Molecular Immunology 131 (2021) 89–96

Contents lists available at ScienceDirect

Molecular Immunology

journal homepage: www.elsevier.com/locate/molimm

Integrative analysis identifies potential causal methylation-mRNA regulation chains for rheumatoid arthritis

Xing-Bo Mo [a,b], Yong-Hong Zhang [b], Shu-Feng Lei [a,b,*]

[a] Jiangsu Key Laboratory of Preventive and Translational Medicine for Geriatric Diseases, Center for Genetic Epidemiology and Genomics, School of Public Health, Medical College of Soochow University, Suzhou, Jiangsu, 215123, PR China
[b] Jiangsu Key Laboratory of Preventive and Translational Medicine for Geriatric Diseases, Department of Epidemiology, School of Public Health, Medical College of Soochow University, Suzhou, Jiangsu, 215123, PR China

ARTICLE INFO

Keywords:
Rheumatoid arthritis
Methylation
Gene expression
Mendelian randomization
Causal inference

ABSTRACT

Genome-wide association studies have identified many genetic loci for rheumatoid arthritis (RA). However, causal factors underlying these loci were largely unknown. The aim of this study was to identify potential causal methylation-mRNA regulation chains for RA. We identified differentially expressed mRNAs and methylations and conducted summary statistic data-based Mendelian randomization (SMR) analysis to detect potential causal mRNAs and methylations for RA. Then causal inference test (CIT) was performed to determine if the methylation-mRNA pairs formed causal chains. We identified 11,170 mRNAs and 24,065 methylations that were nominally associated with RA. Among them, 197 mRNAs and 104 methylations passed the SMR test. According to physical positions, we defined 16 *cis* methylation-mRNA pairs and inferred 5 chains containing 4 methylations and 4 genes (*BACH2*, *MBP*, *MX1* and *SYNGR1*) to be methylation→mRNA→RA causal chains. The effect of *SYNGR1* expression in peripheral blood mononuclear cells on RA risk was found to be consistent in both the in-house and public data. The identified methylations located in CpG Islands that overlap promoters in the 5′ region of the genes. The promoter regions showed long-range interactions with other enhancers and promoters, suggesting a regulatory potential of these methylations. Therefore, the present study provided a new integrative analysis strategy and highlighted potential causal methylation-mRNA chains for RA. Taking the evidences together, *SYNGR1* promoter methylations most probably affect mRNA expressions and then affect RA risk.

1. Introduction

Rheumatoid arthritis (RA) is one of the most common forms of autoimmune arthritis disease (Smolen et al., 2016). RA is caused by combinations of genetic and environmental factors (MacGregor et al., 2000). Genetic and epigenetic factors underlying susceptibility to RA were largely unknown. Previous large scale genome-wide association studies (GWAS) have identified more than one hundred RA susceptibility loci (Eyre et al., 2012; Okada et al., 2012, 2014; Stahl et al., 2010). However most of these observations were statistical associations. How to determine these loci as causal factors and how to find the mediation factors underlying these associations were great challenges at present.

DNA methylation, primarily on CpG dinucleotides, is one of the most studied epigenetic modifications. The best-known function of DNA methylation is to regulate nearby gene expression. DNA methylations are most commonly observed in promoter regions where they control transcription of the nearby target genes. Previous studies have found that the changes in DNA methylation patterns in the genome can cause durable changes of gene expression (Ball et al., 2009; Bell et al., 2011; Hadji et al., 2016; Jones, 1999; Kass et al., 1997; Yang et al., 2014). Methylation is a key player in medicine (Dor and Cedar, 2018; Feinberg, 2018). Growing evidence has suggested that heritable DNA methylation plays a role in RA (Ai et al., 2016, 2014; Klein and Gay, 2013; Liu et al., 2013b; Nakano et al., 2013; Zhu et al., 2019). On the other hand, genome-wide mRNA expression profile studies have been widely carried out to identify genes involved in RA etiology (Lee et al., 2011; Teixeira et al., 2009). However, these were small sample studies and always focused on the top significant association signals for follow-up. Such strategy would miss some methylations and/or genes with moderate or small effects. Besides, the effects of methylations or mRNA expressions

* Corresponding author at: Jiangsu Key Laboratory of Preventive and Translational Medicine for Geriatric Diseases, Department of Epidemiology, School of Public Health, Soochow University, 199 Renai Road, Suzhou, Jiangsu, 215123, PR China.
E-mail address: leisf@suda.edu.cn (S.-F. Lei).

https://doi.org/10.1016/j.molimm.2020.12.021
Received 14 June 2020; Received in revised form 13 December 2020; Accepted 14 December 2020
Available online 30 December 2020

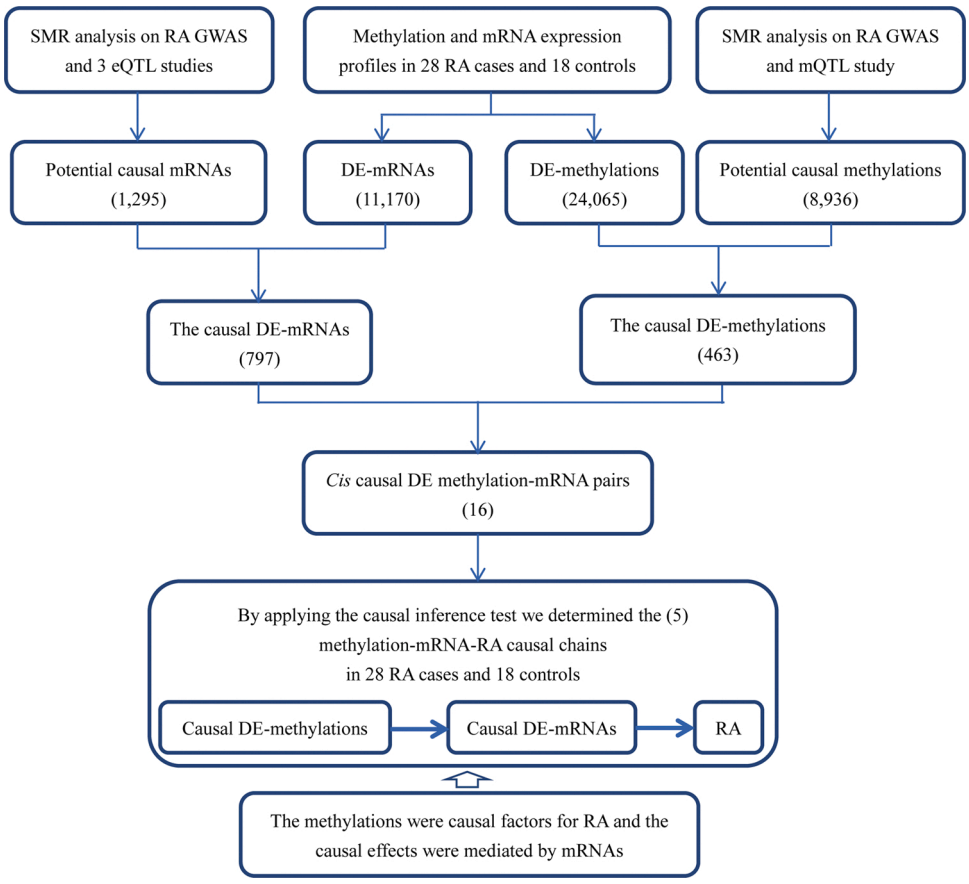

Fig. 1. The flow chart of the study design.
This figure presented the analysis flow and the main results of this study. The numbers in the brackets indicated the count of the identified mRNAs, methylations or the pairs.

on RA were always evaluated in separate studies. Thus, the effects of methylations on expressions of the identified genes and their causal roles in RA (i.e., methylation→mRNA→RA causal chains) were less known.

