PNAS
www.pnas.org

Supplementary Information for

Specific hypomethylation programs underpin B cell activation in early multiple sclerosis

Qin Ma[1*], Stacy J. Caillier[1], Shaun Muzic[1], University of California San Francisco MS-EPIC Team[1], Michael R. Wilson[1], Roland G. Henry[1], Bruce A.C. Cree[1], Stephen L. Hauser[1], Alessandro Didonna[1], Jorge R. Oksenberg[1*]

[1]Weill Institute for Neurosciences, Department of Neurology, University of California San Francisco, San Francisco, CA 94158, USA

***Corresponding Author:** Qin Ma, Jorge R. Oksenberg
***Email:** qin.ma@ucsf.edu, jorge.oksenberg@ucsf.edu

This PDF file includes:

SI Material and Methods
SI References
Figures S1 to S8
Legends for Datasets S1 to S14
Members of the University of California San Francisco MS-EPIC Team

Other supplementary materials for this manuscript include the following:

Datasets S1 to S14

SI Materials and Methods

Subjects

Peripheral blood was sampled by venipuncture from 29 multiple sclerosis (MS) patients and 24 healthy controls recruited at the University of California San Francisco (UCSF) MS Center as part of the UCSF multiple sclerosis EPIC and ORIGINS studies. All patients were treatment naive (i.e., no glucocorticoids or disease-modifying MS therapy at any time prior to blood draw) and met clinical and radiographic diagnostic criteria for relapsing-remitting MS (RRMS). Individuals selected were all of white-western European ancestry, female, with no history of smoking (**Dataset S3**). Mean age at sample collection was 38 (range from 21 to 56), mean disease duration was 1.76 +/- 2.8 years and mean Expanded Disability Status Scale (EDSS) was 2.15 +/- 0.74. Healthy controls were of white-western European ancestry, sex and age matched (mean age: 37, range: 22-58), with no history of autoimmune disease or smoking. The individuals are further divided into two datasets, the first cohort (17 MS patients and 12 healthy controls) was used for the analysis of $CD4^+$, $CD8^+$, $CD14^+$ and $CD19^+$ cells, whereas the second dataset was only used for the analysis of $CD19^+$ cells. There was no significant difference in age between MS and control samples, and between two datasets (Wilcoxon test, P>0.1). There was no significant difference in EDSS at time of sample collection between two datasets (Wilcoxon test, P=0.77). On the other hand, 23.53% of patients in the discovery cohort displayed more than 20 brain lesions detected by the baseline MRI, whereas only 9.09% of patients in the second cohort had more than 20 brain lesions (***SI Appendix,* Fig. S2A**). All studies were approved by the UCSF Institutional Review Board, and informed consent was obtained from all subjects.

Sample preparation

Peripheral blood mononuclear cells (PBMCs) were isolated immediately after blood collection using a standard Ficoll (GE Healthcare) gradient procedure and cells were frozen down in freezing media made of fetal bovine serum (FBS) and 10% DMSO. PBMC aliquots were stored in liquid nitrogen (LN2, gas phase) for further use. At time of experimentation, PBMCs were removed from LN2 and thawed in a 37ºC water bath. Four MS patients and four healthy controls were processed at each time. Cells were washed and incubated in warm media (RPMI/10%FBS) with benzonase for 5 minutes then washed with media and rested in 37C/5% CO_2 incubator for 1 hour. Cells were then washed with warm phosphate buffered saline (PBS), blocked with FcR Block (Miltenyi Biotec), and incubated in the dark on ice for 30 minutes with the following antibodies: CD3 APC (UCHT1, Beckman Coulter), CD4 Brilliant Violet 421(OKT4, BioLegend), CD8 APC AlexaFluor750 (B9.11, Beckman Coulter), CD19 PerCP Cy5.5 (J3-119, Beckman Coulter), CD14 PE (61D3, eBioscience). Cells were washed and resuspended in PBS containing 1% FBS and then $CD4^+$ and $CD8^+$ T cells, $CD14^+$ monocytes, and $CD19^+$ B cells were sorted on a MoFlo Astrios (Beckman Coulter). Cell subsets were then pelleted, flash frozen on dry ice and stored at -80ºC for DNA or RNA extraction.

Genomic DNA extraction and BS-seq library preparation

Genomic DNA from 100,000-500,000 cells was extracted using the Quick DNA kit (Zymo Research) and processed for BS-seq library preparations as described (1). For library preparations, we processed each time 16-22 samples from one cell type across cases and controls following the same protocol. Briefly, genomic DNA was fragmented, end-repaired, A-tailed, and ligated to methylated adaptors. Ligated fragments were then bisulfite converted using the EZ DNA Methylation Gold kit (Zymo Research) at 64℃ for 2.5 hours and amplified using the KAPA HiFi HotStart Uracil+ ReadyMix PCR Kit. Libraries were sequenced on a NovaSeq platform (Illumina). The sequencing reads were analyzed following an established bioinformatics pipeline (1, 2). Briefly, after filtering out low quality reads, the remaining paired reads were uniquely mapped to the reference genome (hg38, UCSC) using the Bismark software (3). The bisulfite conversion rates were calculated using the lambda DNA. Subsequently, the 5mC level at each CpG site was computed as described (1). The number of “C” bases from the sequencing reads

were counted as methylated (denoted as NC) and the number of "T" bases as unmodified (denoted as NT). The methylation levels were then estimated as NC/ (NC + NT).

