Journal of Neurology (2019) 266:2699–2709
https://doi.org/10.1007/s00415-019-09476-w

ORIGINAL COMMUNICATION

Integrative analysis revealed potential causal genetic and epigenetic factors for multiple sclerosis

Xing-Bo Mo[1,2,3] · Shu-Feng Lei[1,2,3] · Qi-Yu Qian[1,3] · Yu-Fan Guo[4] · Yong-Hong Zhang[1,3] · Huan Zhang[1,3]

Received: 21 March 2019 / Revised: 14 July 2019 / Accepted: 15 July 2019 / Published online: 18 July 2019

Abstract

Objective Many genomic loci have been identified for multiple sclerosis (MS) by genome-wide association studies (GWAS). Discrimination of the most functionally relevant genes in these loci remains challenging. The aim of this study was to highlight potential causal genes for MS.

Methods We detected potential causal DNA methylations and gene expressions for MS by integrating data from large scale GWAS and quantitative trait locus (QTL) studies using the summary data-based Mendelian randomization method. Potential functional SNPs in the identified genes were searched.

Results We found 178 DNA methylation sites and mRNA expressions of 29 genes that were causally associated with MS. The identified genes enriched in 21 specific KEGG pathways and 80 GO terms (e.g., antigen processing and presentation, interferon gamma mediated signaling pathway). Among the identified non-MHC genes, *METTL21B*, *METTL1* and *TSFM* were strongly connected. MS-associated SNPs in *DDR1* were strongly associated with plasma MHC class I polypeptide-related sequence B (MICB) and Granzyme A levels. And plasma MICB and Granzyme A levels were causally associated with MS. Many SNPs in the causal genes showed QTL effects. The association between m^6A-SNPs rs923829 and *METTL21B* expression level was validated in 40 unrelated Chinese Han individuals.

Conclusions This study identified many DNA methylations and genes as important risk factors for MS and provided novel evidence on the association between circulating MICB and Granzyme A and MS. We also showed that the interaction among *DDR1*, *MICB* and *GZMA* and interaction among *METTL21B*, *METTL1* and *TSFM* may participate in the pathogenesis of MS.

Keywords Multiple sclerosis · Genome-wide association study · Methylation · Mendelian randomization

Introduction

Multiple sclerosis (MS) is an autoimmune disorder mainly affecting young adults and characterized by destruction of myelin in the central nervous system. Genetic factors play important roles in the etiology of MS. According to the genome-wide association studies (GWAS) catalog (https://www.ebi.ac.uk/gwas/), a total of 41 MS GWAS have been published and 395 genetic associations have been found for MS. The International Multiple Sclerosis Genetics Consortium (IMSGC) has identified many genetic factors contributing to a susceptibility to MS by GWAS since 2007 [1–5]. The important effects on MS exerted by variants in the human leukocyte antigen (HLA) genes from the major histocompatibility complex (MHC) have been well established by these GWAS, and 103 discrete loci outside of the MHC have been found [1].

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00415-019-09476-w) contains supplementary material, which is available to authorized users.

✉ Huan Zhang
hzhang3@suda.edu.cn

1 Jiangsu Key Laboratory of Preventive and Translational Medicine for Geriatric Diseases, Soochow University, 199 Renai Road, Suzhou 215123, Jiangsu, People's Republic of China

2 Center for Genetic Epidemiology and Genomics, School of Public Health, Soochow University, 199 Renai Road, Suzhou 215123, Jiangsu, People's Republic of China

3 Jiangsu Key Laboratory of Preventive and Translational Medicine for Geriatric Diseases, Department of Epidemiology, School of Public Health, Soochow University, 199 Renai Road, Suzhou 215123, Jiangsu, People's Republic of China

4 Department of Rheumatology, The First Affiliated Hospital of Soochow University, Suzhou 215000, Jiangsu, People's Republic of China

The GWAS exploits the linkage disequilibrium (LD) of the genomic variants so the identified associations will generally point to genomic regions that harbor many genes [6]. Prioritize among these genes to identify the most functionally relevant genes using GWAS data alone is out of the question. Integration of GWAS data with data from gene expression and methylation GWAS, which allowed identification of expression and methylation quantitative trait loci (eQTL and mQTL), respectively, is considered to be a possible way to prioritize relevant genes in the GWAS identified regions [7–9].

Mendelian randomization (MR) is an instrumental variable analysis approach that uses genetic variants as instrumental variables (e.g., QTLs) to test whether an exposure (e.g., gene expression and DNA methylation level) has a causal effect on an outcome (e.g., MS). Zhu et al. have proposed a method called summary data-based MR (SMR) that integrates independent GWAS summary statistics with QTL data to identify potential functionally relevant genes at the loci identified in GWAS [9]. By applying this method, Zhu et al. have successfully identified novel trait-associated genes for several diseases [9, 10]. However, this method has never been applied to MS. In this study, we applied the SMR method to MS to prioritize a list of genes, and further explored the functional relevance of these genes using a combination of data from MS GWAS, gene expression profile, eQTL, mQTL, plasma protein level QTL (pQTL, SNPs associated with plasma metabolite levels) [11], metabolite QTLs (metabQTL, SNPs associated with blood metabolite levels) studies [12] and SNP databases.

Methods

Study design

This study was designed to identify potential causal factors (e.g., genetic variants, DNA methylations, gene expressions and plasma protein levels) for MS. First, we conducted SMR analysis to identified DNA methylations that were causally associated with MS. Second, we conducted SMR analysis to identified gene expressions that were causally associated with MS. Meanwhile, differential expression analysis was performed for the identified genes based on expression profile data from four studies available in the GEO database (https://www.ncbi.nlm.nih.gov/geo). Third, we looked for eQTLs, mQTLs, pQTLs and metabQTLs and identified potential functional variants within the identified causal genes. We also validated some eQTLs in sample from the Chinese Han population. Finally, for proteins that were strongly affected by SNPs within the identified causal genes, we tested if the protein levels were causally associated with MS using MR methods such as GSMR and MR-PRESSO.

MS GWAS datasets

We used two GWAS datasets in this study. The first dataset was from the GWAS conducted by the IMSGC and Wellcome Trust Case Control Consortium 2 (MS GWAS 2011), which comprised 9772 MS cases and 17,376 controls of European ancestry [4]. The second dataset was from the GWAS (ImmunoChip custom genotyping array) conducted by the IMSGC (MS GWAS 2013), which examined the association between 161,311 autosomal variants and MS in 14,498 MS cases and 24,091 healthy controls [1]. These two dataset can be freely downloaded from the ImmunoBase (https://www.immunobase.org/), a web-based resource focused on the genetics and genomics of immunologically related human diseases. The available summary data contained information on the rs number, chromosome, position, alleles, odds ratio and the corresponding confidence interval and *P* values for each SNP. These two datasets were separately analyzed in the following analyses.

QTL datasets

We conducted MR analyses to test for associations between gene expressions (using eQTL data), DNA methylation levels (mQTL data) and plasma protein levels (pQTL data) and MS association signals (MS GWAS data) at genome-wide scale, by which we could prioritize functionally relevant genes in the GWAS loci. The eQTL summary data from four studies were used in our SMR analysis. The first is the study conducted by Westra et al. which is the largest eQTL meta-analysis so far in peripheral blood samples of 5311 European healthy individuals [13]. The second is the genetic architecture of gene expression (GAGE) study which detected eQTLs in peripheral blood in 2765 European individuals [14]. The third and fourth datasets contain the *cis*-eQTL summary data from the GTEx project (whole blood and brain tissue) [15]. 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 ($n=1366$) [16]. The pQTLs data was from a large scale protein level GWAS that measured 509,946 SNPs for genome-wide associations with 1124 protein levels measured in 1000 blood samples from the KORA study [11]. The metabQTLs data were released from a large scale metabolites GWAS, which measured genome-wide associations of 6.69 million common SNPs and 4.66 million low-frequency ($0.5\% \leq \text{MAF} < 5\%$) variants with each of the 644 metabolites in 1960 blood samples [12].