Genetic variants can affect disease risk through altering gene expressions (Westra et al., 2013). Meanwhile the effect of genetic variants on gene expression can be mediated by DNA methylation (Hannon et al., 2016, 2017). Integration of GWAS data with data from gene expression quantitative trait loci (eQTL) and methylation QTL (mQTL) studies is suggested to be a possible way to prioritize functionally relevant methylation sites or genes in the GWAS loci (Gamazon et al., 2015; Gusev et al., 2016; Zhu et al., 2016). The recently developed two-sample multi-instrumental Mendelian randomization (MR) method is an available way to infer the causal effects between methylations and RA or between mRNAs and RA through instrument variable SNPs (Zhu et al., 2016).

Since the expressions of RA-associated genes are influenced by DNA methylation and the effects of methylations on RA could be mediated by gene expression levels, the methylations and mRNA may be causal factors for RA. This study performed an integrative analysis to identify potential causal methylation-mRNA regulation chains for RA. We first applied a powerful summary data-based MR analysis method, named SMR (Zhu et al., 2016), to identify potential methylations and mRNAs for RA separately. Because the SMR analysis is not feasible at present to infer the casual effects of more than two factors (e.g., methylation-mRNA chain) on RA, then, based on our in-house multi-omics data (methylation and mRNA) in the same sample, we performed causal inference test (CIT) to construct putative methylation→mRNA→RA causal chain. The methylations and genes in the identified chains may be causal factors for RA. This strategy may increase the efficiency of identification of functionally relevant genes for RA.

2. Methods

2.1. Study design

The analysis flow of study was presented in Fig. 1. First, we identified differentially expressed (DE) mRNAs and methylations (denoted as DE-mRNAs and DE-methylations, respectively). Second, we performed SMR analysis to detect potential causal mRNAs and methylations for RA by integrating data from RA GWAS with data from eQTL and mQTL studies and then the causal DE-mRNAs and DE-methylations were determined. And then the *cis* causal DE methylation-mRNA pairs were found out. Finally, for the *cis* pairs we performed CIT to determine if the causal effects of methylations on RA were mediated by mRNAs.

2.2. Study sample

A total of 28 female patients suffering from RA and 18 female subjects without RA or other autoimmune diseases were included into this study. The 46 individuals were Han Chinese individuals living in Suzhou city of the Jiangsu province. The RA patients were diagnosed as RA following the 2010 American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) criteria (Aletaha et al., 2010). The EULAR Disease Activity Score (DAS28) (Prevoo et al., 1995) of the 28 patients ranged from 2.91 to 6.41. All of the cases and controls were female and the age was matched (mean ages were 47.39 (±10.71) and

47.11 (±14.09) for the patients and the controls, respectively) (P = 0.75). Any RA patient or control with the following diseases and/or conditions were excluded: taking antibiotics, consuming probiotics, or having a known history of inflammatory bowel disease or other autoimmune diseases like systemic lupus erythematosus, diabetes and multiple sclerosis.

Peripheral blood (15 mL) was collected by phlebotomy and stored in sodium citrate supplemented vacuum tubes. Peripheral blood mononuclear cells (PBMCs) were isolated by density gradient centrifugation using Lymphoprep (Sigma, life science, USA) within 4 h after phlebotomy, and then treated with TRIzol reagent (invitrogen, Carlsbad, California, USA) and stored at -80℃ to avoid RNA degradation. Total RNA and DNA were extracted from PBMCs according to the instructions recommended by the manufacture and then quantified by using NanoDrop ND-1000 (Thermo Scientific, Wilmington, Delaware, USA) spectrophotometer. The study was approved by the ethical committee of Soochow University. The written informed consent was obtained from all the subjects.

2.3. Genome-wide expression profiling

DNA methylation profiling was performed using Illumina 450 K Infinium Methylation BeadChip (Illumina, Inc., USA) according to the manufacturer's instructions. The methylation level of a locus was measured as $\beta = M/(M + U)$ where M was the methylated signals and U was the methylation level of the unmethylated signals. This methylation value β continuously ranges from 0 (unmethylated) to 1 (fully methylated). Quality control processes that were used to identify DE-methylations have been described in our previous paper (Zhu et al., 2019). After normalization, methylation β values at 484,770 methylation sites for the 46 samples were available for the analysis.

Genome-wide mRNA expression was profiled using Human Gene Expression Microarray V4.0 (CapitalBio Corp, Beijing, China). The data was extracted by Agilent Feature Extraction (V10.7) and was processed using GeneSpring GX software (V12.0). Then, $\log_2$ transformation was applied to the data using the Adjust Data function of Multi experiment Viewer software. Probes with detection rate less than 80 % and/or incomplete annotation information were filtered out. Finally, a total of 21,323 unique mRNA probes were left and used for the analysis. To increase the probability to detect the potential factors, methylations and RNA probes with a loose cutoff ($P < 0.05$) were considered to be DE-methylations and DE-mRNAs and included in further analysis.

2.4. Mendelian randomization analysis

The SMR (http://cnsgenomics.com/software/smr/) is an effective method that integrates summary data from published GWAS with data from QTL studies (Zhu et al., 2016). These method applies the principles of MR (Davey Smith and Hemani, 2014; Smith and Ebrahim, 2003) to integrate independent GWAS summary statistics data with QTL data to identify potential functionally relevant genes at the loci identified in GWAS at genome-wide scale (Pavlides et al., 2016; Zhu et al., 2016). Default parameters were used in SMR analysis. Genotype data of HapMap r23 CEU was used as reference panel to calculate the linkage disequilibrium correlation for SMR analysis. Heterogeneity in dependent instruments (HEIDI) test for heterogeneity in the resulting association statistics was performed. $P_{HEIDI} > 0.05$ means that there was no significant heterogeneity underlying the QTL signals. The genome-wide significance level for the SMR test was defined as 5.0×10^{-6}. Those probes with $P_{SMR} < 5.0 \times 10^{-6}$ and little evidence of heterogeneity ($P_{HEIDI} > 0.05$) were used to highlight genome-wide significant genes.

GWAS have examined the associations of millions SNPs with RA. A large scale GWAS comprised 19,234 RA cases and 61,565 controls has been published (Okada et al., 2014). Among them, 14,361 RA cases and 43,923 controls were Europeans and 4873 RA cases and 17,642 controls were Asians. The datasets downloaded contain the association P values of almost 6.6 million SNPs for Europeans, Asians and the overall samples (http://plaza.umin.ac.jp/~yokada/datasource/software.htm). The dataset for Europeans was used in this analysis.

The eQTL summary data from three studies were used in our study. Westra et al. performed the largest eQTL meta-analysis so far in peripheral blood samples of 5311 European healthy individuals (Westra et al., 2013). The genetic architecture of gene expression (GAGE) study detected eQTLs in peripheral blood in 2765 European individuals (Lloyd-Jones et al., 2017). The cis-eQTL summary data from the GTEx whole blood (Battle et al., 2017) were used. Only SNPs within 1Mb of the transcription start site are available for these two GTEx datasets. The mQTL summary data from the study conducted by McRae et al., which measured mQTL in the Brisbane Systems Genetics Study (n = 614) and the Lothian Birth Cohorts (n = 1366)(McRae et al., 2018). These QTL summary datasets can be downloaded at http://cnsgenomics.com/software/smr/#DataResource.

2.5. Causal inference test

The CIT is an important powerful method for detection of molecular mediators of the effects of causal factors on disease phenotypes (Liu et al., 2013a; Millstein et al., 2009). For the *cis* methylation-mRNA pairs (the methylation locates in the mRNA), the relationship of causal DE-methylation, causal DE-mRNA and RA in each trio were further assessed by CIT to test the causal effect of methylation and the mediation effect of mRNA on RA. The CIT is a model selection approach based on conditional correlation, in which causality can be inferred if the following four conditions are met (Schadt et al., 2005): 1) methylation and RA is associated (P value 1 < 0.05); 2) methylation is associated with mRNA after adjusting for RA (P value 2 < 0.05); 3) mRNA is associated with RA after adjusting for methylation (P value 3 < 0.05); and 4) methylation is independent of RA after adjusting for mRNA (P value 4 > 0.05).