BS-seq data analysis

First, the metilene package (version 0.2-8) was used to call the differentially methylated regions (DMRs), including only CpGs with at least 3× in coverage (4). DMRs were defined to have a 10% minimum absolute mean methylation difference, as recommended by multiple DMR detection tools (4-6), with a maximum distance of 1000 nt between CpGs within a DMR and a minimum of 3 CpGs per DMR (parameter –M 1000 –m 3 –d 0.1). DMRs were first calculated in $CD4^+$ and $CD8^+$T cells (n=29), $CD14^+$ monocytes (n=22) and $CD19^+$ B cells (n=29) respectively, between RRMS and HC individuals. Subsequently, the DMR analysis in B cells were extended to 29 RRMS and 23 HC individuals. After each DMR calling, the regions with false discovery rate (FDR) of 0.05 or less were considered as final DMRs. For heatmap visualization, the mean methylation levels for the different DMRs were calculated and plotted using the heatmap function of the R package ComplexHeatmap (7). The rainfall track and genomic density track for the hypo-DMRs and hyper-DMRs were plotted using the circlize R package (8). PCA analysis of DMRs was performed using the ggfortify package in R and Bioconductor. The P values for the enrichments in hypo-DMR and hyper-DMRs at different genomic loci were calculated using the HOMER annotatePeaks.pl function. The Genomic Regions Enrichment of Annotations Tool (GREAT) was employed to predict the DMR functional significance (9).

RNA extraction and mRNA-seq

Transcriptomic analysis was performed on the B cell fractions from 23 MS patients and 14 healthy controls selected within our study cohort. Total RNA was extracted using Quick-RNA Microprep Kit (Zymo Research). The poly(A) selected paired-end sequencing libraries were constructed using the mRNA Hyper Prep Kit (KAPA), following the manufacturer's instructions. The sequencing was performed on HiSeq X Ten platform (Illumina), generating ~30 million pair-end 150 bp reads per sample. For the analysis, low-quality and adapter sequences were trimmed using the Trimmomatic tool (10). Sequencing reads were then aligned to the human reference (hg38) using hisat2 (11). The htseq-count tool and human gtf file from Ensembl were used to count aligned reads for each gene (12). Differentially expressed genes (DEGs) between RRMS and HC individuals were identified using the edgeR package (13). Genes were considered significant differentially expressed with an absolute fold change of 1.5 or more and FDR of 0.05 or less.

Publicly available B cell datasets

The B cell subset-specific datasets obtained by reduced-representation bisulfite sequencing (RRBS) (accession number: GSE118255) and RNA-seq (accession number: GSE118254) were downloaded from the GEO repository (14). The count matrix of methylated and unmethylated reads for each CpG covered by the RRBS data was used to calculate the mean methylation level for each DMR in each cell sample and then averaged across each cell type. The methylation differences between cell types were assessed by two-tailed Student's t-test. For the DEGs identified in our data, the kilobase per million reads (RPKM) values were searched in the RNA-seq data and then averaged across each cell type. The log2 normalized RPKM values for DEGs in each cell type were used to generate the heatmap and assess the expression differences between cell types by two-tailed Student's t-test. The B cell line (GM12878) transcription factor ChIP-seq dataset was downloaded from the ENCODE Consortium. The BEDTools intersectBed function was used to get the overlap between transcription factor binding peaks and hypo-DMRs. To test the statistical significance, the peaks were randomly permuted 10,000 times among human genome by BEDTools shuffleBed function, while retaining the length of each peak to assess the distribution of background overlap.

Overlap with MS susceptibility genes

The recently reported MS susceptibility protein coding genes were used to test the overlap between DMRs and MS-associated loci (15). A DMR was assigned to the nearest coding gene when the DMR was located within 100-kb flanking region of the transcriptional start site (TSS). Subsequently, based on the distance between the DMR and TSS, DMRs were assigned to the following categories: TSS-proximal (3 kb upstream and downstream of the TSS), TSS-distal 10 kb (10 kb upstream and downstream of the TSS), and TSS-distal 100 kb (100 kb upstream and downstream of the TSS). The statistical significance of the overlap between DMR-coupled coding genes in each category and MS susceptibility genes was assessed by Fisher's exact test using the GeneOverlap package (URL: http://shenlab-sinai.github.io/shenlab-sinai/). The DMRs were further divided into 10 groups according to the distance and the statistical significance in each group was tested. A linear regression model was built to find association between the distance to TSS and the overlap significance.

Cell lines

The HeLa cell line was obtained from ATCC biobank. The cells were cultured in Dulbecco's Modified Eagle's Medium (GIBCO) supplemented with 10% v/v fetal bovine serum (GIBCO) and maintained at 37°C in a humidified atmosphere with 5% CO2.

Plasmids

The CpG-free firefly luciferase reporter vectors pCpGL-basic and pCpGL-CMV were provided by Dr. Michael Rehli (University Hospital, Regensburg, Germany). The Renilla luciferase reporter vector pRL-TK was purchased from Promega. To study the effects of methylation on transcription, the promoter regions (1 kb upstream the TSS) of the genes *DCAF12*, *PRDM1*, *SEM4A*, *TNFRSF17*, and *ZWINT* were amplified from genomic DNA with Phusion High-Fidelity DNA Polymerase (New England Biolabs), using the primers listed in **Dataset S14**. The PCR products were then double-digested with the restriction enzymes SpeI and NcoI (New England Biolabs) and cloned into the pCpGL-basic vector, previously digested with the same pair of enzymes. Individual clones were confirmed by Sanger sequencing. In preparation for luciferase assays, the different pCpGL constructs were methylated in vitro using the DNA methyltransferase SssI (New England Biolabs). In each reaction, 6 µg of plasmid was incubated with 12 U of enzyme in the presence of 320 µM S-Adenosylmethionine (SAM; New England Biolabs) for 4 hours at 37°C, with additional 320 µM SAM supplemented after the first two hours of incubation. Mock reactions without the methyltransferases were carried out in parallel for the non-methylated vectors. All the plasmids were finally precipitated overnight with ethanol and resuspended in sterile water.

Luciferase assays

HeLa cells were grown in 12-well plates and co-transfected at 90% confluency with 700 ng of the different pCpGL constructs, either methylated or non-methylated, and 70 ng of pRL-TK vector using Lipofectamine 2000 (Invitrogen), according to the manufacturer's instructions. After 24 hours, cells washed once with PBS and were lysed in 250 mL of Passive Lysis Buffer (Promega), for 15 minutes at room temperature with agitation. Twenty mL of each cell lysate were subsequently loaded in duplicate in black 96-well plates and the activities of both firefly and Renilla luciferases were quantified by means of Dual-Luciferase Reporter Assay System (Promega), using a Veritas Microplate Luminometer (Turner BioSystems). The firefly luciferase activity of each construct was first normalized against Renilla luciferase activity, and then the normalized signals of the methylated constructs were expressed as fold-changes of the corresponding non-methylated vectors. Data from 3 biologically independent experiments were expressed as mean ± SEM. Differences between methylated and non-methylated vectors were assessed with two-tailed Student's t test, and P values equal to 0.05 or less were considered significant.