SMR analysis

SMR is a two-sample multi-instrumental MR approach. In this kind of MR methods the exposure and outcome do not

necessary measure in the same samples. It can jointly analyze summary statistics from independent GWAS to test for association between an exposure and a trait due to a shared variant at a locus. The shared SNPs tested in two GWAS provided the effect data (the regression coefficient beta values) of SNPs on exposure and outcome. Using this effect data the MR method can estimate the causal effect of the exposure on outcome.

The SMR method implements a transcriptome-wide association analysis in a formal statistical framework using summary data from large-scale GWAS so that the statistical power is increased by the large sample size. We ran SMR (version 0.712) with default parameters in a command-line program which was downloaded from https://cnsgenomics.com/software/smr/. The requisite data (i.e., SNP rs number, allele 1, allele 2, frequency of allele 1, *P* value and sample size) were extracted from this GWAS dataset and beta and standard error were calculated using the odds ratio and the corresponding confidence interval. This data were organized to the specific format (the.ma file with 8 columns) for the SMR analysis softwares using the R language. The QTL summary data in SMR binary format can be downloaded from https://cnsgenomics.com/software/smr/#DataResource. Genotype data of HapMap r23 CEU were used as a reference panel to calculate the LD correlation for SMR analysis. The genome-wide significance level for the SMR test was set to 5.0×10^{-6}. The heterogeneity in dependent instruments (HEIDI) test for no pleiotropy, the basic assumption of MR study, was conducted to test whether there is a single causal variant affecting MS, methylation or gene expression. Those probes with little evidence of heterogeneity ($P_{\mathrm{HEIDI}}\geq0.05$) were retained. The SMR locus plots were generated using the R code provided by Zhu et al. (https://cnsgenomics.com/software/smr/#SMRlocusplot19). The gene range list glist-hg19 was used for plotting.

Differential expression analysis

We also tried to determine if the expression levels of genes identified by SMR analysis were associated with MS based on expression profile data from four studies available in the GEO database (https://www.ncbi.nlm.nih.gov/geo), as supporting evidence. GSE16461 contained data of gene expression levels in $CD8^+$ T cells from eight MS cases and eight controls. GSE21942 contained data of gene expression levels in peripheral blood mononuclear cells from 14 MS cases and 15 controls [17]. GSE23832 contained data of gene expression levels in peripheral blood mononuclear cells from eight MS cases and four controls [18]. GSE26484 contained data of gene expression levels in $CD4^+$ T cells from six MS cases and four controls [19]. Differential expression was tested by comparing mean gene expression signals between cases and controls using *t* test. The significance level of $P=0.05$ was used for the differential expression analysis.

Relationship between the identified genes

We searched for the connection between the identified genes to clarify the relationship, by which we can find pathways of MS. Functional enrichment analysis was conducted using the DAVID online tool (https://david.ncifcrf.gov/). Protein–protein interactions were found in the STRING (https://string-db.org/) and LENS databases (https://hagrid.dbmi.pitt.edu/LENS/).

Identification of potential regulatory SNPs

We searched for SNPs with specific functions, including the N^6-methyladenosine (m^6A)-associated SNPs (m^6A-SNPs) [20] and phosphorylation-related SNPs (phosSNPs) [21]. m^6A is a kind of reversible RNA methylation which has a crucial role in regulating many fundamental biological processes. The m^6A-SNPs are a kind of SNPs that would influence m^6A by changing the RNA sequences of the target sites or key flanking nucleotides [20]. This kind of SNPs was identified in the m6AVar database (https://m6avar.renlab.org/). Phosphorylation is a kind of important post-translational modifications of proteins. The phosSNPs are nonsynonymous variants that might influence protein phosphorylation status. The phosSNPs were identified in the phosSNP 1.0 database (https://phossnp.biocuckoo.org/index.php).

We carried out QTL analysis for SNPs in the identified genes. The *cis*- and *trans*-QTL could show the relationship between these genes, which was expected to show us clues to begin to understand the regulatory mechanism. SNPs located in the identified genes were obtained from the UCSC database. Then we looked for eQTLs in this SNPs list based on data downloaded from the HaploReg browser (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php). HaploReg is a database which manually collates and updates data from recent eQTL studies [22]. The mQTLs were found from the study conducted by McRae et al. [16]. We looked for pQTLs according to the protein level GWAS from the KORA study [11]. In addition, we looked for metabQTLs from the metabolites GWAS described above [12].

Validation of the eQTLs

To validate the eQTLs we tested genotypes and mRNA expression levels in peripheral blood mononuclear cells (PBMCs) of 40 unrelated Chinese Han individuals (age range from 27 to 67) to obtain additional evidence to support the identified SNPs. These participants did not have diseases such as MS. PBMCs were isolated from 15 ml peripheral blood by density gradient centrifugation using Lymphoprep

(Sigma, life science, USA). Total RNA and DNA were extracted in the same lab according to the instructions recommended by the manufacturer. The mRNA levels and genotypes were tested using RT-PCR method. The mean mRNA levels of individuals with different genotypes were compared and the association between SNP and mRNA level was assessed in linear regression model. SNP was analyzed as 0, 1 or 2 copies of the minor allele in an additive genetic model. The study was approved by the ethical committee of Soochow University. The written informed consent was obtained from all of the participants.

MR analysis of protein levels

To obtain additional supporting evidence, we took the advantage of two recent developed R-packages which each implements the Generalised SMR (GSMR) [23] and MR pleiotropy residual sum and outlier (MR-PRESSO) [24] methods to test for putative causal association between plasma protein levels and MS. In GSMR analysis, the summary data were from the GWAS conducted by Beecham et al. [1], and the summary data of association between SNPs and plasma protein levels ($P < 5 \times 10^{-8}$) were obtained from the pQTL GWAS mentioned above [11]. From the MS GWAS dataset, SNP rs number, allele 1, allele2, frequency of allele 1, beta, standard error, *P* value and sample size were extracted. From the pQTL datasets, SNP rs number, beta, standard error, *P* value and sample size were extracted. Then the two extracted datasets were merged by SNP rs number to form the file with specific format (a plain file with 12 columns) for the GSMR analysis softwares using the R language. The consistency of the effect allele for each SNP in the MS GWAS and pQTL study was checked. SNPs with pQTL *P* value less than 5.0×10^{-6} and not in LD were used as instrumental variables. The HapMap r23 CEU genotype data was used to calculate the LD correlation. The HEIDI test was used to test the no pleiotropy hypothesis. The parameters were left at the default setting in this analysis.

We used an implementation of MR-PRESSO in R (https://github.com/rondolab/MR-PRESSO) to detect horizontal pleiotropy which is the MR-PRESSO global test and the outlier corrected causal estimate [24]. Data used in the MR-PRESSO analysis were the same as the GSMR analysis. The requisite data (i.e., SNP rs number, beta, standard error, and *P* value) were extracted from each of the MS GWAS and pQTL datasets and then merged by SNP to the specific format (a plain file with 7 columns) for the MR-PRESSO analysis using the R language. The consistency of the effect allele for each SNP in the MS GWAS and pQTL study was checked. SNPs with *P* value less than 5.0×10^{-6} in the pQTL study were used. The default values were used for all of the MR-PRESSO analysis parameters.

Results

MS-associated methylation sites

We performed a SMR analysis to determine if DNA methylations were causally associated with MS. By integrating data from large scale MS GWAS and mQTL study, we found 178 methylation sites in 23 loci (83 genes) which were associated with MS ($P_{SMR} < 5 \times 10^{-6}$, $P_{HEIDI} > 0.05$) (Table 1). There are CpG islands in most of the associated genes (Table 1, Supplementary Table S1). The detail information about these 178 MS-associated methylation sites was presented in table S1. $P_{HEIDI} > 0.05$ means that there was no significant heterogeneity underlying the mQTL signals. Among the 23 loci, 21 have been reported in GWAS and 2 (7p22.2 and 19q13.33) have never been reported.