2.6. Differential expression analysis

Additionally, for DE-mRNAs in putative causal chains identified by CIT, we validated the associations based on public data. Two gene expression datasets, GSE17755 (peripheral blood cells)(Lee et al., 2011) and GSE15573 (PBMCs)(Teixeira et al., 2009), were downloaded from GEO database (http://www.ncbi.nlm.nih.gov/geo). GSE17755 contained data of gene expression levels in peripheral blood cells from 112 RA patients and 45 controls (Lee et al., 2011). GSE15573 contained data of gene expression levels in PBMCs from 18 patients and 15 control subjects (Teixeira et al., 2009). Differential expression was tested by comparing mean gene expression signals between cases and controls using *t*-test.

3. Results

3.1. RA-associated causal mRNAs and methylations

Based on our in-house methylation and mRNA expression data in PBMCs of 28 RA cases and 18 controls, we performed differential expression tests and found 11,170 mRNAs (DE-mRNAs) and 24,065 methylation sites (DE-methylations) that were potentially associated with RA by adopting a loose cutoff ($P < 0.05$). These mRNAs and methylations were considered in the subsequent screening analyses. The SMR analysis identified 1295 potential causal mRNAs underlying RA GWAS association ($P_{SMR} < 0.05$) and 197 of them passed the significance level of 5×10^{-6}. Of these mRNAs, 41 passed the HEIDI tests ($P_{HEIDI} > 0.05$), suggesting that there was no heterogeneity, and the expression of these genes and RA were affected by the same variant. Among the 1295 potential causal mRNAs, 797 were DE-mRNAs.

Besides, the SMR analysis identified 8936 potential causal methylations underlying RA GWAS association ($P_{SMR} < 0.05$) and 1514 of them

Table 1
The cis methylation-mRNA pairs.

Methylation							mRNA							
			DE_*P* value	SMR				DE		SMR			*P* DE validation	
Site	CHR	Position[#]		beta	$P_{\_SMR}$	$P_{\_HEIDI}$	Gene	*P* value	Regulation	beta	$P_{\_SMR}$	$P_{\_HEIDI}$	GSE17755	GSE15573
cg25341925	1	11,908,058	4.04E-02	−0.0544	1.11E-02	4.22E-01	*NPPA*	2.72E-04	up	0.1052	1.24E-02	5.74E-01	3.90E-07	–
cg12432846	1	85,741,899	1.43E-03	−0.0398	4.45E-02	1.58E-01	*BCL10*	1.01E-03	up	−0.1437	4.74E-02	8.68E-01	3.57E-03	–
cg02934600	1	107,600,376	3.92E-02	−0.2133	1.01E-03	5.75E-01	*PRMT6*	6.66E-02	up	0.1019	8.89E-03	3.24E-01	–	8.30E-03
cg27576241	1	155,034,570	2.37E-02	0.3036	9.24E-03	6.19E-02	*ADAM15*	2.45E-02	down	−0.0849	7.00E-03	8.96E-02	2.55E-03	–
cg11674355	2	65,610,261	2.19E-02	0.1734	2.04E-09	1.23E-01	*SPRED2*	7.54E-05	up	−0.1913	4.65E-02	1.49E-01	–	–
cg09521647	2	100,722,408	1.72E-02	−0.1606	1.21E-03	3.38E-02	*AFF3*	4.92E-03	down	0.5414	6.25E-05	9.72E-02	–	–
cg18144655	2	114,007,341	3.55E-03	−0.0982	4.58E-02	6.69E-01	*PAX8*	2.44E-02	down	0.0772	4.20E-03	5.25E-01	1.92E-07	–
cg13621396	4	26,861,261	3.66E-02	0.1258	1.80E-02	4.82E-01	*STIM2*	9.36E-05	up	0.1176	4.53E-02	9.31E-02	3.57E-02	9.65E-04
cg23953133	5	126,112,875	4.07E-02	0.1015	3.76E-03	4.56E-02	*LMNB1*	2.49E-04	up	0.1014	8.28E-03	7.08E-01	1.17E-02	–
cg19110902	6	30,698,936	3.13E-02	0.2432	2.40E-03	1.55E-10	*FLOT1*	3.81E-02	up	1.1212	5.36E-07	5.08E-02	4.20E-02	1.47E-02
cg25204543	6	91,005,756	1.64E-02	0.1910	5.86E-03	2.13E-01	*BACH2*	1.19E-02	up	−0.3898	7.07E-03	2.24E-01	2.99E-03	1.75E-02
cg15176413	7	4,764,443	1.48E-02	0.0834	4.36E-02	5.12E-02	*FOXK1*	7.60E-04	up	0.2531	4.48E-03	2.87E-01	2.45E-06	4.76E-02
cg26800803	7	4,765,095	2.61E-02	0.1050	4.49E-02	1.52E-01	*FOXK1*	7.60E-04	up	0.2061	4.81E-03	3.98E-01	2.45E-06	4.76E-02
cg13985518	7	27,143,788	3.41E-02	−0.1528	2.88E-02	1.11E-02	*HOXA2*	3.59E-02	down	−0.0815	1.30E-02	2.11E-01	2.25E-06	–
cg24861686	8	11,418,058	2.08E-02	−0.1354	9.49E-03	1.45E-01	*BLK*	1.32E-03	down	−0.1949	2.84E-11	7.10E-02	–	–
cg26820168	8	124,217,896	1.25E-02	0.0519	2.34E-02	3.24E-01	*FAM83A*	1.50E-03	down	0.2334	2.70E-02	2.77E-01	–	–
cg19640130	10	64,028,056	8.27E-03	0.1752	5.57E-04	6.04E-01	*RTKN2*	1.84E-02	up	0.1464	1.93E-03	6.21E-02	–	–
cg03108697	11	9,732,066	2.39E-02	−0.1278	9.36E-03	4.63E-01	*SWAP70*	3.17E-03	up	−0.1705	6.62E-05	8.13E-01	–	–
cg14911132	11	61,596,755	3.69E-02	0.1871	6.30E-06	2.76E-02	*FADS2*	1.09E-01	up	−0.3362	9.70E-06	2.32E-01	1.54E-03	–
cg01400685	11	61,598,025	3.45E-02	0.1232	3.32E-06	4.67E-02	*FADS2*	1.09E-01	up	−0.0770	4.44E-06	5.14E-01	1.54E-03	–
cg09619347	11	64,107,520	7.11E-03	−0.0519	1.10E-02	2.95E-02	*CCDC88B*	7.92E-04	down	0.1221	1.62E-04	2.79E-01	4.02E-22	–
cg25186680	15	90,931,462	2.43E-02	0.0611	4.53E-02	8.16E-03	*IQGAP1*	3.74E-03	up	0.1728	3.80E-04	6.48E-02	9.27E-12	4.35E-02
cg05799962	16	31,404,385	2.69E-02	−0.0341	4.00E-02	4.71E-02	*ITGAD*	7.98E-03	up	0.1667	4.81E-02	3.83E-01	–	–
cg13876325	17	16,190,098	2.64E-02	−0.2680	4.57E-03	3.72E-01	*PIGL*	5.27E-02	up	0.3072	1.70E-03	2.14E-01	1.54E-42	–
cg07392740	17	73,061,848	2.58E-02	0.1056	2.42E-02	1.14E-03	*KCTD2*	9.31E-05	down	0.1172	1.02E-02	4.04E-01	1.24E-14	–
cg12648893	17	78,305,896	1.94E-02	0.1852	1.27E-02	7.47E-01	*RNF213*	6.21E-03	up	−0.0892	1.04E-02	2.55E-01	6.75E-09	–
cg25786273	18	74,801,026	3.80E-02	0.2308	3.43E-03	9.53E-02	*MBP*	4.58E-04	up	−0.0573	7.47E-04	6.05E-02	–	1.77E-02
cg26312951	21	42,797,847	4.66E-03	0.0914	8.07E-03	8.13E-01	*MX1*	5.34E-03	up	−0.0512	7.15E-03	8.87E-01	1.18E-08	–
cg19713460	22	39,745,530	9.44E-03	−0.4705	1.70E-05	6.14E-01	*SYNGR1*	4.94E-03	down	0.2761	1.61E-09	5.05E-02	3.30E-19	3.27E-02
cg24268161	22	39,747,459	3.79E-02	−0.10187	3.78E-10	5.78E-01	*SYNGR1*	4.94E-03	down	0.1945	1.89E-09	4.03E-01	3.30E-19	3.27E-02

GRCh37.p13.
CHR: Chromosome; DE: Differential expression; SMR: summary statistic data-based Mendelian randomization; HEIDI: Heterogeneity in dependent instruments.