Enzyme-linked immunosorbent assays

Serum samples were processed immediately after phlebotomy and stored at −80°C. Commercially available enzyme-linked immunosorbent assay (ELISA) kits used for measuring IgM, IgG and cytokines levels in the serum were as follows: IgM ELISA kit from Invitrogen (catalog: 88-50620-22); IgG ELISA kit from Invitrogen (catalog: 88-50550-88); BAFF ELISA kit from R&D Systems (catalog: DY124-05), TRAIL ELISA kit from R&D Systems (catalog: DTRL00), IL6 ELISA kit from Invitrogen (catalog: BMS213HS); IL10 ELISA kit from Invitrogen (catalog: BMS215HS). A total of 46 samples (18 HC and 28 MS) were tested. Serum IgM and IgG levels were assessed in triplicates. Serum TRAIL, BAFF, IL6 and IL10 levels were assessed in duplicates. Differences between MS and HC were assessed with Mann–Whitney U test, and P values equal to 0.05 or less were considered significant. To assess the associations between DNA methylation level and IgM level, the patients were divided in three groups according to IgM levels. The statistical significance of methylation differences between groups was calculated by the two-tailed Student's t-test.

Lymphocyte immunophenotyping

The antibody mix consisted of CD19 (PerCPCy5.5, Beckman Coulter), CD27 (PE, BD Biosciences), IgD (FITC, Beckman Coulter), CD38 (Pacific Blue, Beckman Coulter), CD40 (APC, BD Biosciences), CD80 (PECy7, BD Biosciences), and CD86 (PE-CF594, BD Biosciences) for all samples. B cell subtypes were defined with the following surface markers: Naïve B cells (CD19+CD27-IgD+), unswitched (CD19+CD27+IgD+) and switched memory B cells (CD19+CD27+IgD-), double-negative B cells (CD19+CD27-IgD-) and plasma cells (CD19+IgD-CD27hiCD38hi). Within the CD19+ B cell population, cell activation was assessed by measuring the surface expression of the activation markers CD40, CD80 and CD86 by flow cytometry. Surface expression was analyzed using a LSRFortessa instrument equipped with FACSDiva software (BD Biosciences).

Bioinformatics analyses

The DMRs and every 500-bp window in ±2 kb flanking DMRs were used to search for the enrichment of known motifs using the HOMER tool (16). The default parameters with a fragment size for motif searching of 200 bp were used. To look at the association between DNA methylation changes and gene expression changes, DMRs were annotated to the nearest TSS using in-house Perl script and Ensembl gtf file. The statistical significance of the overlap between hypo-DMR or hyper-DMR coupled coding genes and up- or down-regulated DEGs was assessed by Fisher's exact test using the GeneOverlap package, respectively. The differences on cytokines, IgM and IgG, cell surface expression levels between two groups were assessed with the Mann–Whitney U test. The differences on the proportion of each B cell subset between two groups were assessed with the Mann–Whitney U test. Gene ontology (GO) analysis was performed using Metascape, providing gene annotation and enrichment analysis including GO processes and KEGG pathways (17). The known protein-protein interaction (PPI) networks were retrieved from the STRING database using Cytoscape v3.7.2 and stringApp v1.5.1 (18-20).

SI References

1. K. Chen *et al.*, Loss of 5-hydroxymethylcytosine is linked to gene body hypermethylation in kidney cancer. *Cell Res* **26**, 103-118 (2016).
2. Q. Ma *et al.*, Distal regulatory elements identified by methylation and hydroxymethylation haplotype blocks from mouse brain. *Epigenetics Chromatin* **11**, 75 (2018).
3. F. Krueger, S. R. Andrews, Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571-1572 (2011).
4. F. Juhling *et al.*, metilene: fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome Res* **26**, 256-262 (2016).
5. K. D. Hansen, B. Langmead, R. A. Irizarry, BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. *Genome Biology* **13**, R83 (2012).
6. D. E. Condon *et al.*, Defiant: (DMRs: easy, fast, identification and ANnoTation) identifies differentially Methylated regions from iron-deficient rat hippocampus. *BMC Bioinformatics* **19**, 31 (2018).
7. Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847-2849 (2016).
8. Z. Gu, L. Gu, R. Eils, M. Schlesner, B. Brors, circlize Implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811-2812 (2014).
9. C. Y. McLean *et al.*, GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**, 495-501 (2010).
10. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
11. D. Kim, J. M. Paggi, C. Park, C. Bennett, S. L. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907-915 (2019).
12. S. Anders, P. T. Pyl, W. Huber, HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169 (2015).
13. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).
14. C. D. Scharer *et al.*, Epigenetic programming underpins B cell dysfunction in human SLE. *Nat Immunol* **20**, 1071-1082 (2019).
15. International Multiple Sclerosis Genetics Consortium, Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science* **365** (2019).
16. S. Heinz *et al.*, Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589 (2010).
17. Y. Zhou *et al.*, Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523 (2019).
18. N. T. Doncheva, J. H. Morris, J. Gorodkin, L. J. Jensen, Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J Proteome Res* **18**, 623-632 (2019).
19. D. Szklarczyk *et al.*, The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* **45**, D362-D368 (2017).
20. M. S. Cline *et al.*, Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc* **2**, 2366-2382 (2007).