MS-associated genes

We found that expressions of 29 genes in 12 loci were significantly associated with MS ($P_{SMR} < 5 \times 10^{-6}$), and there was no significant heterogeneity underlying the eQTL signals ($P_{HEIDI} > 0.05$) (Table 2). Among these 29 genes, *RMI2*, *JUND* and *CD37* have not been reported to be associated with MS. All of these 29 genes passed HEIDI tests ($P_{HEIDI} > 0.05$), suggesting that there was no heterogeneity, and the expression of these genes and MS were affected by the same variant, suggesting that the same causal SNPs contributed to both MS risk and gene expression. We compared mRNA expression signals in gene expression studies for the 29 genes and found that 17 of them were differentially expressed ($P < 0.05$) (Supplementary Table S2). For the 83 genes identified in mQTL SMR analysis (Table 1), the eQTL SMR tests identified that *DDR1*, *POU5F1*, *PSMB9*, *ZC2HC1A*, *NADSYN1*, *METTL1*, *RMI2* and *TNFSF14* were significantly associated with MS (Table 2).

Potential regulatory SNPs

We further looked for SNPs with specific functions and QTLs in the 29 genes which were found in the eQTL SMR analysis. A total of 14,309 SNPs were obtained from the UCSC database for these genes. We found 13 m^6A-SNPs in this SNP list. Among them, rs923829 in *METTL21B* and rs2288481 in *DKKL1* gene were significantly associated with MS ($P = 1.35 \times 10^{-10}$ and 4.36×10^{-6}). Besides, a total of 55 phosSNPs were found. Among them, nine were associated with MS (Supplementary Table S4).

According to the SMR analysis, SNPs in the identified genes are strongly associated with DNA methylation and gene expression levels. And this eQTLs and mQTLs were

Table 1 Summary of the MS-associated methylations identified by Mendelian randomization analysis

Loci	Genes	Reported genes	CpG sites[¶]	CpG islands[†]
1p22.3	*BCL10*	Yes	1	1
1p22.1	*GFI1, RPL5, FAM69A*	Yes	3	4
1p13.1	*CD58*	Yes	1	2
1q31.2	*RGS1, C1orf106*	Yes	2	2
3q13.33	*ILDR1*	Yes	1	1
5q35.3	*RGS14*	Yes	4	2
6p22.1	*OR12D3, HCG4, RNF39, RNF39, RNF39, TRIM31, TRIM40, TRIM15, HLA-E*	Yes	11	8
6p21.33	*GNL1, TUBB, DDR1, VARS2, SFTA2, HCG22, C6orf15, PSORS1C1, CDSN, POU5F1, HLA-B, MICA, DDX39B, NFKBIL1;ATP6V1G2, C6orf47, LY6G6C, DDAH2, VARS, SLC44A4, EHMT2, TNXB*	Yes	57	23
6p21.32	*PRRT1, PPT2, AGPAT1, RNF5, PBX2, GPSM3, NOTCH4, C6orf10, BTNL2, HLA-DRA, HLA-DQB1, HLA-DQA2, HLA-DQB2, TAP2, PSMB8, TAP1, PSMB9, HLA-DMA, BRD2, SYNGAP1*	Yes	55	19
6q15	*BACH2*	Yes	1	1
7p22.2	*LOC100129603*	No	1	0
8q21.13	*ZC2HC1A*	Yes	1	1
10p15.1	*IL2RA*	Yes	1	0
11q12.2	*CD6*	Yes	1	1
11q13.4	*DHCR7, NADSYN1*	Yes	4	4
12p13.31	*CLECL1*	Yes	1	0
12q14.1	*CDK4, METTL1, AVIL, CTDSP2*	Yes	6	3
14q31.3	*GPR65*	Yes	3	0
16p13.13	*CLEC16A, RMI2*	Yes	3	5
17q22	*YPEL2*	Yes	1	1
19p13.3	*TNFSF14*	Yes	2	0
19p13.11	*MED26, MAST3, PDE4C*	Yes	10	12
19q13.33	*CCDC155*	No	3	1

[¶]Number of methylation sites which were significantly associated with MS in the gene regions

[†]Number of CpG islands in the gene regions according to UCSC genome browser

mostly significantly associated with MS ($P < 5.0 \times 10^{-8}$). pQTLs significantly associated with MS ($P < 5.0 \times 10^{-8}$) were found for six genes, including *DDR1*, *HLA-DQA1*, *POU5F1*, *PSMB9*, *SKIV2L* and *TAGAP* (Table 3, Supplementary Table S5). metabQTLs significantly associated with MS ($P < 5.0 \times 10^{-8}$) were found for *DDR1*, *SKIV2L* and *HLA-DQA1* (Supplementary Table S6).

Relationship between the identified genes

The genes identified by the SMR analyses were found to enrich in 21 specific KEGG pathways (e.g., antigen processing and presentation, rheumatoid arthritis, systemic lupus erythematosus, and others) and 80 GO biological process terms (e.g., antigen processing and presentation of exogenous peptide antigen, interferon-gamma-mediated signaling pathway, and others) (Supplementary Table S3). Genes in the identified terms were mostly the MHC genes, except *GFI1*, *RPL5*, *CD58*, *IL2RA*, *CDK4*, *CTDSP2*, *TNFSF14*, *IL7* and *JUND*.

Interactions between the identified genes were found in the STRING database (Supplementary Fig. S1). Among the non-MHC genes identified by SMR analysis ($P_{SMR} < 5 \times 10^{-6}$, $P_{HEIDI} > 0.05$), *ZC2HC1A*, *METTL21B*, *METTL1* and *TSFM* were strongly connected (Fig. 1). Interactions between other genes and their targets were also found in the LENS database (Supplementary Fig. S2–8). Genes identified in pQTL analysis were highly connected (Supplementary Fig. S9). According to the pQTL analysis, MS-associated SNPs in *DDR1* was strongly associated with plasma Granzyme A and MHC class I polypeptide-related sequence B (MICB) protein levels (Table 3). The *DDR1* gene does not directly interact with other genes according to STRING. But via other interactors, *DDR1* was connected with *GZMA* and *MICB* (Supplementary Fig. S10). We further performed GSMR analysis and found that plasma MICB and Granzyme A protein levels may be causally associated with MS ($P = 0.0032$ and 1.27×10^{-7}), respectively. In MR-PRESSO analysis, the causal association between plasma MICB level (beta = −0.1090, se = 0.0356, $P = 0.0135$) and

Table 2 The MS-associated methylations and mRNA expressions identified by Mendelian randomization analysis