Table 2
The identified potential causal relationships among methylations, mRNAs and RA.

Methylation			mRNA					*P* cor	*P* value 1	*P* value 2	*P* value 3	*P* value 4
Probe ID	CHR	Position[†]	Probe ID	Gene	CHR	Start[a]	End[a]					
cg25204543	6	91,005,756	A_33_P3302916	*BACH2*	6	90,636,247	91,006,627	5.65E-03	1.64E-02	2.98E-02	9.02E-03	0.1418
cg25786273	18	74,801,026	A_24_P402080	*MBP*	18	74,690,789	74,844,774	7.00E-03	3.80E-02	4.81E-02	2.84E-04	0.4952
cg26312951	21	42,797,847	A_23_P17663	*MX1*	21	42,792,520	42,831,141	7.32E-09	4.66E-03	4.28E-07	4.21E-02	0.4868
cg24268161	22	39,747,459	A_23_P348063	*SYNGR1*	22	39,745,954	39,781,593	2.19E-04	3.79E-02	2.03E-03	9.57E-03	0.6066
cg24268161	22	39,747,459	A_23_P211522	*SYNGR1*	22	39,745,954	39,781,593	1.57E-09	3.79E-02	1.96E-08	4.22E-02	0.8911

P cor: the *P* value for the correlation between methylation and mRNA.
P value 1: the crude *P* value for the association between methylation and RA;
P value 2: the *P* value for the association between methylation and mRNA after adjusting for RA;
P value 3: the *P* value for the association between mRNA and RA after adjusting for methylation;
P value 4: the *P* value for the association between methylation and RA after adjusting for mRNA.
[a] Assembly: GRCh37.p13. CHR: Chromosome.

passed the significance level of 5×10^{-6}. Of these methylations, 104 passed the HEIDI tests. Among the 8936 potential causal methylations, a total of 463 were DE-methylations. These 463 potential causal DE-methylations located in 275 mRNAs according to the physical positions of methylation sites in the corresponding genes. Thirty-five of these 275 mRNAs were causal mRNAs, and 30 of them were DE-mRNAs. Thus, in these potential causal DE-mRNAs and DE-methylations identified by expression profile ($P < 0.05$) and SMR analysis ($P_{SMR} < 0.05$), we defined 30 *cis* methylation-mRNA pairs (Table 1). Among these pairs, mRNAs in 16 pairs showed consistent effects on RA risk in both SMR and DE-mRNAs analyses (Table 1).

3.2. Causal chains

For the 16 *cis* methylation-mRNA pairs we set up methylation-mRNA-RA trios and further performed CIT analysis to test whether the trios can form methylation→mRNA→RA causal chains. Based on the conditional correlation results, the mediation effect of mRNA on association between methylation and RA were inferred. A total of 5 chains containing 4 methylations and 5 mRNAs probes in four genes (*BACH2*, *MBP*, *MX1* and *SYNGR1*) were inferred as methylation→mRNA→RA causal chains (Table 2). These methylations and mRNAs all passed the HEIDI tests.

Among these causal chains, the most significant causal gene was *SYNGR1*. The methylation level of cg24268161 and expression levels of *SYNGR1* were found to be causally associated with RA by SMR analysis ($P_{SMR} = 3.78 \times 10^{-10}$ and 1.89×10^{-9}) (Fig. 2A). The methylation level of cg24268161 ($P = 0.04$) and expression level of *SYNGR1* ($P = 4.94 \times 10^{-3}$) were significantly different between RA cases and controls (Fig. 2B) and were strongly correlated in PBMCs (Fig. 2C). The methylation cg24268161 locates in the CpG Island that overlaps a promoter, the so-called GeneHancer regulatory element(Fishilevich et al., 2017), in the 5′ region of *SYNGR1*, suggesting a regulatory potential of these methylations. Moreover, this region shows long-range interactions with four enhancer and another promoter of *SYNGR1* (Fig. 2D). The relationship among cg24268161, *SYNGR1* mRNA and RA met the criteria of the CIT and formed cg24268161→*SYNGR1*→RA causal chains (Table 2).

Similar evidence was found for cg25204543 and *BACH2* (Supplementary Fig. 1). SNPs in *BACH2* gene were associated with the methylation level of cg25204543 and RA risk. And the methylation level of cg25204543 was causally associated with RA according to the SMR analysis ($P_{SMR} = 5.86 \times 10^{-3}$). SNPs in this gene were associated with the *BACH2* expression levels and RA risk. The expression level of *BACH2* were associated with RA ($P_{SMR} = 7.07 \times 10^{-3}$). The methylation level of cg25204543 ($P = 0.02$) and expression of *BACH2* ($P = 0.01$) were different between RA cases and controls and were highly correlated. The methylation cg25204543 locates in the CpG Island that overlaps a promoter. The promoter region interacts with one nearby promoters and twelve enhancers.

SNPs in *MBP* gene were associated with the methylation level of cg25786273 and gene expression levels and RA risk. The methylation level of cg25786273 ($P_{SMR} = 3.43 \times 10^{-3}$) and expression level of *MBP* ($P_{SMR} = 7.47 \times 10^{-4}$) were causally associated with RA (Supplementary Fig. 2). The methylation level of cg25786273 ($P = 0.04$) and expression of *MBP* ($P = 4.58 \times 10^{-4}$) were different between RA cases and controls and were highly correlated. The methylation cg25786273 locates near the CpG Island that overlaps a promoter. This promoter interacts with another promoter that interacts with seven enhancers.

SNPs in *MX1* gene were associated with the methylation level of cg26312951 and gene expression levels and RA risk. According to the SMR analysis the methylation level of cg26312951 ($P_{SMR} = 8.07 \times 10^{-3}$) and expression level of *MX1* ($P_{SMR} = 7.15 \times 10^{-3}$) were associated with RA (Supplementary figure 3). The methylation level of cg26312951 ($P_{SMR} = 4.66 \times 10^{-3}$) and expression of *MX1* ($P_{SMR} = 5.34 \times 10^{-3}$) were different between RA cases and controls and were strongly correlated. The methylation cg26312951 locates near the CpG Island that overlaps a promoter. This promoter interacts with another promoter of *MX1* that interacts with one enhancer of *MX1*.

In addition, based on data from the GEO databases, we found evidence to show that *SYNGR1*, *BACH2*, *MBP* and *MX1* were DE between RA and controls in blood cells (Supplementary figure 4). However, it appears that the direction of effect was opposite in some instances. For example, *BACH2*, *MBP* and *MX1* expressions were increased in RA versus controls in our in-house data (Supplementary Fig. 1 to 3) but were decreased in RA versus controls in public data (Supplementary figure 4). Only the effect of *SYNGR1* expression in PBMCs was found to be consistent in both the in-house (Fig. 2B) and public (Supplementary figure 4, GSE15573) data. Therefore, taking all of these evidences together, we could infer that *SYNGR1* promoter methylations most probably affect mRNA expressions and causally affect RA risk.

