL were variants with a nominal *P* value below the nominal *P* value threshold for that methylation site. Power analysis for the methylation QTL analysis was performed using the R package powerEQTL [43] (function powerEQTL.SLR) across MAF and sample sizes.

Colocalisation

We colocalised [44] fat pad methylation QTL with GWAS signals for knee osteoarthritis and total knee replacement [2]. For this analysis, we applied the coloc.fast function (https://github.com/tobyjohnson/gtx/blob/526120435bb3e29c39fc71604eee03a371ec3753/R/coloc.R) using default settings. We considered mQTL-targeted methylation sites located in the region that spans 100 kb either side of the GWAS signal index variant. For the colocalisation analysis, we included all variants that were included in the cis mQTL analysis for the tested methylation sites. We considered posterior probabilities ("PP4") > 80% as indicator for colocalisation.

Mendelian randomization

To estimate putative causal effects of QTL-targeted methylation sites in fat pad on osteoarthritis traits, we integrated the fat pad mQTL map with GWAS results for knee osteoarthritis and total knee replacement [2]. We applied two sample Mendelian randomization (MR) using the pipeline of the R package TwoSampleMR [45]. We considered methylation sites targeted by at least one mQTL. Per methylation site, we performed clumping (function clump_data, using the European reference panel and setting the R2 threshold to 0.01) to identify independent genetic variants which we included as instruments in the MR models. For methylation sites with one independent instrument, we applied the Wald-ratio, otherwise the inverse variance weighted method.

In total, we applied 64 898 MR models (32 448 and 32 450 for knee osteoarthritis and total knee replacement, respectively) to estimate the putative causal effect of 32 456 methylation sites. We applied the Bonferroni method to correct for multiple testing ($P < 7.70 \times 10\text{^}-07$).

Acknowledgements

We acknowledge the technical support of Core Facility Genomics at Helmholtz Munich. We thank Dr Inti Alberto de la Rosa Velasquez, Dr Peter Lichtner, Susanne Wittmann and Dr Thomas Walzthöni for help with DNA methylation and whole-genome sequencing data generation as well as whole-genome sequencing data preprocessing.

Author contributions

Study design: E.Z., J.M.W.; Clinical collection: J.M.W., D.S.; WGS data preprocessing: A.G., Y.C.P.; Data analysis: P.K.; Interpretation of results: P.K., E.Z.; Manuscript drafting: P.K., E.Z.; Manuscript reviewing and editing: P.K., E.Z., J.M.W., D.S.

Supplementary data

Supplementary data is available at *HMG Journal* online.

Conflict of interest statement: None declared.

Funding

This work was funded by the Wellcome Trust (206194).

Data availability

Full summary statistics will be made openly available through the MSK portal (http://mskkp.org) upon manuscript acceptance. All software used in this study is available from free repositories or manufacturers as referenced in the Materials and Methods section.

References

1. Safiri S, Kolahi A-A, Smith E. *et al.* Global, regional and national burden of osteoarthritis 1990-2017: a systematic analysis of the global burden of disease study 2017. *Ann Rheum Dis* 2020;**79**: 819–28.
2. Boer CG, Hatzikotoulas K, Southam L. *et al.* Deciphering osteoarthritis genetics across 826 690 individuals from 9 populations. *Cell* 2021;**184**:4784–4818.e17.