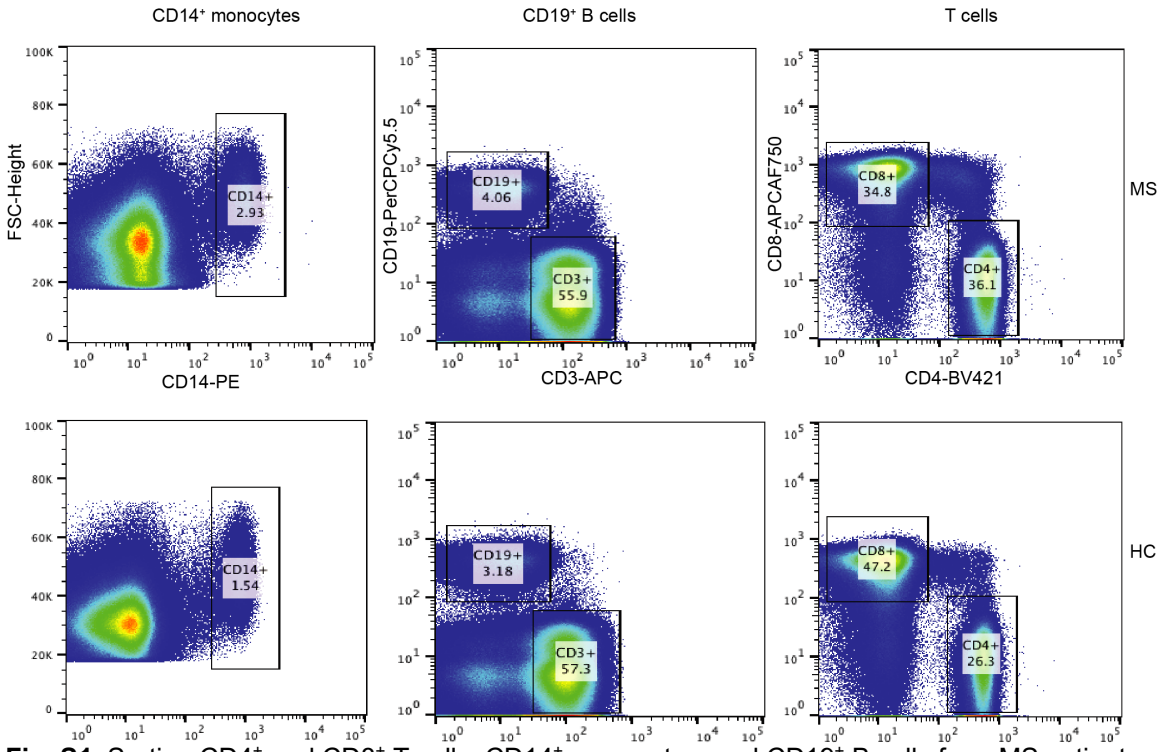

Fig. S1. Sorting CD4⁺ and CD8⁺ T cells, CD14⁺ monocytes, and CD19⁺ B cells from MS patients and healthy controls by flow cytometry (FACS). Flow cytometry data for representative MS and HC subjects. Sample sizes for each cell type can be found in **Dataset S1** and **Dataset S2**.

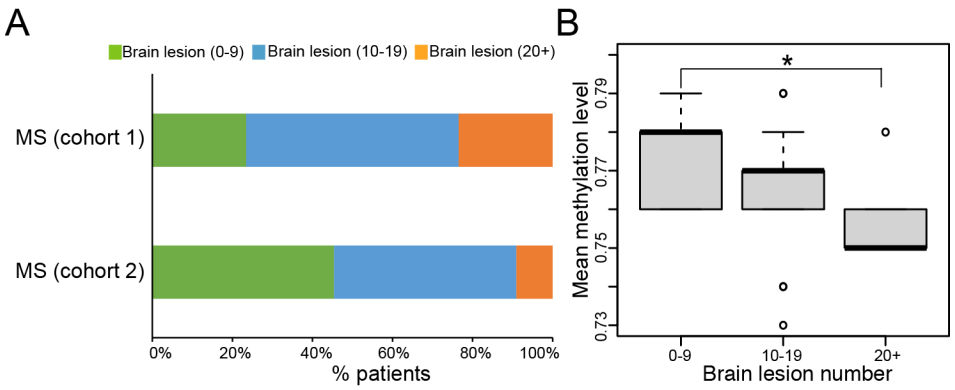

Fig. S2. Methylation level is associated with the accrual of brain lesions. (**A**) Brain lesion numbers in the first and second cohort. (**B**) Box plot of methylation levels in patients with different brain lesion numbers. Significance was determined by two-tailed Student's t-test. *P<0.05.