Locus	Gene	Methylation				mRNA expressions					
		Site (count)	Beta	SE	P_{SMR}*	Probe ID	Beta	SE	P_{SMR}	eQTL study	DEG
6p21.33	*DDR1*	13	− 0.8755	0.1154	3.31E − 14	ENSG00000204580.7	− 0.7426	0.1286	7.78E − 09	GTEx_WB	Yes
6p21.33	*POU5F1*	1	− 0.9682	0.1270	2.51E − 14	ENSG00000204531.11	− 0.6089	0.1257	1.26E − 06	GTEx_WB	No
6p21.32	*PSMB9*	2	0.7431	0.1117	2.86E − 11	ENSG00000240065	− 0.6291	0.1160	5.86E − 08	GTEx_Brain	Yes
8q21.13	*ZC2HC1A*	1	0.1971	0.0343	9.32E − 09	ENSG00000104427.7	− 0.2386	0.0485	8.73E − 07	GTEx_WB	Yes
						ILMN_2057981	− 0.1647	0.0274	1.83E − 09	Westra	
11q13.4	*NADSYN1*	1	0.2842	0.0584	1.14E − 06	ENSG00000254682	− 0.0783	0.0136	8.28E − 09	GTEx_Brain	Yes
		1	0.2624	0.0563	3.18E − 06	ENSG00000254682.1	− 0.2243	0.0419	8.89E − 08	GTEx_WB	
12q14.1	*METTL1*	1	0.3941	0.0722	4.75E − 08	ILMN_1723846	− 0.1330	0.0220	1.39E − 09	CAGE	No
16p13.13	*RMI2*	1	− 0.1062	0.0229	3.47E − 06	ENSG00000175643	− 0.0820	0.0176	2.99E − 06	GTEx_Brain	Yes
19p13.3	*TNFSF14*	2	0.1530	0.0219	2.81E − 12	ILMN_1655414	− 0.4087	0.0633	1.08E − 10	Westra	Yes
						ILMN_1655414	− 0.3801	0.0615	6.43E − 10	CAGE	
						ILMN_1661343	− 0.6232	0.1091	1.11E − 08	Westra	

DEG differentially expressed genes, *eQTL* expression quantitative trait locus, *MS* multiple sclerosis, *SE* standard error, *SMR* summary data-based Mendelian randomization, *WB* whole blood

*If there were more than one methylation sites in the locus, the smallest *P* value (and the corresponding beta and SE) was presented

Granzyme A protein level (beta = − 0.1578, se = 0.0481, $P = 0.0065$) and MS was also significant after removing outliers.

Association between m^{6}A-SNPs and gene expression

The m^{6}A-SNPs may take part in gene expression regulation through exerting influence on RNA modification, so they may be associated with gene expression level. Rs923829 in *METTL21B* and rs2288481 (p.Glu183Lys) in *DKKL1* gene were m^{6}A-SNPs associated with MS. We hypothesized that rs923829 and rs2288481 may act through the regulation of gene expression of *METTL21B* and *DKKL1*, respectively. Evidence on these associations can be found in the HaploReg database. Significant association between rs923829 and *METTL21B* expression was found in 37 tissues. Significant association between rs2288481 and *DKKL1* expression was found in 11 tissues.

We validated the association between rs923829 and *METTL21B* expression and the association between rs2288481 and *DKKL1* expression in PBMCs of 40 unrelated Chinese Han individuals. According to our data we found that the mRNA level of *METTL21B* was higher in rs923829 minor allele A carriers than non-carriers. In linear regression model we found that rs923829 was significantly associated with *METTL21B* expression levels ($P = 0.0003$) (Fig. 2a). This synonymous mutation likely leads to a higher *METTL21B* expression level. Rs2288481 was marginally associated with *DKKL1* expression levels in our sample (linear regression $P = 0.0548$) (Fig. 2b). The results suggested that the m^{6}A-SNPs may affect gene expression and lead to MS.

Discussion

The current study represented the first effort to identify potential causal genetic and epigenetic factors for MS by integrating data of several large-scale GWAS initiatives on DNA methylation, gene expression, plasma protein levels and MS risk. Hundreds of important methylation sites, genes and plasma protein levels (e.g., Granzyme A and MICB) were found to be causally associated with MS using our advantage strategy.

GWAS have confirmed hundreds of genetic variants associated with MS. However, elucidating the causal genes underlying GWAS hits remains challenging given the complexities of a typical genome-wide significant locus and the regulatory process. Laboratory-based evaluation of the associated regions by applying high throughput technology with sufficient sample size is costly and hard to achieve at this stage. On the other hand, although traditional case–control studies have identified DNA methylations [25–27] and gene expressions [17–19, 28] implicated in the pathogenesis of

Table 3 The significant associations between SNPs and MS and plasma protein levels

SNP	Gene	Locus	Position (GRCh37.p13)	MAF*	MS GWAS 2013			Protein	Target	Locus	$P_{protein}$
					Beta	SE	*P*				
rs1264344	*DDR1*	6p21.33	30,800,577	0.4408	− 0.1748	0.0164	7.72E − 26	Granzyme A	*GZMA*	5q11.2	2.18E − 07
rs2535331	*DDR1*	6p21.33	30,816,270	0.3502	− 0.1706	0.0174	5.33E − 23	Granzyme A	*GZMA*	5q11.2	5.64E − 07
rs2844662	*DDR1*	6p21.33	30,817,879	0.3673	− 0.1511	0.0168	7.53E − 19	Granzyme A	*GZMA*	5q11.2	4.28E − 08
rs2535326	*DDR1*	6p21.33	30,832,063	0.3673	− 0.1511	0.0172	6.76E − 19	Granzyme A	*GZMA*	5q11.2	2.53E − 08
rs2844654	*DDR1*	6p21.33	30,838,688	0.3508	− 0.1722	0.0169	2.94E − 23	Granzyme A	*GZMA*	5q11.2	3.07E − 07
rs2535339	*DDR1*	6p21.33	30,839,430	0.3508	− 0.1722	0.0173	2.26E − 23	Granzyme A	*GZMA*	5q11.2	6.94E − 07
rs1264334	*DDR1*	6p21.33	30,844,260	0.3673	− 0.1527	0.0172	2.92E − 19	Granzyme A	*GZMA*	5q11.2	4.47E − 08
rs1264333	*DDR1*	6p21.33	30,844,314	0.3508	− 0.1722	0.0173	2.66E − 23	Granzyme A	*GZMA*	5q11.2	6.16E − 07
rs1264331	*DDR1*	6p21.33	30,846,830	0.3502	− 0.1732	0.0173	1.05E − 23	Granzyme A	*GZMA*	5q11.2	6.05E − 07
rs9348843	*DDR1*	6p21.33	30,786,676	0.3206	− 0.1414	0.0179	7.68E − 15	MHC class I polypeptide-related sequence B	*MICB*	6p21.33	4.57E − 06
rs3130787	*DDR1*	6p21.33	30,809,864	0.4060	0.1561	0.0169	2.26E − 20	MHC class I polypeptide-related sequence B	*MICB*	6p21.33	8.24E − 06
rs3095350	*DDR1*	6p21.33	30,817,866	0.4248	0.1406	0.0167	5.40E − 17	MHC class I polypeptide-related sequence B	*MICB*	6p21.33	1.80E − 06
rs3095344	*DDR1*	6p21.33	30,824,649	0.4317	0.1432	0.0194	9.38E − 14	MHC class I polypeptide-related sequence B	*MICB*	6p21.33	1.16E − 06
rs3130791	*DDR1*	6p21.33	30,831,843	0.4322	0.1415	0.0167	2.31E − 17	MHC class I polypeptide-related sequence B	*MICB*	6p21.33	1.72E − 06
rs3130794	*DDR1*	6p21.33	30,832,810	0.4305	0.1406	0.0167	4.03E − 17	MHC class I polypeptide-related sequence B	*MICB*	6p21.33	3.59E − 06
rs3094609	*POU5F1*	6p21.33	31,165,566	0.1452	0.5602	0.0217	3.02E − 145	Granzyme A	*GZMA*	5q11.2	2.18E − 09
rs2280774	*SKIV2L*	6p21.33	31,928,691	0.3195	− 0.3429	0.0184	2.68E − 77	Complement C4a	*C4A*	6p21.33	4.65E − 07
rs2280774	*SKIV2L*	6p21.33	31,928,691	0.3195	− 0.3429	0.0184	2.68E − 77	Complement C4b	*C4B*	6p21.33	7.59E − 29
rs2280774	*SKIV2L*	6p21.33	31,928,691	0.3195	− 0.3429	0.0184	2.68E − 77	Neutrophil collagenase	*MMP8*	11q22.2	5.68E − 07
rs9271488	*HLA-DQA1*	6p21.32	32,589,000	0.2751	− 0.3778	0.0192	7.79E − 86	Beta-2-microglobulin	*B2M*	15q21.1	7.55E − 06
rs9271588	*HLA-DQA1*	6p21.32	32,590,953	0.4607	0.4536	0.0175	7.04E − 154	Beta-2-microglobulin	*B2M*	15q21.1	6.25E − 06
rs642093	*HLA-DQA1*	6p21.32	32,582,075	0.2369	− 0.2070	0.0195	6.46E − 26	MHC class I polypeptide-related sequence B	*MICB*	6p21.33	6.79E − 06
rs3129763	*HLA-DQA1*	6p21.32	32,590,925	0.2204	− 0.1832	0.0200	2.59E − 20	MHC class I polypeptide-related sequence B	*MICB*	6p21.33	5.84E − 06
rs4530903	*HLA-DQA1*	6p21.32	32,581,889	0.1088	− 0.3528	0.0285	4.73E − 35	Neutrophil collagenase	*MMP8*	11q22.2	9.46E − 06
rs9276909	*PSMB9*	6p21.32	32,850,839	0.2660	− 0.1790	0.0192	8.17E − 21	MHC class I polypeptide-related sequence B	*MICB*	6p21.33	2.38E − 07

