Les canaux calciques Cav1 comme cible thérapeutique dans l'asthme allergique

N. Giang, T. Villeneuve, L. Pelletier, M. Savignac

LES CANAUX CALCIQUES CA_v1 COMME CIBLE THERAPEUTIQUE DANS L'ASTHME ALLERGIQUE

CA_v1 CALCIUM CHANNELS AS A POTENTIEL THERAPEUTIC TARGET IN ALLERGIC ASTHMA

Pharm. D, Nicolas Giang,
Centre de Physiopathologie de Toulouse Purpan INSERM UMR1043, Centre Hospitalier Universitaire Purpan, Place du Dr Baylac, 31024 Toulouse Cedex 3. e-mail : nicolas.giang@inserm.fr, téléphone : 33 5 62 74 45 17, Fax : 33 5 62 74 45 58

Thomas Villeneuve
Centre de Physiopathologie de Toulouse Purpan INSERM UMR1043, Centre Hospitalier Universitaire Purpan, Place du Dr Baylac, 31024 Toulouse Cedex 3. e-mail : thomas.villeneuve@inserm.fr, téléphone : 33 5 62 74 45 17, Fax : 33 5 62 74 45 58

MD-PhD, Lucette Pelletier
Centre de Physiopathologie de Toulouse Purpan INSERM UMR1043, Centre Hospitalier Universitaire Purpan, Place du Dr Baylac, 31024 Toulouse Cedex 3. e-mail : lucette.pelletier@inserm.fr, téléphone : 33 5 62 74 83 78, Fax : 33 5 62 74 45 58

PhD, Magali Savignac
Centre de Physiopathologie de Toulouse Purpan INSERM UMR1043, Centre Hospitalier Universitaire Purpan, Place du Dr Baylac, 31024 Toulouse Cedex 3. e-mail : magali.savignac@inserm.fr, téléphone : 33 5 62 74 45 17, Fax : 33 5 62 74 45 58

Combining evidence from four immune cell types identifies DNA methylation patterns that implicate functionally distinct pathways during Multiple Sclerosis progression

Ewoud Ewing[1], Lara Kular[1], Sunjay Fernandes[2,10], Nestoras Karathanasis [3,4], Vincenzo Lagani[5,6], Sabrina Ruhrmann[1] , Ioannis Tsamardinos[6,7], Jesper Tegner[2,8,10], Fredrik Piehl[1,11], David Gomez-Cabrero[2,11,12,*], Maja Jagodic[1,*,#]

[1]Department of Clinical Neuroscience, Center for Molecular Medicine, Karolinska Institutet, Stockholm, 171 77, Sweden.

[2]Unit of Computational Medicine, Department of Medicine, Solna, Center for Molecular Medicine, Karolinska Institutet, Stockholm, 171 77, Sweden.

[3]Institute of Computer Science, Foundation for Research and Technology-Hellas, Heraklion, Greece.

[4]Computational Medicine Center, Thomas Jefferson University, 1020 Locust Street, Philadelphia, PA 19107 USA.

[5]Institute of Chemical Biology, Ilia State University, Tbilisi, Georgia.

[6]Gnosis Data Analysis PC, Heraklion, Greece.

[7]Department of Computer Science, University of Crete, Heraklion, Greece.

[9]Biological and Environmental Sciences and Engineering Division, Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology, Saudi Arabia.

[10]Science for Life Laboratory, Solna, Sweden.

[11]Center for Neurology, Academic Specialist Clinic, Stockholm Health Services, Stockholm, Sweden

[12]Translational Bioinformatics Unit, Navarrabiomed, Complejo Hospitalario de Navarra (CHN), Universidad Pública de Navarra (UPNA), IdiSNA, Pamplona, Spain.

[13]Centre for Host Microbiome Interactions, Faculty of Dentistry, Oral & Craniofacial Sciences King's College London, UK.

*Equal contribution.

[#]Corresponding author:

Maja Jagodic

Address: CMM, L8:04, Karolinska University Hospital, 171 76 Stockholm Sweden

Email: maja.jagodic@ki.se

Phone: +46-8-517 762 58

Abstract

Background

Multiple Sclerosis (MS) is a chronic inflammatory disease and a leading cause of progressive neurological disability among young adults. DNA methylation, which intersects genes and environment to control cellular functions on a molecular level, may provide insights into MS pathogenesis.