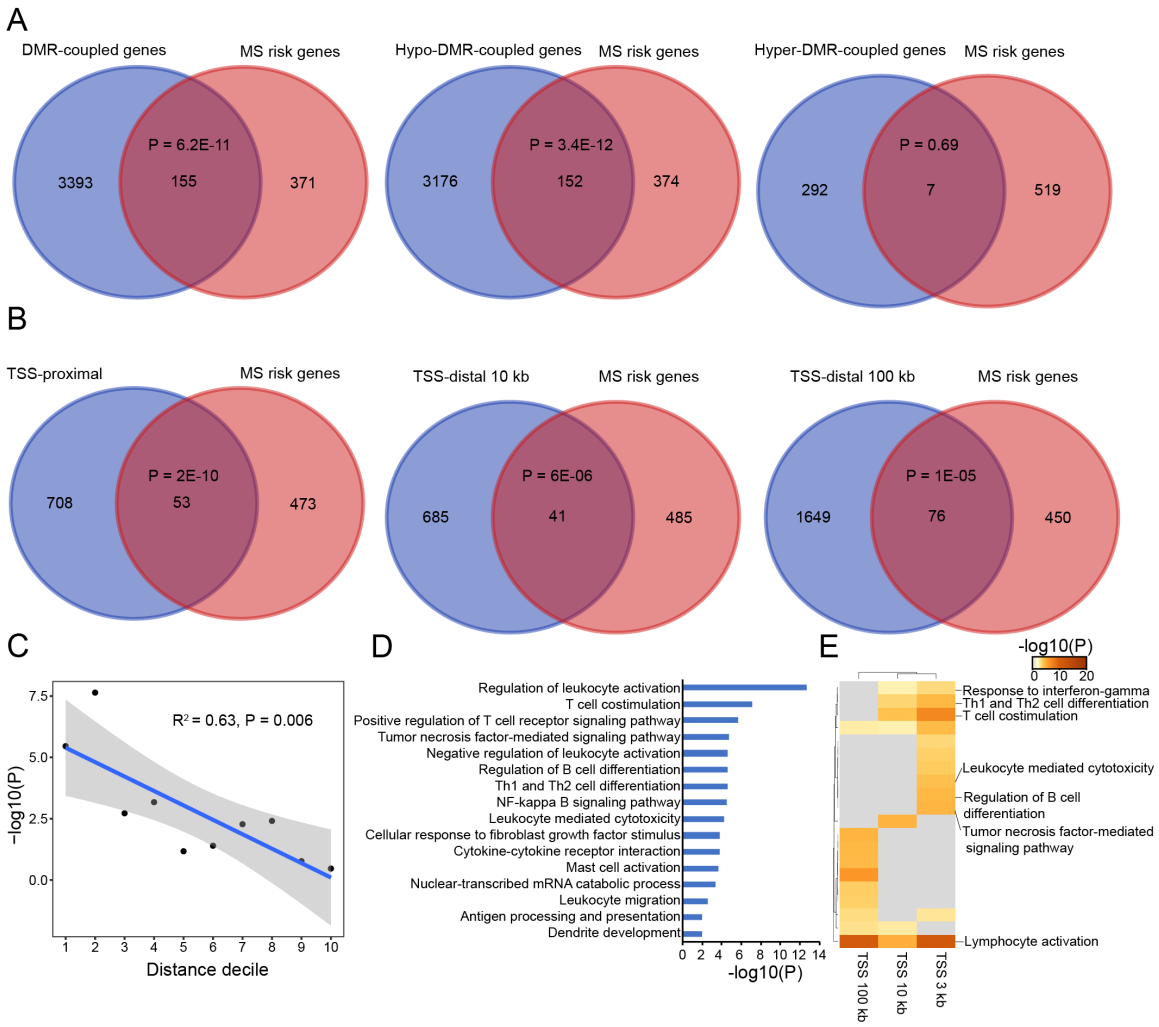

Fig. S3. Overlap between DMRs and MS risk genes. (**A**) Venn diagrams showing the overlap between DMR-coupled (left) or hypo-DMR-coupled (middle) or hyper-DMR-coupled (right) coding genes and MS risk genes. The statistical significance of the overlap was assessed by Fisher's exact test using GeneOverlap package. (**B**) Hypo-DMRs were classified as TSS-proximal (3 kb upstream and downstream of the TSS) (left), TSS-distal 10 kb (10 kb upstream and downstream of the TSS) (middle), and TSS-distal 100 kb (100 kb upstream and downstream of the TSS) (right). Venn diagrams showing the overlap between hypo-DMR-coupled coding genes in each category and MS risk genes. The statistical significance of the overlap was assessed by Fisher's exact test using GeneOverlap package. (**C**) Correlation between DMR distance to TSS and the overlap significance. The DMRs was divided in ten groups according to the distance, and the significance of the overlap in each group was assessed by Fisher's exact test using the GeneOverlap package. Linear regression line with 95% confidence bands was presented. (**D**) Histograms showing the significant GO terms and KEGG pathways enriched for the 53 TSS-proximal hypo-DMR-coupled risk genes in **B**. (**E**) Heatmap visualization of enriched terms across the overlapped genes shown in **B**. The pathways that are significant in all lists or specific enriched in TSS-proximal DMR-coupled risk genes are indicated on the right.