eQTL expression quantitative trait locus, *MAF* minor allele frequency, *MS* multiple sclerosis, *SE* standard error, *SNP* single nucleotide polymorphism

*Minor allele frequency in the European populations

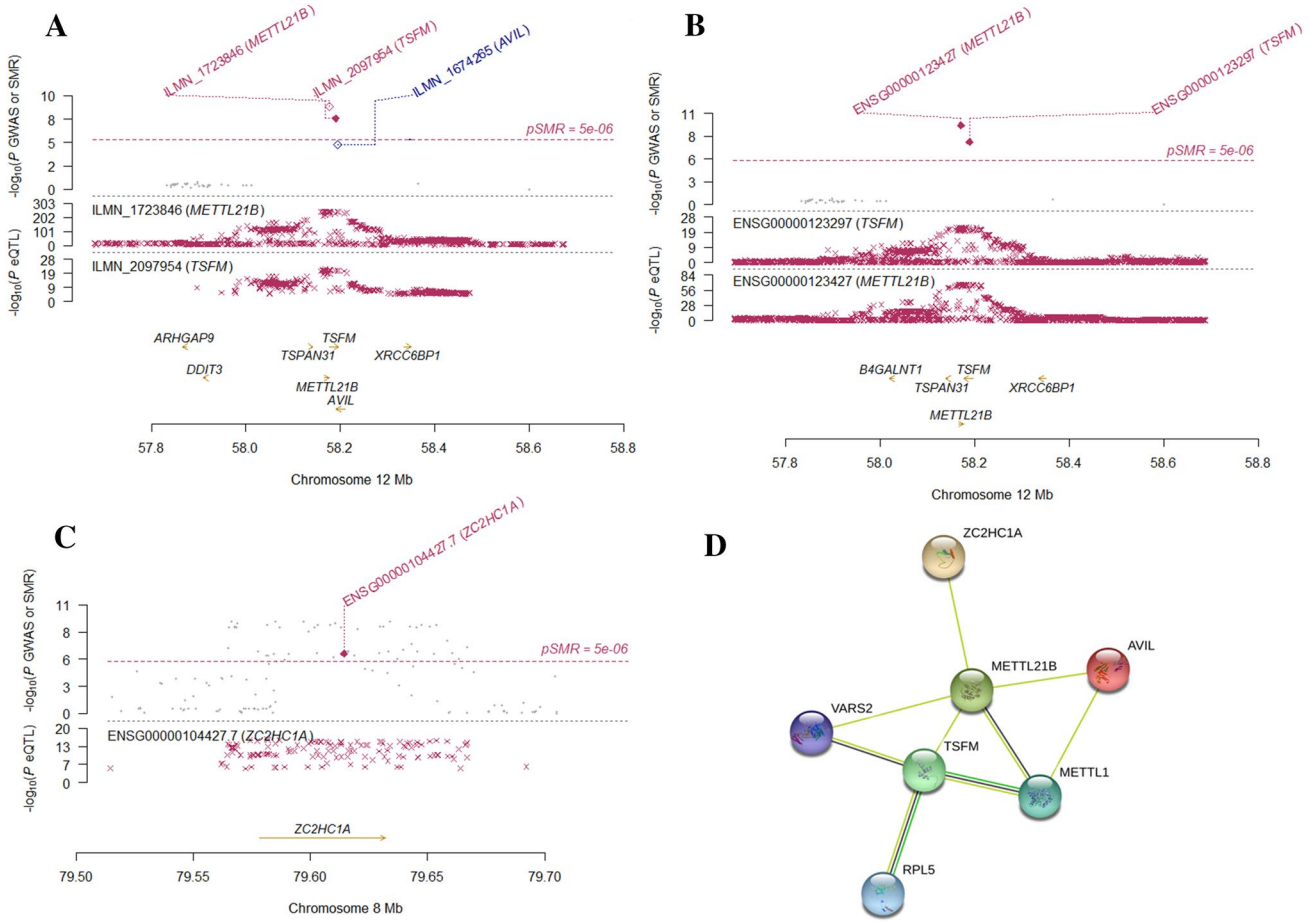

Fig. 1 The association between *METTL21B*, *TSFM*, *ZC2HC1A* and MS. The gene expression levels of *METTL21B* and *TSFM* were significantly associated with MS according to the SMR analysis based on eQTL data from the CAGE study (**a**), GTEx brain tissue (**b**). The gene expression level of *ZC2HC1A* was significantly associated with MS according to the SMR analysis based on eQTL data from GTEx whole blood (**c**). These genes were connected according to protein–protein interaction analysis (**d**)

MS, the power is very limited due to small sample size. Moreover, traditional observational studies are subject to confounding and reverse causation and have not been able to disentangle which genetic, epigenetic and other factors directly influence MS. MR-based studies can circumvent these limitations by use of genetic proxies of putative risk factors when evaluating their associations with disease risk, as they are not subject to reverse causation (i.e., MS causes mutations) [29, 30]. More importantly, the SMR method allows the evaluation of the association between methylation and expression levels and MS risk in very large samples thus the statistical power is increased. As we showed in this study, hundreds of DNA methylations and genes were detected to be important risk factors for MS, after considering multiple testing using the Bonferroni method.

DNA methylation, an important epigenetic modification, plays an important role in MS pathology and could be potential therapeutic targets. Our study identified plenty of DNA methylations and methylation related genes (e.g., *METTL21B* and *METTL1*) which were associated with MS. According to the UCSC database (https://genome.ucsc.edu/) we have noticed that there are CpG islands overlapped promoters in most of the identified gene regions and the promoters show long-range spatial interactions with other promoters or enhancers, which implied a regulatory potential of DNA methylation on the gene expression. The findings raised the possibility that DNA methylation affected gene expression and then caused MS. For example, in the *DDR1* gene region we found 13 methylation sites to be causally associated with MS, which have not been reported. All of these methylation sites were much closed to the CpG island 46 (chr6:30,852,103–30,852,676). This CpG island locates in a promoter (chr6:30,847,554–30,856,766). This CpG island is in long-range interaction with another nearby promoter (chr6:30,842,868–30,846,325) that is in long-range interaction with a nearby enhancer (chr6:30,857,116–30,858,670). Studies have shown that DNA methylations in long-range interactive promoters and enhancers can affect the