Methods

We measured DNA methylation in $CD4^+$ T cells (n=31), $CD8^+$ T cells (n=28), $CD14^+$ monocytes (n=35) and $CD19^+$ B cells (n=27) from relapsing-remitting (RRMS), secondary progressive (SPMS) patients and healthy controls (HC) using Infinium HumanMethylation450 arrays. Monocyte (n=25) and whole blood (n=275) cohorts were used for validations.

Findings

B cells from MS patients displayed most significant differentially methylated positions (DMPs), followed by monocytes, while only few DMPs were detected in T cells. We implemented a non-parametric combination framework (omicsNPC) to increase discovery power by combining evidence from all four cell types. Identified shared DMPs co-localized at MS risk loci and clustered into distinct groups. Functional exploration of changes discriminating RRMS and SPMS from HC implicated lymphocyte signaling, T cell activation and migration. SPMS-specific changes, on the other hand, implicated myeloid cell functions and metabolism. Interestingly, neuronal and neurodegenerative genes and pathways were also specifically enriched in the SPMS cluster.

Interpretation

We utilized a statistical framework (omicsNPC) that combines multiple layers of evidence to identify DNA methylation changes that provide new insights into MS pathogenesis in general, and disease progression, in particular.

Funding

This work was supported by the Swedish Research Council, Stockholm County Council, AstraZeneca, European Research Council, Karolinska Institutet and Margaretha af Ugglas Foundation.

Keywords: DNA methylation, Epigenetics, Multiple Sclerosis, Immune cells, Secondary progressive Multiple Sclerosis, Relapsing-remitting Multiple Sclerosis, 450K, EPIC, omicsNPC

Research in Context

Evidence before this study

While previous studies implicated DNA methylation changes in immune cells from MS patients, there was a very limited overlap between the findings. These studies predominantly focused on the RRMS stage of disease and changes in T cells.

Added value of this study

We investigated DNA methylation changes in both RRMS and SPMS stages and in four immune cell types implicated in MS pathogenesis, i.e. $CD4^+$ and $CD8^+$ T cells, $CD14^+$ monocytes and $CD19^+$ B cells. We observed evidence of shared DNA methylation changes across all cell types and we implemented a non-parametric combination framework (omicsNPC) to identify such differences taking advantage of increased power when multiple layers of evidence are combined. Notably, omicsNPC is applicable in any context where omics from multiple cell types (or multiple omics from the same cell type) are available. Shared disease-associated differences clustered individuals into distinct functional groups suggesting both known and novel pathways in MS pathogenesis.

Implications of all the available evidence

DNA methylation changes, similar to multiple other lines of evidence, implicate dysregulation of adaptive immune mechanisms in the pathogenesis of MS. Additionally, SPMS-specific DNA methylation changes suggest the involvement of myeloid cells, phagocytosis and metabolism, adding to a growing evidence of these mechanisms being important for disease progression. Finally, an intriguing 'brain signature' of neurodegeneration was found for the first time in peripheral immune cells during progressive disease.

Introduction

Multiple Sclerosis (MS) is a leading cause of progressive disability in young adults caused by inflammation, demyelination and axonal loss in the central nervous system (CNS) (1, 2). Patients are typically diagnosed between 20-40 years of age with women being affected nearly three times as often as men (3). The immune response causes the breakdown of the blood-brain barrier, infiltration of immune cells into the CNS and subsequent development of inflammatory and demyelinating lesions in both brain and spinal cord (4). Most MS patients (85-90%) are initially diagnosed with the relapsing-remitting form of MS (RRMS), which is characterized with recurring episodes of acute neurological symptoms (relapses) followed by recovery (remission). The majority of RRMS patients eventually convert to a progressive form of MS, i.e. secondary-progressive MS (SPMS) with accumulating axonal damage and neuronal loss and persistent increase in neurological disability. Current disease modulatory treatments (DMT) are mainly effective in controlling the early inflammatory stage of the disease, while the therapeutic efficacy in progressive stages is poor, likely due to a shift from mainly adaptive immune mechanism to more complex and currently less defined processes also involving innate and local tissue reactions (2).

Although the exact cause of MS remains unknown more than 200 genomic loci have been associated with the risk of developing the disease with the genes in the HLA class II locus (in particular *HLA-DRB1*) exerting the strongest influence (5-7). The risk loci collectively support the immune cause of MS and particularly the role of adaptive immunity and $CD4^+$ T cell pathways in triggering the disease. While genetic and environmental factors independently confer modest effects, their combined impact conveys a dramatic increase in the risk of developing MS, suggesting interactions on a molecular level (8). Thus, studying the epigenetic mechanisms, that integrate instructions from genes and environment to control cellular function on the molecular level, represents one avenue to uncover processes of importance for diseases as complex as MS.