3. Kreitmaier P, Katsoula G, Zeggini E. Insights from multi-omics integration in complex disease primary tissues. *Trends Genet* 2023;**39**:46–58.
4. Katsoula G, Kreitmaier P, Zeggini E. Insights into the molecular landscape of osteoarthritis in human tissues. *Curr Opin Rheumatol* 2022;**34**:79–90.
5. Katsoula G, Steinberg J, Tuerlings M. *et al.* A molecular map of long non-coding RNA expression, isoform switching and alternative splicing in osteoarthritis. *Hum Mol Genet* 2022;**31**: 2090–105.
6. Steinberg J, Southam L, Roumeliotis TI. *et al.* A molecular quantitative trait locus map for osteoarthritis. *Nat Commun* 2021;**12**:1309.
7. Kreitmaier P, Suderman M, Southam L. *et al.* An epigenome-wide view of osteoarthritis in primary tissues. *Am J Hum Genet* 2022;**30**:S48.
8. Tuerlings M, van Hoolwerff M, Houtman E. *et al.* RNA sequencing reveals interacting key determinants of osteoarthritis acting in subchondral bone and articular cartilage: identification of IL11 and CHADL as attractive treatment targets. *Arthritis Rheumatol* 2021;**73**:789–99.
9. Zeng N, Yan Z-P, Chen X-Y. *et al.* Infrapatellar fat pad and knee osteoarthritis. *Aging Dis* 2020;**11**:1317–28.
10. Clockaerts S, Bastiaansen-Jenniskens YM, Runhaar J. *et al.* The infrapatellar fat pad should be considered as an active osteoarthritic joint tissue: a narrative review. *Osteoarthr Cartil* 2010;**18**:876–82.
11. Klein-Wieringa IR, de Lange-Brokaar BJE, Yusuf E. *et al.* Inflammatory cells in patients with endstage knee osteoarthritis: a comparison between the synovium and the infrapatellar fat pad. *J Rheumatol* 2016;**43**:771–8.
12. Clements KM, Ball AD, Jones HB. *et al.* Cellular and histopathological changes in the infrapatellar fat pad in the monoiodoacetate model of osteoarthritis pain. *Osteoarthr Cartil* 2009;**17**: 805–12.
13. Eymard F, Pigenet A, Citadelle D. *et al.* Induction of an inflammatory and prodegradative phenotype in autologous fibroblast-like synoviocytes by the infrapatellar fat pad from patients with knee osteoarthritis. *Arthritis Rheumatol* 2014;**66**:2165–74.
14. Bastiaansen-Jenniskens YM, Wei W, Feijt C. *et al.* Stimulation of fibrotic processes by the infrapatellar fat pad in cultured synoviocytes from patients with osteoarthritis: a possible role for prostaglandin f2α. *Arthritis Rheum* 2013;**65**:2070–80.
15. Gandhi R, Takahashi M, Virtanen C. *et al.* Microarray analysis of the infrapatellar fat pad in knee osteoarthritis: relationship with joint inflammation. *J Rheumatol* 2011;**38**:1966–72.
16. Sorial AK, Hofer IMJ, Tselepi M. *et al.* Multi-tissue epigenetic analysis of the osteoarthritis susceptibility locus mapping to the plectin gene PLEC. *Osteoarthr Cartil* 2020;**28**:1448–58.
17. Parker E, Hofer IMJ, Rice SJ. *et al.* Multi-tissue epigenetic and gene expression analysis combined with epigenome modulation identifies RWDD2B as a target of osteoarthritis susceptibility. *Arthritis Rheum* 2021;**73**:100–9.
18. Rushton MD, Reynard LN, Young DA. *et al.* Methylation quantitative trait locus analysis of osteoarthritis links epigenetics with genetic risk. *Hum Mol Genet* 2015;**24**:7432–44.
19. Belluzzi E, Macchi V, Fontanella CG. *et al.* Infrapatellar fat pad gene expression and protein production in patients with and without osteoarthritis. *Int J Mol Sci* 2020;**21**:6016.
20. Mokuda S, Nakamichi R, Matsuzaki T. *et al.* Wwp2 maintains cartilage homeostasis through regulation of Adamts5. *Nat Commun* 2019;**10**:2429.
21. GTEx Consortium. The GTEx consortium atlas of genetic regulatory effects across human tissues. *Science* 2020;**369**: 1318–30.
22. Dunham I, Kundaje A, Aldred SF. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;**489**: 57–74.
23. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* 2015;**518**:317–30.
24. Berlin I, Schwartz H, Nash PD. Regulation of epidermal growth factor receptor ubiquitination and trafficking by the USP8·STAM complex. *J Biol Chem* 2010;**285**:34909–21.
25. Wang Y, Fan X, Xing L. *et al.* Wnt signaling: a promising target for osteoarthritis therapy. *Cell Commun Signal* 2019;**17**:97.
26. He J, Wang L, Ding Y. *et al.* lncRNA FER1L4 is dysregulated in osteoarthritis and regulates IL-6 expression in human chondrocyte cells. *Sci Rep* 2021;**11**:13032.
27. Hamilton JL, Nagao M, Levine BR. *et al.* Targeting VEGF and its receptors for the treatment of osteoarthritis and associated pain. *J Bone Miner Res* 2016;**31**:911–24.
28. Favero M, El-Hadi H, Belluzzi E. *et al.* Infrapatellar fat pad features in osteoarthritis: a histopathological and molecular study. *Rheumatology (Oxford)* 2017;**56**:1784–93.
29. Wiegertjes R, van de Loo FAJ, Blaney Davidson EN. A roadmap to target interleukin-6 in osteoarthritis. *Rheumatology (Oxford)* 2020;**59**:2681–94.
30. Conde J, Scotece M, Abella V. *et al.* Identification of novel adipokines in the joint. Differential expression in healthy and osteoarthritis tissues. *PLoS One* 2015;**10**:e0123601.
31. Min JL, Hemani G, Davey Smith G. *et al.* Meffil: efficient normalization and analysis of very large DNA methylation datasets. *Bioinformatics* 2018;**34**:3983–9.
32. McCartney DL, Walker RM, Morris SW. *et al.* Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip. *Genom Data* 2016;**9**: 22–4.
33. Pidsley R, Zotenko E, Peters TJ. *et al.* Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol* 2016;**17**:208.
34. Chen Y, Lemire M, Choufani S. *et al.* Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* 2013;**8**:203–9.
35. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 2008;**24**:1547–8.
36. Du P, Zhang X, Huang C-C. *et al.* Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 2010;**11**:587.
37. Arthur R, Schulz-Trieglaff O, Cox AJ. *et al.* AKT: ancestry and kinship toolkit. *Bioinformatics* 2017;**33**:142–4.
38. Leek JT, Johnson WE, Parker HS. *et al.* The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 2012;**28**:882–3.
39. Ritchie ME, Phipson B, Wu D. *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;**43**:e47.
40. Stegle O, Parts L, Piipari M. *et al.* Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. *Nat Protoc* 2012;**7**: 500–7.
41. Ongen H, Buil A, Brown AA. *et al.* Fast and efficient QTL mapper for thousands of molecular phenotypes. *Bioinformatics* 2016;**32**: 1479–85.

42. Price EM, Robinson WP. Adjusting for batch effects in DNA methylation microarray data, a lesson learned. *Front Genet* 2018;**9**:83.
43. Dong X, Li X, Chang T-W. *et al.* powerEQTL: an R package and shiny application for sample size and power calculation of bulk tissue and single-cell eQTL analysis. *Bioinformatics* 2021;**37**: 4269–71.
44. Giambartolomei C, Vukcevic D, Schadt EE. *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* 2014;**10**: e1004383.
45. Hemani G, Zheng J, Elsworth B. *et al.* The MR-base platform supports systematic causal inference across the human phenome. *elife* 2018;**7**:e34408.