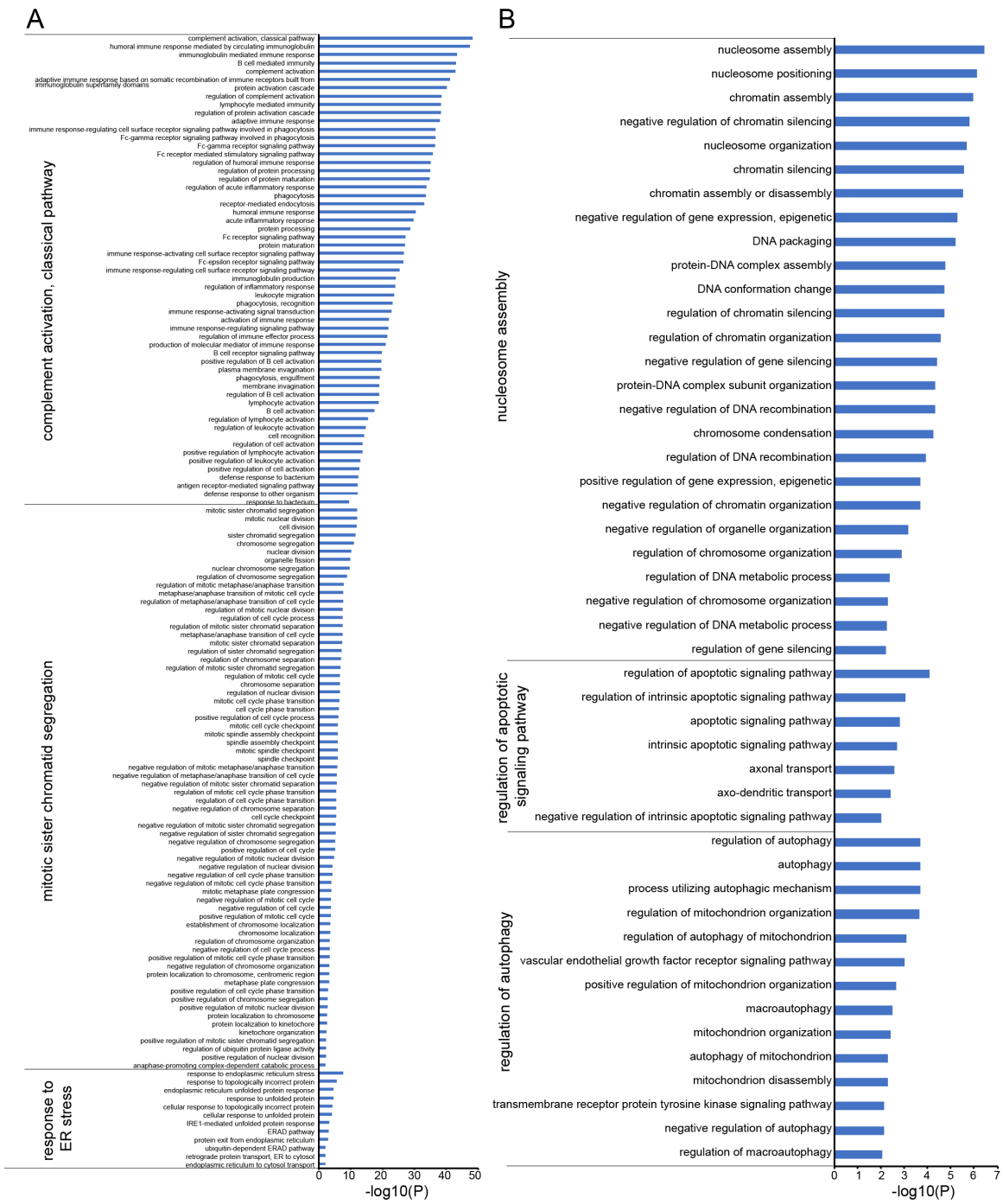

Fig. S4. GO enrichment of the differentially expressed genes (DEGs). (**A**) Bar charts of top significant enriched GO biological processes for the up regulated genes in MS. (**B**) Bar charts of top significant enriched GO biological processes for the down regulated genes in MS. The terms are associated into functional groups based on shared genes. Full lists of enriched categories are provided in **Dataset S9** and **Dataset S10**.

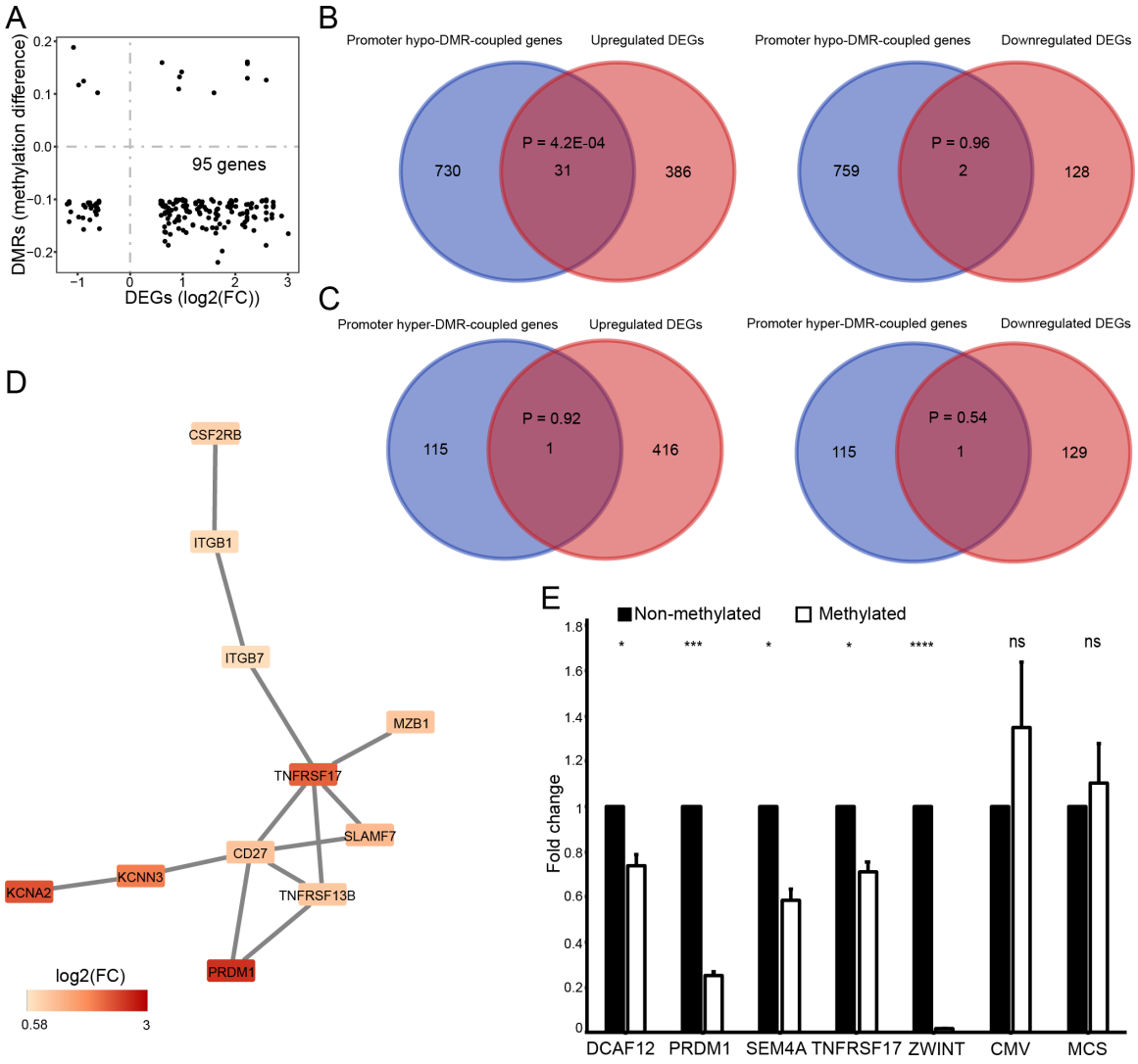

Fig. S5. The overlap between DMRs and DEGs. (**A**) Scatter plots of the correlation between DMRs and DEGs. The y-axis corresponds to methylation difference between MS and HC in each DMR and the x-axis corresponds to the log2 fold change (FC) of the DMR-coupled gene. 95 hypo-DMR-coupled gene are significant upregulated in MS. (**B**) Venn diagrams showing the overlap between promoter hypo-DMR-coupled genes and up- (left) or down-regulated (right) genes. Hypo-DMRs located within 3 kb up- and down-stream regions of the TSS were defined as promoter hypo-DMR. The statistical significance of the overlap was assessed by Fisher's exact test using GeneOverlap package. (**C**) Venn diagrams showing the overlap between promoter hyper-DMR-coupled genes and up- or down-regulated genes. Hyper-DMRs located within 3 kb up- and down-stream regions of the TSS were defined as promoter hyper-DMR. (**D**) Largest connected component of the PPI network using the 31 overlapping genes in left panel of **B** as input. (**E**) Luciferase assay validation. Luciferase reporter vectors bearing the promoter regions of *DCAF12*, *PRDM1*, *SEM4A*, *TNFRSF17*, and *ZWINT* were methylated in vitro and their luciferase activity was compared to the same constructs but not methylated. A luciferase vector carrying the CpG-free CMV promoter and another without any promoter region were used as controls.