4. Discussion

The current study represented an effort to identify RA-associated genes by integrating GWAS data with QTL data using the SMR method and further identify causal factors such as methylation-mRNA chains based on expression profile data using the CIT method. We highlighted potential causal methylation-mRNA chains for RA. The identified methylations and mRNAs were highly connected and seemed to be functionally related to RA.

GWAS have confirmed many susceptibility loci for RA. However, it is difficult to elucidate the causal factors among them. The two-sample MR methods such as SMR use QTLs as instrumental variables to test whether gene expression level has a causal effect on RA. MR analysis presents several advantages over traditional observational epidemiology studies, including the ability to control for environmental confounders and to evaluate the impact of an exposure on an outcome without necessitating the measurement of the exposure and outcome in individuals of the same group. This is a feasible approach to integrate multi-omics data

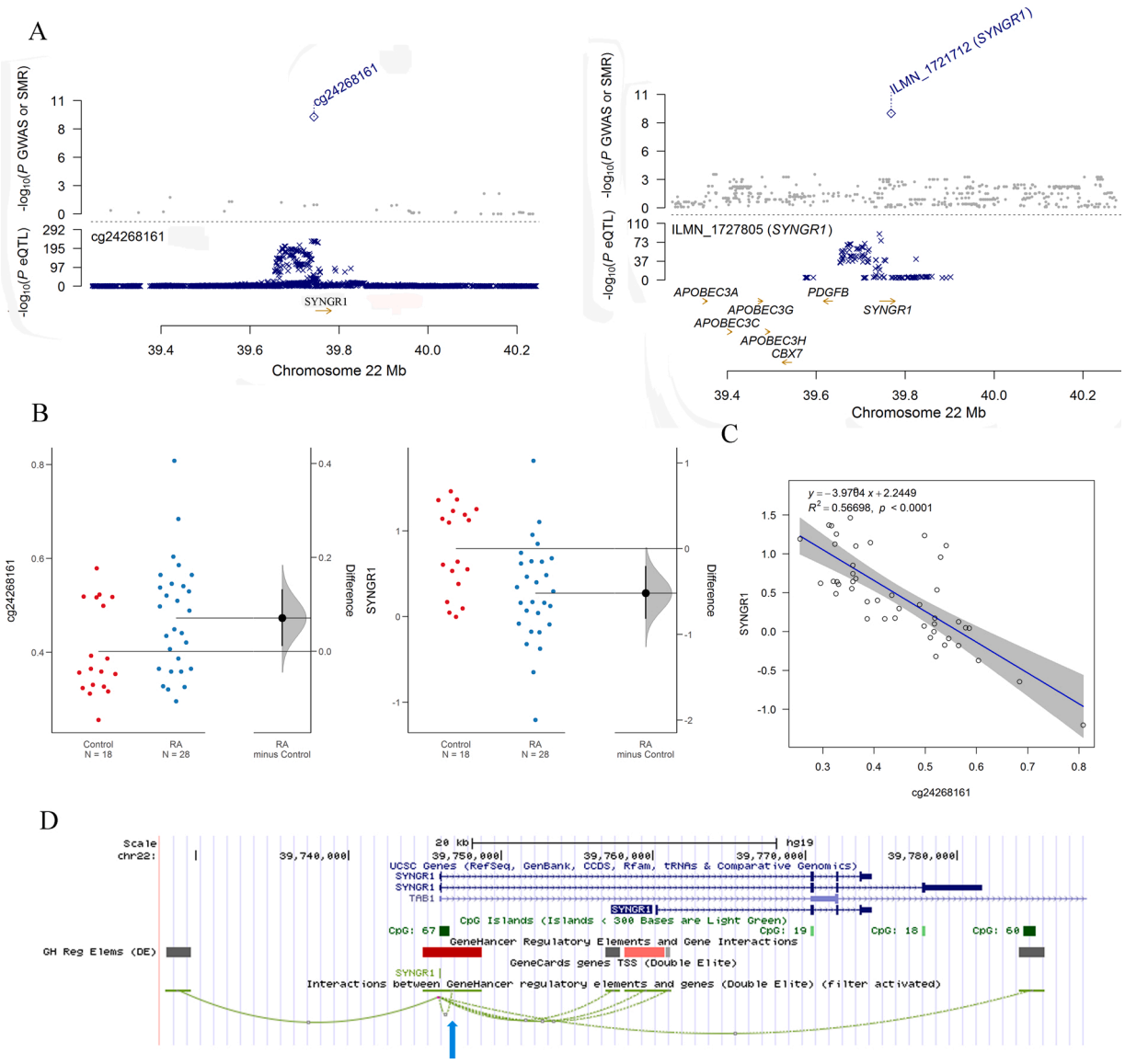

Fig. 2. The relationships among cg24268161, *SYNGR1* and RA.
(A) SNPs in this gene were associated with the methylation level of cg24268161 and gene expression levels and RA risk. According to the SMR analysis the methylation level of cg24268161 and expression level of *SYNGR1* were associated with RA.
(B) The methylation expression level of cg24268161 and gene expression of *SYNGR1* were different between RA cases and controls. The estimation graphics was generated by using the DABEST program (Ho et al., 2019). In each of the estimation graphics there are two plots. The left scatter plot shows the observed methylation/gene expression levels in PBMCs of RA cases and controls. The right part is the estimation plot for the difference of the methylation/gene expression levels between cases and controls. An estimation graphic uses the difference axis to display an effect size (the difference). The curve indicates the resampled distribution of the difference given the observed data. Horizontally aligned with the mean of the RA group, the difference is indicated by the black dot. The 95 % confidence interval of the difference is illustrated by the black vertical line.
(C) The methylation level of cg24268161 and expression level of *SYNGR1* in PBMCs were highly correlated. The scatter plot presents the relationship between the levels of the methylations and mRNAs. The blue line represents the linear regression line and the gray area represents the 95 % confidence intervals.
(D) The genomic region of this site shows long-range interactions. The arrow points to the genomic position of cg24268161. In the track named "Enhancers and promoters from GeneHancer", red and gray colors represent promoters and enhancers, respectively. Methylation cg24268161 (chr22: 39,747,459) locates in a promoter (chr22:39744856−39748743) with high confidence score and interacts with nearby promoters and enhancers.

and this integration approach facilitates the discovery of causal factors underlying the GWAS-identified loci. In addition, CIT is commonly used method to detect causal effects. Previous CIT study has found the mediation effect of mRNA levels on the association between methylation and RA (Zhu et al., 2019). Combined application of MR approach and CIT is an effective strategy not only can help identifying causal factors, but also can strengthen the evidence of the causal relationships. As we shown in this study that by applying this strategy we have identified a list of methylations and genes as potential causal factors for RA.

The genes highlighted in this study may play critical roles in RA pathology. We found some well-known RA-related genes such as *SYNGR1*(Okada et al., 2014), *BACH2*(McAllister et al., 2013), *MBP* (Terao et al., 2011) and *MX1*(Eloranta et al., 2007; Nordborg et al., 2009; Zhu et al., 2019). *SYNGR1* (Synaptogyrin 1) encodes an integral membrane protein associated with presynaptic vesicles in neuronal cells. It may be involved in the regulation of short-term and long-term synaptic plasticity (Janz et al., 1999). *BACH2* (BTB Domain and CNC Homolog 2) is a transcriptional regulator that is related to NF-kappa-B Signaling pathways (Afzali et al., 2017). MBP (Myelin Basic Protein) are present in the bone marrow and the immune system(Li et al., 2005;

Mantovani et al., 2010). MX1 (Interferon-Regulated Resistance GTP-Binding Protein MxA) has antiviral activity against a wide range of RNA viruses and some DNA viruses(Mantovani et al., 2010). The cg26312951→*MX1* expression→RA causal chain has been shown in previous study (Zhu et al., 2019).