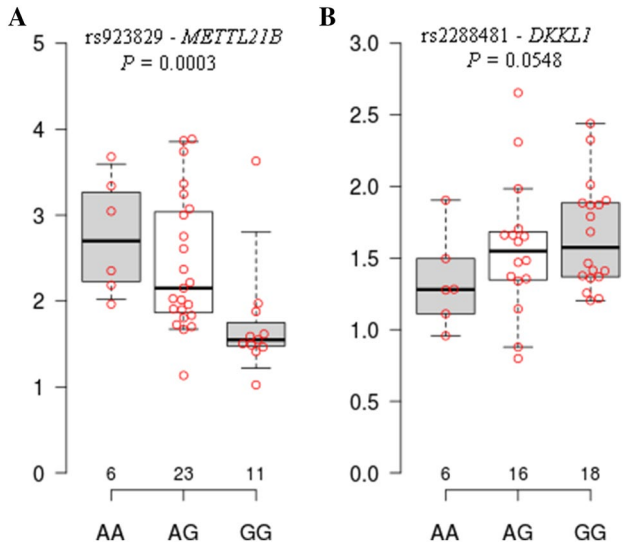

Fig. 2 The association between m^{6}A-SNPs and gene expression. We validated the association between rs923829 and rs2288481 and gene expressions of *METTL21B* (**a**) and *DKKL1* (**b**) in PBMCs of 40 individuals from the Chinese Han population, respectively

long-range interactions and may be associated with disease risk [31–33]. Therefore, these methylations were supposed to affect *DDR1* expression then affect MS risk. As we have found that *DDR1* was differentially expressed between MS cases and controls and *DDR1* expression level was causally associated with MS which was identified by SMR analysis.

This study showed the advantage of integration of multi-omics data in identification of regulatory mechanisms underlying the genetic associations. The *DDR1* gene encodes the Discoidin Domain Receptor Tyrosine Kinase 1. The importance of this gene in MS has not been discussed. We showed that SNPs in *DDR1* was strongly associated with MS in the GWAS data, while these SNPs were strongly associated with plasma MICB and Granzyme A protein levels. Moreover, we demonstrated that plasma MICB and Granzyme A protein levels may be causally associated with MS. We found that plasma MICB was associated with an increased risk of MS, which was consistent with previous case control study [34]. The relationship between Granzyme A protein levels and MS is first reported in the present study. How *DDR1* interacts with *MICB* and *GZMA* is unknown. As suggested by the protein–protein interaction analysis, these genes might interact with each other via other MS-related genes such as *PLCG1*, *ZAP70*, *KLRK1*, *TTR*, *SYK*, *RGS2*, *XRCC6* and *APEX1*. The interaction between these genes may point to a pathway for MS. Further studies are suggested to elucidate the mechanisms.

This study also showed the advantage in interpretation of the relationship among genes in GWAS-identified loci. A valuable discovery in this study was the interaction among *METTL21B*, *METTL1* and *TSFM* gene in 12q14.1. We showed that genetic variants and expression levels of these non-MHC genes were associated with MS. However, how this interaction affect MS is still unclear. Genetic variants (e.g., rs703842) in *METTL21B*, also named *FAM119B* and *EEF1AKMT3* (EEF1A Lysine Methyltransferase 3), have been showed to affect *METTL21B* expression [35]. In our study, we confirmed that the m^{6}A-SNP rs923829 was significantly associated with *METTL21B* expression levels, which has also been showed in other eQTL studies. So it is possible that genetic variants may affect *METTL21B* gene expression and then caused MS. The role of *METTL1* (Methyltransferase Like 1) and *TSFM* (Ts translation elongation factor, mitochondrial) in MS etiology is unknown but only the genetic association was reported. *METTL21B* has been showed to affect mRNA translation through inducible and dynamic methylation of Lys-165 in human eukaryotic elongation factor 1 alpha (eEF1A) [36, 37]. So it is reasonable to propose the hypothesis that the methyltransferases (*METTL21B* and *METTL1*) affect translation elongation factor (*TSFM*). And this interaction may affect MS.

In summary, we applied the SMR method in this study to provide robust evidence on the important roles of DNA methylations and expressions of many genes as important risk factors of MS. This study provided novel evidence of an important role of circulating MICB and Granzyme A in MS etiology. We also show that the interaction among *DDR1*, *MICB* and *GZMA* and interaction among *METTL21B*, *METTL1* and *TSFM* may participate in the pathogenesis of MS. Further researches are needed to fully elucidate the important relationship between these factors and MS.

Acknowledgments The study was supported by Natural Science Foundation of China (81773508), the Key Research Project (Social Development Plan) of Jiangsu Province (BE2016667), the Startup Fund from Soochow University (Q413900313, Q413900412), Project funded by China Postdoctoral Science Foundation (2013M530269 and 2014M551649, 2014T70547), and a Project of the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Compliance with ethical standards

Conflicts of interest The authors have declared no conflicts of interest.

Ethical standard All aspects of the study were performed in accordance with the Declaration of Helsinki, and the study protocol was approved by the local Ethics Board of Soochow university.

Informed consent Written informed consent was obtained from every participant.

References

1. Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kemppinen A, Cotsapas C, Shah TS, Spencer C, Booth D, Goris A, Oturai A,

Saarela J, Fontaine B, Hemmer B, Martin C, Zipp F, D'Alfonso S, Martinelli-Boneschi F, Taylor B, Harbo HF, Kockum I, Hillert J, Olsson T, Ban M, Oksenberg JR, Hintzen R, Barcellos LF, Agliardi C, Alfredsson L, Alizadeh M, Anderson C, Andrews R, Sondergaard HB, Baker A, Band G, Baranzini SE, Barizzone N, Barrett J, Bellenguez C, Bergamaschi L, Bernardinelli L, Berthele A, Biberacher V, Binder TM, Blackburn H, Bomfim IL, Brambilla P, Broadley S, Brochet B, Brundin L, Buck D, Butzkueven H, Caillier SJ, Camu W, Carpentier W, Cavalla P, Celius EG, Coman I, Comi G, Corrado L, Cosemans L, Cournu-Rebeix I, Cree BA, Cusi D, Damotte V, Defer G, Delgado SR, Deloukas P, di Sapio A, Dilthey AT, Donnelly P, Dubois B, Duddy M, Edkins S, Elovaara I, Esposito F, Evangelou N, Fiddes B, Field J, Franke A, Freeman C, Frohlich IY, Galimberti D, Gieger C, Gourraud PA, Graetz C, Graham A, Grummel V, Guaschino C, Hadjixenofontos A, Hakonarson H, Halfpenny C, Hall G, Hall P, Hamsten A, Harley J, Harrower T, Hawkins C, Hellenthal G, Hillier C, Hobart J, Hoshi M, Hunt SE, Jagodic M, Jelcic I, Jochim A, Kendall B, Kermode A, Kilpatrick T, Koivisto K, Konidari I, Korn T, Kronsbein H, Langford C, Larsson M, Lathrop M, Lebrun-Frenay C, Lechner-Scott J, Lee MH, Leone MA, Leppa V, Liberatore G, Lie BA, Lill CM, Linden M, Link J, Luessi F, Lycke J, Macciardi F, Mannisto S, Manrique CP, Martin R, Martinelli V, Mason D, Mazibrada G, McCabe C, Mero IL, Mescheriakova J, Moutsianas L, Myhr KM, Nagels G, Nicholas R, Nilsson P, Piehl F, Pirinen M, Price SE, Quach H, Reunanen M, Robberecht W, Robertson NP, Rodegher M, Rog D, Salvetti M, Schnetz-Boutaud NC, Sellebjerg F, Selter RC, Schaefer C, Shaunak S, Shen L, Shields S, Siffrin V, Slee M, Sorensen PS, Sorosina M, Sospedra M, Spurkland A, Strange A, Sundqvist E, Thijs V, Thorpe J, Ticca A, Tienari P, van Duijn C, Visser EM, Vucic S, Westerlind H, Wiley JS, Wilkins A, Wilson JF, Winkelmann J, Zajicek J, Zindler E, Haines JL, Pericak-Vance MA, Ivinson AJ, Stewart G, Hafler D, Hauser SL, Compston A, McVean G, De Jager P, Sawcer SJ, McCauley JL (2013) Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet 45:1353–1360

2. Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, de Bakker PI, Gabriel SB, Mirel DB, Ivinson AJ, Pericak-Vance MA, Gregory SG, Rioux JD, McCauley JL, Haines JL, Barcellos LF, Cree B, Oksenberg JR, Hauser SL (2007) Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 357:851–862
3. De Jager PL, Jia X, Wang J, de Bakker PI, Ottoboni L, Aggarwal NT, Piccio L, Raychaudhuri S, Tran D, Aubin C, Briskin R, Romano S, Baranzini SE, McCauley JL, Pericak-Vance MA, Haines JL, Gibson RA, Naeglin Y, Uitdehaag B, Matthews PM, Kappos L, Polman C, McArdle WL, Strachan DP, Evans D, Cross AH, Daly MJ, Compston A, Sawcer SJ, Weiner HL, Hauser SL, Hafler DA, Oksenberg JR (2009) Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. Nat Genet 41:776–782
4. Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, Moutsianas L, Dilthey A, Su Z, Freeman C, Hunt SE, Edkins S, Gray E, Booth DR, Potter SC, Goris A, Band G, Oturai AB, Strange A, Saarela J, Bellenguez C, Fontaine B, Gillman M, Hemmer B, Gwilliam R, Zipp F, Jayakumar A, Martin R, Leslie S, Hawkins S, Giannoulatou E, D'Alfonso S, Blackburn H, Martinelli Boneschi F, Liddle J, Harbo HF, Perez ML, Spurkland A, Waller MJ, Mycko MP, Ricketts M, Comabella M, Hammond N, Kockum I, McCann OT, Ban M, Whittaker P, Kemppinen A, Weston P, Hawkins C, Widaa S, Zajicek J, Dronov S, Robertson N, Bumpstead SJ, Barcellos LF, Ravindrarajah R, Abraham R, Alfredsson L, Ardlie K, Aubin C, Baker A, Baker K, Baranzini SE, Bergamaschi L, Bergamaschi R, Bernstein A, Berthele A, Boggild M, Bradfield JP, Brassat D, Broadley SA, Buck D, Butzkueven H, Capra R, Carroll WM, Cavalla P, Celius EG, Cepok S, Chiavacci R, Clerget-Darpoux F, Clysters K, Comi G, Cossburn M, Cournu-Rebeix I, Cox MB, Cozen W, Cree BA, Cross AH, Cusi D, Daly MJ, Davis E, de Bakker PI, Debouverie M, D'Hooghe MB, Dixon K, Dobosi R, Dubois B, Ellinghaus D, Elovaara I, Esposito F, Fontenille C, Foote S, Franke A, Galimberti D, Ghezzi A, Glessner J, Gomez R, Gout O, Graham C, Grant SF, Guerini FR, Hakonarson H, Hall P, Hamsten A, Hartung HP, Heard RN, Heath S, Hobart J, Hoshi M, Infante-Duarte C, Ingram G, Ingram W, Islam T, Jagodic M, Kabesch M, Kermode AG, Kilpatrick TJ, Kim C, Klopp N, Koivisto K, Larsson M, Lathrop M, Lechner-Scott JS, Leone MA, Leppa V, Liljedahl U, Bomfim IL, Lincoln RR, Link J, Liu J, Lorentzen AR, Lupoli S, Macciardi F, Mack T, Marriott M, Martinelli V, Mason D, McCauley JL, Mentch F, Mero IL, Mihalova T, Montalban X, Mottershead J, Myhr KM, Naldi P, Ollier W, Page A, Palotie A, Pelletier J, Piccio L, Pickersgill T, Piehl F, Pobywajlo S, Quach HL, Ramsay PP, Reunanen M, Reynolds R, Rioux JD, Rodegher M, Roesner S, Rubio JP, Ruckert IM, Salvetti M, Salvi E, Santaniello A, Schaefer CA, Schreiber S, Schulze C, Scott RJ, Sellebjerg F, Selmaj KW, Sexton D, Shen L, Simms-Acuna B, Skidmore S, Sleiman PM, Smestad C, Sorensen PS, Sondergaard HB, Stankovich J, Strange RC, Sulonen AM, Sundqvist E, Syvanen AC, Taddeo F, Taylor B, Blackwell JM, Tienari P, Bramon E, Tourbah A, Brown MA, Tronczynska E, Casas JP, Tubridy N, Corvin A, Vickery J, Jankowski J, Villoslada P, Markus HS, Wang K, Mathew CG, Wason J, Palmer CN, Wichmann HE, Plomin R, Willoughby E, Rautanen A, Winkelmann J, Wittig M, Trembath RC, Yaouanq J, Viswanathan AC, Zhang H, Wood NW, Zuvich R, Deloukas P, Langford C, Duncanson A, Oksenberg JR, Pericak-Vance MA, Haines JL, Olsson T, Hillert J, Ivinson AJ, De Jager PL, Peltonen L, Stewart GJ, Hafler DA, Hauser SL, McVean G, Donnelly P, Compston A (2011) Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 476:214–219
5. Patsopoulos NA, Esposito F, Reischl J, Lehr S, Bauer D, Heubach J, Sandbrink R, Pohl C, Edan G, Kappos L, Miller D, Montalban J, Polman CH, Freedman MS, Hartung HP, Arnason BG, Comi G, Cook S, Filippi M, Goodin DS, Jeffery D, O'Connor P, Ebers GC, Langdon D, Reder AT, Traboulsee A, Zipp F, Schimrigk S, Hillert J, Bahlo M, Booth DR, Broadley S, Brown MA, Browning BL, Browning SR, Butzkueven H, Carroll WM, Chapman C, Foote SJ, Griffiths L, Kermode AG, Kilpatrick TJ, Lechner-Scott J, Marriott M, Mason D, Moscato P, Heard RN, Pender MP, Perreau VM, Perera D, Rubio JP, Scott RJ, Slee M, Stankovich J, Stewart GJ, Taylor BV, Tubridy N, Willoughby E, Wiley J, Matthews P, Boneschi FM, Compston A, Haines J, Hauser SL, McCauley J, Ivinson A, Oksenberg JR, Pericak-Vance M, Sawcer SJ, De Jager PL, Hafler DA, de Bakker PI (2011) Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. Ann Neurol 70:897–912
6. Visscher PM, Brown MA, McCarthy MI, Yang J (2012) Five years of GWAS discovery. Am J Hum Genet 90:7–24
7. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, Jansen R, de Geus EJ, Boomsma DI, Wright FA, Sullivan PF, Nikkola E, Alvarez M, Civelek M, Lusis AJ, Lehtimaki T, Raitoharju E, Kahonen M, Seppala I, Raitakari OT, Kuusisto J, Laakso M, Price AL, Pajukanta P, Pasaniuc B (2016) Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet 48:245–252
8. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, Eyler AE, Denny JC, Nicolae DL, Cox NJ, Im HK (2015) A gene-based association method for mapping traits using reference transcriptome data. Nat Genet 47:1091–1098
9. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, Montgomery GW, Goddard ME, Wray NR, Visscher PM, Yang J (2016) Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 48:481–487