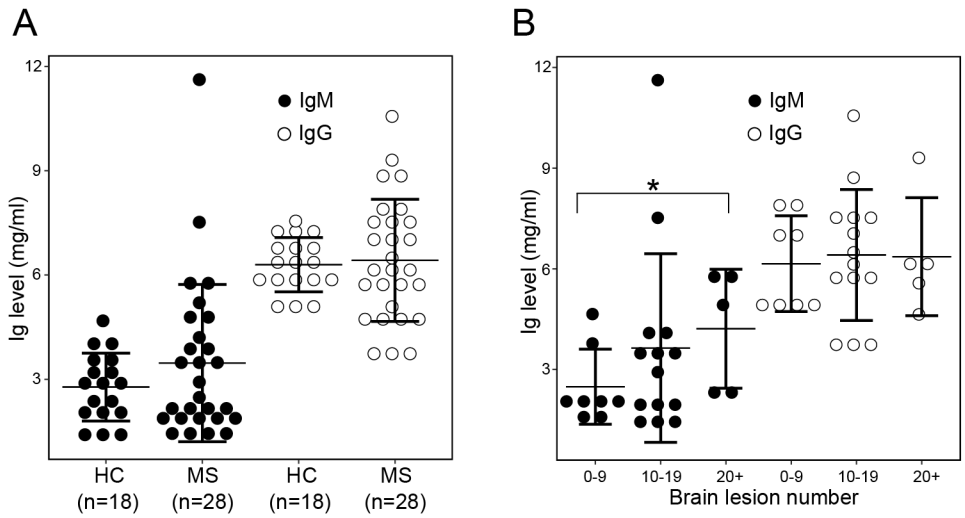

Fig. S6. Serum IgM levels are associated with accrual of brain lesions and global methylation levels. (**A**) The serum levels of IgM and IgG between MS patients and controls. (**B**) The serum levels of IgM and IgG between patients with different brain lesion numbers. Significance was determined by Mann–Whitney U test. *P<0.05.

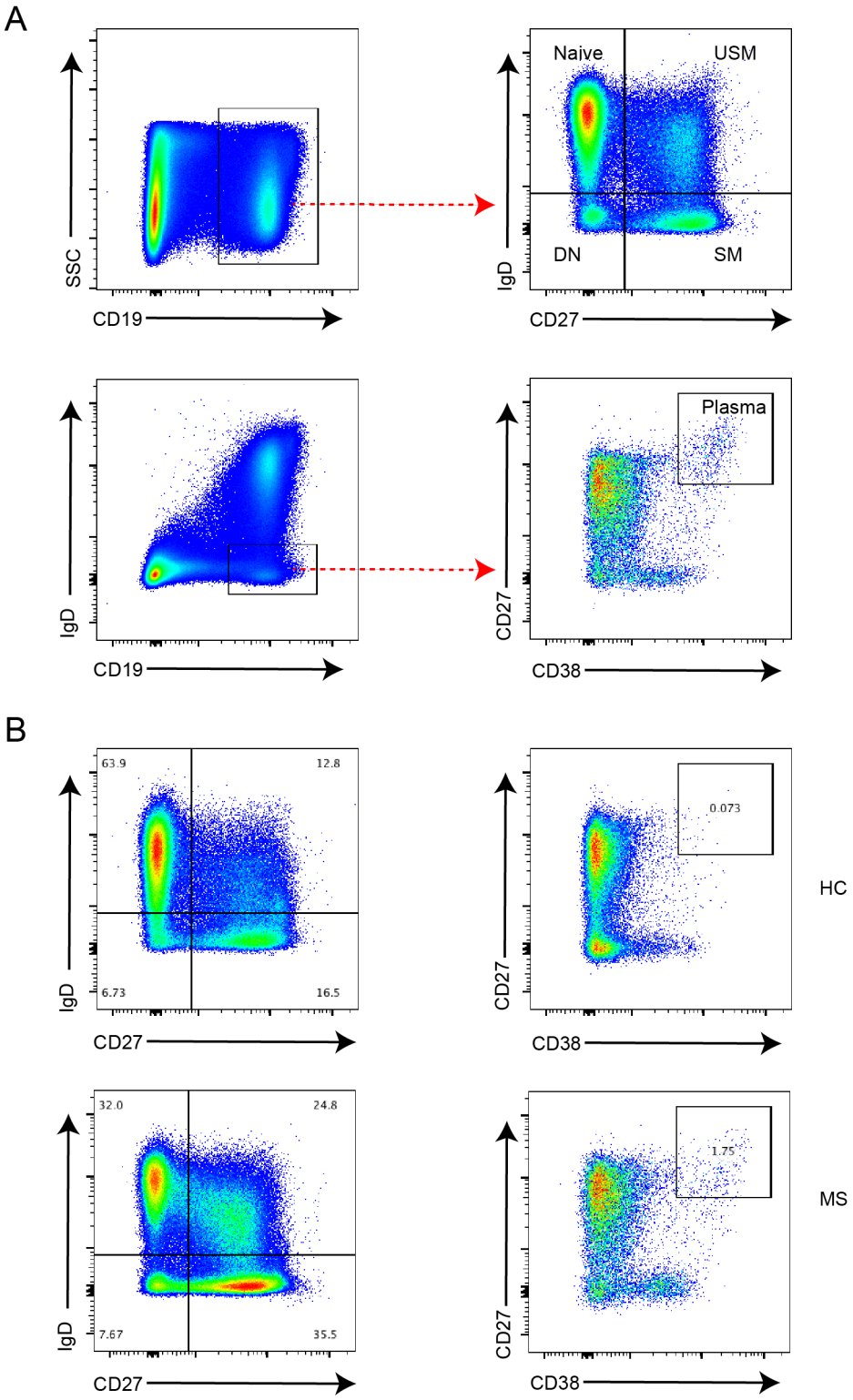

Fig. S7. Identification of B cell subsets in MS patients and healthy controls by flow cytometry. (**A**) Flow cytometry gating strategy for identification of B cell subsets: Naïve B cells (CD19+CD27-IgD+), unswitched memory (USM) B cells (CD19+CD27+IgD+), switched memory (SM) B cells (CD19+CD27+IgD-), double-negative (DN) B cells (CD19+CD27-IgD-) and plasma cells (CD19+IgD-CD27hiCD38hi). **(B)** Flow cytometry data for representative MS and HC subjects.