Although these protein coding genes have been well-confirmed by GWAS and showed very close relationship with RA, the regulatory processes and the mechanisms is far from understood. DNA methylation is one of the most important elements during the transcription process. The identified methylations locate in the CpG Island which overlaps the GeneHancer regulatory elements. Moreover, this genomic regions show long-range interactions with other nearby enhancers and promoters, which are critical elements in transcription. The identified methylations inside these interaction sites may be critical functional regulatory elements that may have effects on disease risk (Flavahan et al., 2016; Park et al., 2017; Schoenfelder and Fraser, 2019). Therefore, the identified RA-associated promoter methylations may be potentially functional elements that affect RA risk. The identified methylations and genes can be suggested as important candidates for RA functional studies. The role of these methylations in the RA pathology is suggested to be clarified in future studies.

The present study has some potential limitations. First, we were unable to validate the results of the SMR analysis in an independent sample because of lacking in data. Second, the sample size of our methylation and mRNA expression profile studies is relatively small. So the results of the differential expression analysis and CIT were needed to validate, even though we have validated same genes by using public data. Third, some factors (e.g., cell counts/proportions) that varied between patients with an autoimmune disease and healthy controls were not accounted for and the impacts on methylation were unable to assess in this study. Finally, although our study strategy was able to identify causal chains, 14/30 of *cis* methylation-mRNA pairs showed inconsistent effects on RA risk in SMR versus DE-mRNAs, and inconsistent effects are also seen in the validation data. The difference in tissue type and ethnicity as well as treatments of study populations may be the reasons and these may also be limitations to the generalisability of the approach.

In summary, the present study provided a new integrative analysis strategy that may increase the efficiency of identification of functionally relevant genes for RA and other diseases. This study found some potential causal methylations and mRNAs for RA and demonstrated that the effect of the GWAS-identified genes on RA may be regulated by methylations. The new findings may provide new insight into the mechanism underlying the GWAS-identified RA susceptibility loci. The identified methylations and mRNAs were potential candidates for RA studies. The roles of these methylations and mRNAs and their interactions in RA pathology needed to be elucidated in future studies.

Author's contribution

SFL planned and designed the study. XBM analyzed the data and write the manuscript. YHZ and SFL revised the manuscript. All authors have read and accepted the final manuscript.

Funding

The study was supported by Natural Science Foundation of China (81872681, 31401079 and 81473046), the Science and Technology Project of Suzhou (SS202050), Natural science fund for colleges and universities in Jiangsu Province (18KJB330005), and a Project of the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Declaration of Competing Interest

The authors declare that they do not have any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.molimm.2020.12.021.

References

Afzali, B., Gronholm, J., Vandrovcova, J., O'Brien, C., Sun, H.W., Vanderleyden, I., Davis, F.P., Khoder, A., Zhang, Y., Hegazy, A.N., Villarino, A.V., Palmer, I.W., Kaufman, J., Watts, N.R., Kazemian, M., Kamenyeva, O., Keith, J., Sayed, A., Kasperaviciute, D., Mueller, M., Hughes, J.D., Fuss, I.J., Sadiyah, M.F., Montgomery-Recht, K., McElwee, J., Restifo, N.P., Strober, W., Linterman, M.A., Wingfield, P.T., Uhlig, H.H., Roychoudhuri, R., Aitman, T.J., Kelleher, P., Lenardo, M.J., O'Shea, J. J., Cooper, N., Laurence, A.D.J., 2017. BACH2 immunodeficiency illustrates an association between super-enhancers and haploinsufficiency. Nat. Immunol. 18, 813–823.

Ai, R.Z., Whitaker, J.W., Boyle, D.L., Tak, P.P., Gerlag, D.M., Wang, W., Firestein, G.S., 2014. Distinctive DNA Methylome Signatures in Early Rheumatoid Arthritis (RA) Synoviocytes Compared with Longstanding (RA) and Other Inflammatory Arthritides. Arthritis Rheumatol. 66. S1229-S1229.

Ai, R., Hammaker, D., Boyle, D.L., Morgan, R., Walsh, A.M., Fan, S.C., Firestein, G.S., Wang, W., 2016. Joint-specific DNA methylation and transcriptome signatures in rheumatoid arthritis identify distinct pathogenic processes. Nat. Commun. 7.

Aletaha, D., Neogi, T., Silman, A.J., Funovits, J., Felson, D.T., Bingham 3rd, C.O., Birnbaum, N.S., Burmester, G.R., Bykerk, V.P., Cohen, M.D., Combe, B., Costenbader, K.H., Dougados, M., Emery, P., Ferraccioli, G., Hazes, J.M., Hobbs, K., Huizinga, T.W., Kavanaugh, A., Kay, J., Kvien, T.K., Laing, T., Mease, P., Menard, H. A., Moreland, L.W., Naden, R.L., Pincus, T., Smolen, J.S., Stanislawska-Biernat, E., Symmons, D., Tak, P.P., Upchurch, K.S., Vencovsky, J., Wolfe, F., Hawker, G., 2010. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League against Rheumatism collaborative initiative. Arthritis Rheum. 62, 2569–2581.

Ball, M.P., Li, J.B., Gao, Y., Lee, J.H., LeProust, E.M., Park, I.H., Xie, B., Daley, G.Q., Church, G.M., 2009. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat. Biotechnol. 27, 361–368.

Battle, A., Brown, C.D., Engelhardt, B.E., Montgomery, S.B., 2017. Genetic effects on gene expression across human tissues. Nature 550, 204–213.

Bell, J.T., Pai, A.A., Pickrell, J.K., Gaffney, D.J., Pique-Regi, R., Degner, J.F., Gilad, Y., Pritchard, J.K., 2011. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol. 12, R10.

Davey Smith, G., Hemani, G., 2014. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum. Mol. Genet. 23, R89–98.

Dor, Y., Cedar, H., 2018. Principles of DNA methylation and their implications for biology and medicine. Lancet 392, 777–786.

Eloranta, M.L., Barbasso Helmers, S., Ulfgren, A.K., Ronnblom, L., Alm, G.V., Lundberg, I. E., 2007. A possible mechanism for endogenous activation of the type I interferon system in myositis patients with anti-Jo-1 or anti-Ro 52/anti-Ro 60 autoantibodies. Arthritis Rheum. 56, 3112–3124.

Eyre, S., Bowes, J., Diogo, D., Lee, A., Barton, A., Martin, P., Zhernakova, A., Stahl, E., Viatte, S., McAllister, K., Amos, C.I., Padyukov, L., Toes, R.E., Huizinga, T.W., Wijmenga, C., Trynka, G., Franke, L., Westra, H.J., Alfredsson, L., Hu, X., Sandor, C., de Bakker, P.I., Davila, S., Khor, C.C., Heng, K.K., Andrews, R., Edkins, S., Hunt, S.E., Langford, C., Symmons, D., Concannon, P., Onengut-Gumuscu, S., Rich, S.S., Deloukas, P., Gonzalez-Gay, M.A., Rodriguez-Rodriguez, L., Arlsetig, L., Martin, J., Rantapaa-Dahlqvist, S., Plenge, R.M., Raychaudhuri, S., Klareskog, L., Gregersen, P. K., Worthington, J., 2012. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. Nat. Genet. 44, 1336–1340.

Feinberg, A.P., 2018. The key role of epigenetics in human disease prevention and mitigation. N. Engl. J. Med. 378, 1323–1334.

Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Kohn, A., Twik, M., Safran, M., Lancet, D., Cohen, D., 2017. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database (Oxford) 2017, 1–17.

Flavahan, W.A., Drier, Y., Liau, B.B., Gillespie, S.M., Venteicher, A.S., Stemmer-Rachamimov, A.O., Suva, M.L., Bernstein, B.E., 2016. Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature 529, 110–114.

Gamazon, E.R., Wheeler, H.E., Shah, K.P., Mozaffari, S.V., Aquino-Michaels, K., Carroll, R.J., Eyler, A.E., Denny, J.C., Nicolae, D.L., Cox, N.J., Im, H.K., 2015. A gene-based association method for mapping traits using reference transcriptome data. Nat. Genet. 47, 1091–1098.