10. Pavlides JM, Zhu Z, Gratten J, McRae AF, Wray NR, 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
11. Suhre K, Arnold M, Bhagwat AM, Cotton RJ, Engelke R, Raffler J, Sarwath H, Thareja G, Wahl A, DeLisle RK, Gold L, Pezer M, Lauc G, El-Din Selim MA, Mook-Kanamori DO, Al-Dous EK, Mohamoud YA, Malek J, Strauch K, Grallert H, Peters A, Kastenmuller G, Gieger C, Graumann J (2017) Connecting genetic risk to disease end points through the human blood plasma proteome. Nat Commun 8:14357
12. Long T, Hicks M, Yu HC, Biggs WH, Kirkness EF, Menni C, Zierer J, Small KS, Mangino M, Messier H, Brewerton S, Turpaz Y, Perkins BA, Evans AM, Miller LA, Guo L, Caskey CT, Schork NJ, Garner C, Spector TD, Venter JC, Telenti A (2017) Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. Nat Genet 49:568–578
13. Kerekes G, Nurmohamed MT, Gonzalez-Gay MA, Seres I, Paragh G, Kardos Z, Barath Z, Tamasi L, Soltesz P, Szekanecz Z (2014) Rheumatoid arthritis and metabolic syndrome. Nat Rev Rheumatol 10:691–696
14. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, Kochi Y, Ohmura K, Suzuki A, Yoshida S, Graham RR, Manoharan A, Ortmann W, Bhangale T, Denny JC, Carroll RJ, Eyler AE, Greenberg JD, Kremer JM, Pappas DA, Jiang L, Yin J, Ye L, Su DF, Yang J, Xie G, Keystone E, Westra HJ, Esko T, Metspalu A, Zhou X, Gupta N, Mirel D, Stahl EA, Diogo D, Cui J, Liao K, Guo MH, Myouzen K, Kawaguchi T, Coenen MJ, van Riel PL, van de Laar MA, Guchelaar HJ, Huizinga TW, Dieude P, Mariette X, Bridges SL Jr, Zhernakova A, Toes RE, Tak PP, Miceli-Richard C, Bang SY, Lee HS, Martin J, Gonzalez-Gay MA, Rodriguez-Rodriguez L, Rantapaa-Dahlqvist S, Arlestig L, Choi HK, Kamatani Y, Galan P, Lathrop M, Eyre S, Bowes J, Barton A, de Vries N, Moreland LW, Criswell LA, Karlson EW, Taniguchi A, Yamada R, Kubo M, Liu JS, Bae SC, Worthington J, Padyukov L, Klareskog L, Gregersen PK, Raychaudhuri S, Stranger BE, De Jager PL, Franke L, Visscher PM, Brown MA, Yamanaka H, Mimori T, Takahashi A, Xu H, Behrens TW, Siminovitch KA, Momohara S, Matsuda F, Yamamoto K, Plenge RM (2014) Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature 506:376–381
15. Battle A, Brown CD, Engelhardt BE, Montgomery SB (2017) Genetic effects on gene expression across human tissues. Nature 550:204–213
16. McRae AF, Marioni RE, Shah S, Yang J, Powell JE, Harris SE, Gibson J, Henders AK, Bowdler L, Painter JN, Murphy L, Martin NG, Starr JM, Wray NR, Deary IJ, Visscher PM, Montgomery GW (2018) Identification of 55,000 Replicated DNA Methylation QTL. Sci Rep 8:17605
17. Kemppinen AK, Kaprio J, Palotie A, Saarela J (2011) Systematic review of genome-wide expression studies in multiple sclerosis. BMJ Open 1:e000053
18. Zhang F, Shi Y, Wang L, Sriram S (2011) Role of HDAC3 on p53 expression and apoptosis in T cells of patients with multiple sclerosis. PLoS ONE 6:e16795
19. Nakatsuji Y, Okuno T, Moriya M, Sugimoto T, Kinoshita M, Takamatsu H, Nojima S, Kimura T, Kang S, Ito D, Nakagawa Y, Toyofuku T, Takata K, Nakano M, Kubo M, Suzuki S, Matsui-Hasumi A, Uto-Konomi A, Ogata A, Mochizuki H, Sakoda S, Kumanogoh A (2012) Elevation of Sema4A implicates Th cell skewing and the efficacy of IFN-beta therapy in multiple sclerosis. J Immunol 188:4858–4865
20. Zheng Y, Nie P, Peng D, He Z, Liu M, Xie Y, Miao Y, Zuo Z, Ren J (2018) m6AVar: a database of functional variants involved in m6A modification. Nucleic Acids Res 46:D139–D145
21. Ren J, Jiang C, Gao X, Liu Z, Yuan Z, Jin C, Wen L, Zhang Z, Xue Y, Yao X (2010) PhosSNP for systematic analysis of genetic polymorphisms that influence protein phosphorylation. Mol Cell Proteom 9:623–634
22. Ward LD, Kellis M (2012) HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res 40:D930–934
23. Zhu Z, Zheng Z, Zhang F, Wu Y, Trzaskowski M, Maier R, Robinson MR, McGrath JJ, Visscher PM, Wray NR, Yang J (2018) Causal associations between risk factors and common diseases inferred from GWAS summary data. Nat Commun 9:224
24. Verbanck M, Chen CY, Neale B, Do R (2018) Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 50:693–698
25. Li X, Xiao B, Chen XS (2017) DNA methylation: a new player in multiple sclerosis. Mol Neurobiol 54:4049–4059
26. Casaccia-Bonnefil P, Pandozy G, Mastronardi F (2008) Evaluating epigenetic landmarks in the brain of multiple sclerosis patients: a contribution to the current debate on disease pathogenesis. Prog Neurobiol 86:368–378
27. Huynh JL, Casaccia P (2013) Epigenetic mechanisms in multiple sclerosis: implications for pathogenesis and treatment. Lancet Neurol 12:195–206
28. Annibali V, Ristori G, Angelini DF, Serafini B, Mechelli R, Cannoni S, Romano S, Paolillo A, Abderrahim H, Diamantini A, Borsellino G, Aloisi F, Battistini L, Salvetti M (2011) CD161(high)CD8 + T cells bear pathogenetic potential in multiple sclerosis. Brain 134:542–554
29. Davey Smith G, Hemani G (2014) Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet 23:R89–98
30. Smith GD, Ebrahim S (2003) 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol 32:1–22
31. Park SM, Choi EY, Bae M, Choi JK, Kim YJ (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
32. Flavahan WA, Drier Y, Liau BB, Gillespie SM, Venteicher AS, Stemmer-Rachamimov AO, Suva ML, Bernstein BE (2016) Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature 529:110–114
33. Schoenfelder S, Fraser P (2019) Long-range enhancer-promoter contacts in gene expression control. Nat Rev Genet. https://doi.org/10.1038/s41576-41019-40128-41570
34. Fernandez-Morera JL, Rodriguez-Rodero S, Lahoz C, Tunon A, Astudillo A, Garcia-Suarez O, Martinez-Borra J, Lopez-Vazquez A, Rodrigo L, Gonzalez S, Lopez-Larrea C (2008) Soluble MHC class I chain-related protein B serum levels correlate with disease activity in relapsing-remitting multiple sclerosis. Hum Immunol 69:235–240
35. Gandhi KS, McKay FC, Cox M, Riveros C, Armstrong N, Heard RN, Vucic S, Williams DW, Stankovich J, Brown M, Danoy P, Stewart GJ, Broadley S, Moscato P, Lechner-Scott J, Scott RJ, Booth DR (2010) The multiple sclerosis whole blood mRNA transcriptome and genetic associations indicate dysregulation of specific T cell pathways in pathogenesis. Hum Mol Genet 19:2134–2143
36. Malecki J, Aileni VK, Ho AYY, Schwarz J, Moen A, Sorensen V, Nilges BS, Jakobsson ME, Leidel SA, Falnes PO (2017) The novel lysine specific methyltransferase METTL21B affects mRNA translation through inducible and dynamic methylation of Lys-165 in human eukaryotic elongation factor 1 alpha (eEF1A). Nucleic Acids Res 45:4370–4389
37. Hamey JJ, Wienert B, Quinlan KGR, Wilkins MR (2017) METTL21B is a novel human lysine methyltransferase of translation elongation factor 1A: discovery by CRISPR/Cas9 Knockout. Mol Cell Proteom 16:2229–2242