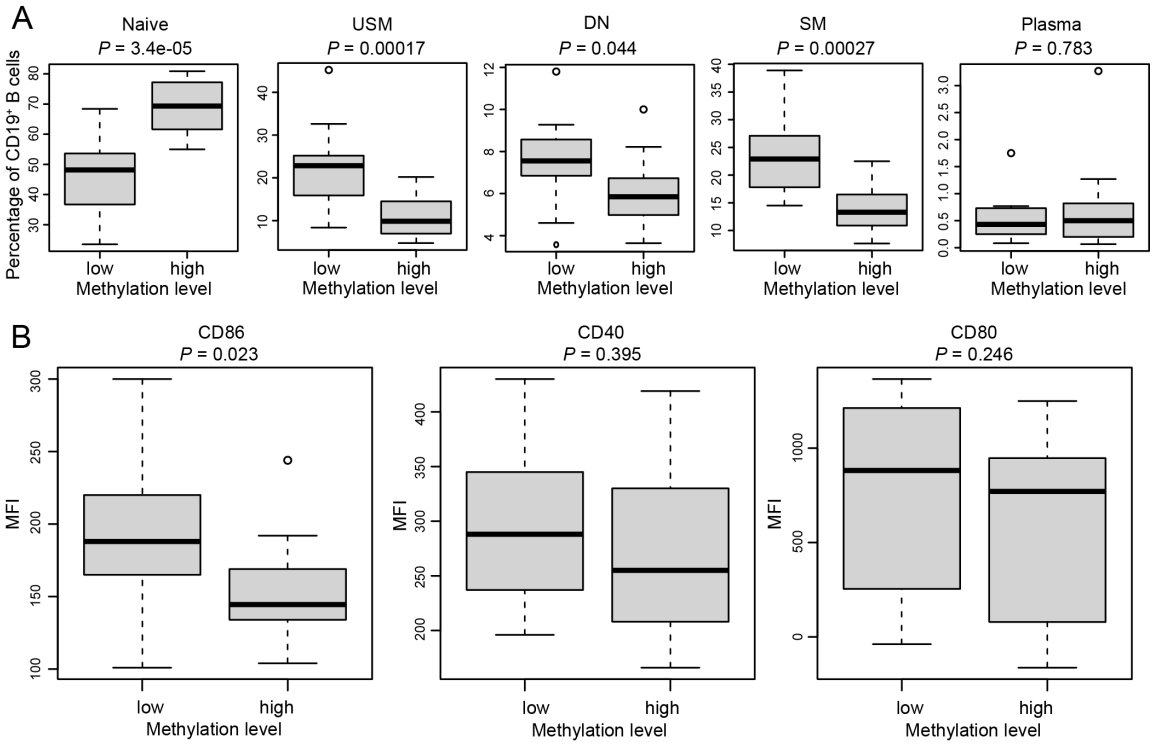

Fig. S8. The association between DNA hypomethylation and B cell differentiation and activation. (**A**) Box plots showing the percentage of each B cell subset in subjects with different methylation levels. The subjects were stratified into two equal groups according to methylation levels. Statistical tests were performed using the Mann–Whitney U test. USM, unswitched memory; DN, double-negative; SM, switched memory. (**B**) Box plots showing the mean fluorescence intensity (MFI) of CD86, CD40 and CD80 expression on $CD19^+$ B cells in subjects with different methylation levels. The subjects were stratified into two equal groups according to methylation levels. Statistical tests were performed using the Mann–Whitney U test.

Dataset S1. Summary of BS-seq in four immune cell types from 29 subjects.

Dataset S2. Summary of BS-seq in B cells from 53 subjects.

Dataset S3. Demographic and clinical information for study participants.

Dataset S4. Genes assigned with DMRs for four cell populations.

Dataset S5. Statistical significance of overlap of hypo-DMRs and transcription factor binding peaks.

Dataset S6. Transcriptional factor binding motif enriched in hyper-DMRs.

Dataset S7. Hypo-DMR-coupled coding genes that are overlapped with MS risk genes.

Dataset S8. Differentially expressed genes between MS patients and healthy controls.

Dataset S9. List of GO terms and KEGG pathways associated with up-regulated genes.

Dataset S10. List of GO terms and KEGG pathways associated with down-regulated genes.

Dataset S11. Differentially expressed genes that are annotated as epigenetic factors.

Dataset S12. List of GO terms and KEGG pathways associated with 95 up-regulated hypo-DMR-associated genes.

Dataset S13. Serum TRAIL, BAFF, IL6 and IL10 levels in patients and controls.

Dataset S14. List of primers used in PCR experiments.

Members of the UCSF MS-EPIC Team (alphabetize by last name)

Jessa Alexander, Riley Bove, Sergio Baranzini, Bruce A. C. Cree, Eduardo Caverzasi, Richard Cuneo, Stacy J. Caillier, Tiffany Cooper, Ari J. Green, Chu-Yueh Guo, Jeffrey M. Gelfand, Refujia Gomez, Sasha Gupta, Jill Hollenbach, Meagan Harms, Roland G. Henry, Stephen L. Hauser, Myra Mendoza, Jorge R. Oksenberg, Nico Papinutto, Sam Pleasure, Adam Santaniello, Joseph J. Sabatino Jr., William A. Stern, Michael R. Wilson, Scott Zamvil