Gusev, A., Ko, A., Shi, H., Bhatia, G., Chung, W., Penninx, B.W., Jansen, R., de Geus, E.J., Boomsma, D.I., Wright, F.A., Sullivan, P.F., Nikkola, E., Alvarez, M., Civelek, M., Lusis, A.J., Lehtimaki, T., Raitoharju, E., Kahonen, M., Seppala, I., Raitakari, O.T., Kuusisto, J., Laakso, M., Price, A.L., Pajukanta, P., Pasaniuc, B., 2016. Integrative approaches for large-scale transcriptome-wide association studies. Nat. Genet. 48, 245–252.

Hadji, F., Boulanger, M.C., Guay, S.P., Gaudreault, N., Amellah, S., Mkannez, G., Bouchareb, R., Marchand, J.T., Nsaibia, M.J., Guauque-Olarte, S., Pibarot, P., Bouchard, L., Bosse, Y., Mathieu, P., 2016. Altered DNA methylation of long noncoding RNA H19 in calcific aortic valve disease promotes mineralization by silencing NOTCH1. Circulation 134, 1848–1862.

Hannon, E., Dempster, E., Viana, J., Burrage, J., Smith, A.R., Macdonald, R., St Clair, D., Mustard, C., Breen, G., Therman, S., Kaprio, J., Toulopoulou, T., Hulshoff Pol, H.E., Bohlken, M.M., Kahn, R.S., Nenadic, I., Hultman, C.M., Murray, R.M., Collier, D.A., Bass, N., Gurling, H., McQuillin, A., Schalkwyk, L., Mill, J., 2016. An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization of genetic associations and differential DNA methylation. Genome Biol. 17, 176.

Hannon, E., Weedon, M., Bray, N., O'Donovan, M., Mill, J., 2017. Pleiotropic effects of trait-associated genetic variation on DNA methylation: utility for refining GWAS loci. Am. J. Hum. Genet. 100, 954–959.

Ho, J., Tumkaya, T., Aryal, S., Choi, H., Claridge-Chang, A., 2019. Moving beyond P values: data analysis with estimation graphics. Nat. Methods 16, 565–566.

Janz, R., Sudhof, T.C., Hammer, R.E., Unni, V., Siegelbaum, S.A., Bolshakov, V.Y., 1999. Essential roles in synaptic plasticity for synaptogyrin I and synaptophysin I. Neuron 24, 687–700.

Jones, P.A., 1999. The DNA methylation paradox. Trends Genet. 15, 34–37.

Kass, S.U., Landsberger, N., Wolffe, A.P., 1997. DNA methylation directs a time-dependent repression of transcription initiation. Curr. Biol. 7, 157–165.

Klein, K., Gay, S., 2013. Epigenetic modifications in rheumatoid arthritis, a review. Curr. Opin. Pharmacol. 13, 420–425.

Lee, H.M., Sugino, H., Aoki, C., Nishimoto, N., 2011. Underexpression of mitochondrial-DNA encoded ATP synthesis-related genes and DNA repair genes in systemic lupus erythematosus. Arthritis Res. Ther. 13, R63.

Li, Y., Huang, Y., Lue, J., Quandt, J.A., Martin, R., Mariuzza, R.A., 2005. Structure of a human autoimmune TCR bound to a myelin basic protein self-peptide and a multiple sclerosis-associated MHC class II molecule. EMBO J. 24, 2968–2979.

Liu, Y., Aryee, M.J., Padyukov, L., Fallin, M.D., Hesselberg, E., Runarsson, A., Reinius, L., Acevedo, N., Taub, M., Ronninger, M., Shchetynsky, K., Scheynius, A., Kere, J., Alfredsson, L., Klareskog, L., Ekstrom, T.J., Feinberg, A.P., 2013a. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. Nat. Biotechnol. 31, 142–147.

Liu, Y., Aryee, M.J., Padyukov, L., Fallin, M.D., Hesselberg, E., Runarsson, A., Reinius, L., Acevedo, N., Taub, M., Ronninger, M., Shchetynsky, K., Scheynius, A., Kere, J., Alfredsson, L., Klareskog, L., Ekstrom, T.J., Feinberg, A.P., 2013b. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. Nat. Biotechnol. 31, 142–147.

Lloyd-Jones, L.R., Holloway, A., McRae, A., Yang, J., Small, K., Zhao, J., Zeng, B., Bakshi, A., Metspalu, A., Dermitzakis, M., Gibson, G., Spector, T., Montgomery, G., Esko, T., Visscher, P.M., Powell, J.E., 2017. The genetic architecture of gene expression in peripheral blood. Am. J. Hum. Genet. 100, 228–237.

MacGregor, A.J., Snieder, H., Rigby, A.S., Koskenvuo, M., Kaprio, J., Aho, K., Silman, A. J., 2000. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum. 43, 30–37.

Mantovani, C., Mahay, D., Kingham, M., Terenghi, G., Shawcross, S.G., Wiberg, M., 2010. Bone marrow- and adipose-derived stem cells show expression of myelin mRNAs and proteins. Regen. Med. 5, 403–410.

McAllister, K., Yarwood, A., Bowes, J., Orozco, G., Viatte, S., Diogo, D., Hocking, L.J., Steer, S., Wordsworth, P., Wilson, A.G., Morgan, A.W., Kremer, J.M., Pappas, D., Gregersen, P., Klareskog, L., Plenge, R., Barton, A., Greenberg, J., Worthington, J., Eyre, S., 2013. Identification of BACH2 and RAD51B as rheumatoid arthritis susceptibility loci in a meta-analysis of genome-wide data. Arthritis Rheum. 65, 3058–3062.

McRae, A.F., Marioni, R.E., Shah, S., Yang, J., Powell, J.E., Harris, S.E., Gibson, J., Henders, A.K., Bowdler, L., Painter, J.N., Murphy, L., Martin, N.G., Starr, J.M., Wray, N.R., Deary, I.J., Visscher, P.M., Montgomery, G.W., 2018. Identification of 55,000 replicated DNA methylation QTL. Sci. Rep. 8, 17605.

Millstein, J., Zhang, B., Zhu, J., Schadt, E.E., 2009. Disentangling molecular relationships with a causal inference test. BMC Genet. 10, 23.

Nakano, K., Whitaker, J.W., Boyle, D.L., Wang, W., Firestein, G.S., 2013. DNA methylome signature in rheumatoid arthritis. Ann. Rheum. Dis. 72, 110–117.

Nordborg, C., Larsson, K., Aman, P., Nordborg, E., 2009. Expression of the class I interferon-related MxA protein in temporal arteries in polymyalgia rheumatica and temporal arteritis. Scand. J. Rheumatol. 38, 144–148.

Okada, Y., Terao, C., Ikari, K., Kochi, Y., Ohmura, K., Suzuki, A., Kawaguchi, T., Stahl, E. A., Kurreeman, F.A., Nishida, N., Ohmiya, H., Myouzen, K., Takahashi, M., Sawada, T., Nishioka, Y., Yukioka, M., Matsubara, T., Wakitani, S., Teshima, R., Tohma, S., Takasugi, K., Shimada, K., Murasawa, A., Honjo, S., Matsuo, K., Tanaka, H., Tajima, K., Suzuki, T., Iwamoto, T., Kawamura, Y., Tanii, H., Okazaki, Y., Sasaki, T., Gregersen, P.K., Padyukov, L., Worthington, J., Siminovitch, K.A., Lathrop, M., Taniguchi, A., Takahashi, A., Tokunaga, K., Kubo, M., Nakamura, Y., Kamatani, N., Mimori, T., Plenge, R.M., Yamanaka, H., Momohara, S., Yamada, R., Matsuda, F., Yamamoto, K., 2012. Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population. Nat. Genet. 44, 511–516.

Okada, Y., Wu, D., Trynka, G., Raj, T., Terao, C., Ikari, K., Kochi, Y., Ohmura, K., Suzuki, A., Yoshida, S., Graham, R.R., Manoharan, A., Ortmann, W., Bhangale, T., Denny, J.C., Carroll, R.J., Eyler, A.E., Greenberg, J.D., Kremer, J.M., Pappas, D.A., Jiang, L., Yin, J., Ye, L., Su, D.F., Yang, J., Xie, G., Keystone, E., Westra, H.J., Esko, T., Metspalu, A., Zhou, X., Gupta, N., Mirel, D., Stahl, E.A., Diogo, D., Cui, J., Liao, K., Guo, M.H., Myouzen, K., Kawaguchi, T., Coenen, M.J., van Riel, P.L., van de Laar, M.A., Guchelaar, H.J., Huizinga, T.W., Dieude, P., Mariette, X., Bridges Jr., S. L., Zhernakova, A., Toes, R.E., Tak, P.P., Miceli-Richard, C., Bang, S.Y., Lee, H.S., Martin, J., Gonzalez-Gay, M.A., Rodriguez-Rodriguez, L., Rantapaa-Dahlqvist, S., Arlestig, L., Choi, H.K., Kamatani, Y., Galan, P., Lathrop, M., Eyre, S., Bowes, J., Barton, A., de Vries, N., Moreland, L.W., Criswell, L.A., Karlson, E.W., Taniguchi, A., Yamada, R., Kubo, M., Liu, J.S., Bae, S.C., Worthington, J., Padyukov, L., Klareskog, L., Gregersen, P.K., Raychaudhuri, S., Stranger, B.E., De Jager, P.L., Franke, L., Visscher, P.M., Brown, M.A., Yamanaka, H., Mimori, T., Takahashi, A., Xu, H., Behrens, T.W., Siminovitch, K.A., Momohara, S., Matsuda, F., Yamamoto, K., Plenge, R.M., 2014. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature 506, 376–381.

Park, S.M., Choi, E.Y., Bae, M., Choi, J.K., Kim, Y.J., 2017. A long-range interactive DNA methylation marker panel for the promoters of HOXA9 and HOXA10 predicts survival in breast cancer patients. Clin. Epigenetics 9, 73.

Pavlides, J.M., Zhu, Z., Gratten, J., McRae, A.F., Wray, N.R., Yang, J., 2016. Predicting gene targets from integrative analyses of summary data from GWAS and eQTL studies for 28 human complex traits. Genome Med. 8, 84.

Prevoo, M.L., van' t Hof, M.A., Kuper, H.H., van Leeuwen, M.A., van de Putte, L.B., van Riel, P.L., 1995. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum. 38, 44–48.

Schadt, E.E., Lamb, J., Yang, X., Zhu, J., Edwards, S., Guhathakurta, D., Sieberts, S.K., Monks, S., Reitman, M., Zhang, C., Lum, P.Y., Leonardson, A., Thieringer, R., Metzger, J.M., Yang, L., Castle, J., Zhu, H., Kash, S.F., Drake, T.A., Sachs, A., Lusis, A. J., 2005. An integrative genomics approach to infer causal associations between gene expression and disease. Nat. Genet. 37, 710–717.

Schoenfelder, S., Fraser, P., 2019. Long-range enhancer-promoter contacts in gene expression control. Nat. Rev. Genet. 20, 437–455.

Smith, G.D., Ebrahim, S., 2003. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int. J. Epidemiol. 32, 1–22.

Smolen, J.S., Aletaha, D., McInnes, I.B., 2016. Rheumatoid arthritis. Lancet 388, 2023–2038.

Stahl, E.A., Raychaudhuri, S., Remmers, E.F., Xie, G., Eyre, S., Thomson, B.P., Li, Y., Kurreeman, F.A., Zhernakova, A., Hinks, A., Guiducci, C., Chen, R., Alfredsson, L., Amos, C.I., Ardlie, K.G., Barton, A., Bowes, J., Brouwer, E., Burtt, N.P., Catanese, J. J., Coblyn, J., Coenen, M.J., Costenbader, K.H., Criswell, L.A., Crusius, J.B., Cui, J., de Bakker, P.I., De Jager, P.L., Ding, B., Emery, P., Flynn, E., Harrison, P., Hocking, L.J., Huizinga, T.W., Kastner, D.L., Ke, X., Lee, A.T., Liu, X., Martin, P., Morgan, A.W., Padyukov, L., Posthumus, M.D., Radstake, T.R., Reid, D.M., Seielstad, M., Seldin, M.F., Shadick, N.A., Steer, S., Tak, P.P., Thomson, W., van der Helm-van Mil, A.H., van der Horst-Bruinsma, I.E., van der Schoot, C.E., van Riel, P. L., Weinblatt, M.E., Wilson, A.G., Wolbink, G.J., Wordsworth, B.P., Wijmenga, C., Karlson, E.W., Toes, R.E., de Vries, N., Begovich, A.B., Worthington, J., Siminovitch, K.A., Gregersen, P.K., Klareskog, L., Plenge, R.M., 2010. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat. Genet. 42, 508–514.

Teixeira, V.H., Olaso, R., Martin-Magniette, M.L., Lasbleiz, S., Jacq, L., Oliveira, C.R., Hilliquin, P., Gut, I., Cornelis, F., Petit-Teixeira, E., 2009. Transcriptome analysis describing new immunity and defense genes in peripheral blood mononuclear cells of rheumatoid arthritis patients. PLoS One 4, e6803.

Terao, C., Ohmura, K., Katayama, M., Takahashi, M., Kokubo, M., Diop, G., Toda, Y., Yamamoto, N., Shinkura, R., Shimizu, M., Gut, I., Heath, S., Melchers, I., Manabe, T., Lathrop, M., Mimori, T., Yamada, R., Matsuda, F., 2011. Myelin basic protein as a novel genetic risk factor in rheumatoid arthritis–a genome-wide study combined with immunological analyses. PLoS One 6, e20457.

Westra, H.J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M.W., Fairfax, B.P., Schramm, K., Powell, J.E., Zhernakova, A., Zhernakova, D.V., Veldink, J.H., Van den Berg, L.H., Karjalainen, J., Withoff, S., Uitterlinden, A.G., Hofman, A., Rivadeneira, F., Hoen, P.A.C., Reinmaa, E., Fischer, K., Nelis, M., Milani, L., Melzer, D., Ferrucci, L., Singleton, A.B., Hernandez, D.G., Nalls, M.A., Homuth, G., Nauck, M., Radke, D., Volker, U., Perola, M., Salomaa, V., Brody, J., Suchy-Dicey, A., Gharib, S.A., Enquobahrie, D.A., Lumley, T., Montgomery, G.W., Makino, S., Prokisch, H., Herder, C., Roden, M., Grallert, H., Meitinger, T., Strauch, K., Li, Y., Jansen, R.C., Visscher, P.M., Knight, J. C., Psaty, B.M., Ripatti, S., Teumer, A., Frayling, T.M., Metspalu, A., van Meurs, J.B. J., Franke, L., 2013. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat. Genet. 45, 1238–1243.

Yang, X., Han, H., De Carvalho, D.D., Lay, F.D., Jones, P.A., Liang, G., 2014. Gene body methylation can alter gene expression and is a therapeutic target in cancer. Cancer Cell 26, 577–590.

Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M.R., Powell, J.E., Montgomery, G.W., Goddard, M.E., Wray, N.R., Visscher, P.M., Yang, J., 2016. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat. Genet. 48, 481–487.

Zhu, H., Wu, L.F., Mo, X.B., Lu, X., Tang, H., Zhu, X.W., Xia, W., Guo, Y.F., Wang, M.J., Zeng, K.Q., Wu, J., Qiu, Y.H., Lin, X., Zhang, Y.H., Liu, Y.Z., Yi, N.J., Deng, F.Y., Lei, S.F., 2019. Rheumatoid arthritis-associated DNA methylation sites in peripheral blood mononuclear cells. Ann. Rheum. Dis. 78, 36–42.