The most commonly studied epigenetic mechanism is DNA methylation, which is the covalent addition of a methyl group to the 5th carbon of cytosine, known as 5-methylcytosine (5mC) in a CpG dinucleotide context (9). Generally, DNA methylation within CpG rich promoters of genes is associated with transcriptional repression, while higher methylation in gene bodies has been shown to positively correlate with expression (10). We have recently demonstrated that DNA methylation mediates risk of developing MS (11). Several studies have compared DNA methylation changes between MS patients and controls in CD4$^+$, CD8$^+$, CD14$^+$, CD19$^+$ cells and bulk peripheral blood mononuclear cells using the same methodology to measure DNA methylation genome-wide, i.e. Illumina methylation arrays (11-18). While each study reports potentially interesting candidates, changes in *HLA-DRB1* seem most reproducible likely owing to the strong genetic regulation of methylation in the locus. This lack of reproducibility is caused by the fact that MS is a heterogeneous disease, thus warranting larger cohorts of sorted cells, which is typically challenging, and new analytical methods.

Here we analyzed DNA methylation in four cell types implicated in MS immunopathology (19-21) that were sorted from peripheral blood of RRMS and SPMS patients and healthy controls. We show that immune cells from MS patients share epigenetic changes and we demonstrate a statistical framework to identify such changes, thus increasing the power of identifying disease-associated methylation patterns in complex heterogeneous diseases.

Methods

Cohorts

A discovery cohort comprising persons affected with RRMS (n=16) and SPMS (n=14) and healthy controls (n=21), and an independent validation cohort, comprising persons affected with RRMS (n=14) and healthy controls (n=11), were recruited at the Neurology clinic at Karolinska University Hospital in Stockholm. The RRMS patients were primarily selected based on recent evidence of disease activity, either manifested as relapses or contrast enhancing MRI lesions, and the majority (87.5%) of the RRMS patients have not been treated at the time of sampling. Cohort details, with the exact number of RRMS, SPMS and HC individuals profiled for each cell type, are provided in Table 1, and detailed patient information, including treatment history and disease activity, is supplied in Supplementary Table 1. The Regional Ethical Review Board in Stockholm approved the study and methods were carried out in accordance with institutional guidelines for experiments with human subjects. Informed consent was obtained from all subjects.

The whole blood cohort used to replicate functional pathways in SPMS, consisting of RRMS (n=119), SPMS (n=17) and healthy controls (n=139), was described in detail elsewhere (11, 22).

Sample preparation

Peripheral blood mononuclear cells (PBMCs) from the discovery and validation cohort were isolated directly after collection using a standard Ficoll (GE Healthcare) and sodium citrate-containing preparation tubes (Becton Dickinson) procedures, respectively. Monocytes were isolated using $CD14^+$ positive selection on MACS microbeads magnetic separation (Miltenyi), according to manufacturer's instructions (> 95% purity). Sorting of $CD4^+$ and $CD8^+$ T cells and $CD19^+$ B cells was performed from the negative fraction obtained after sorting of monocytes by adding fluorochrome-conjugated antibodies against human CD4 (clone SK3, APC-conjugated, Becton Dickinson), CD8 (clone SK1, FITC-conjugated, Becton Dickinson), CD3

(clone UCHT1, PE-conjugated, BD Bioscience) and CD19 (clone SJ25C1, APC-Cy7-conjugated, Becton Dickinson) using high-speed MoFlo™ cell sorter (Beckman Coulter, Inc, > 99% purity). Extraction of genomic DNA was performed using Gen Elute Mammalian Genomic DNA Miniprep kit (Sigma-Aldrich). The amount and quality of DNA were assessed with a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies Inc). The four cells types were sorted from all individuals in the discovery cohort and samples with sufficient DNA amounts were used in further analysis. The numbers of RRMS, SPMS and HC used for each cell type are provided in Table 1 and details of individuals are given in Supplementary Table 1. Processing of the discovery cohort samples for Infinium HumanMethylation450 arrays (Illumina), including bisulfite conversion, was done at the Bioinformatics and Expression Analysis core facility (BEA), Karolinska Institutet (Stockholm) for $CD14^+$ monocytes and $CD4^+$ T cells, and at Johns Hopkins University School of Medicine (Baltimore) for $CD8^+$ T cells and $CD19^+$ B cells. Processing of the validation cohort samples for Infinium MethylationEPIC arrays (Illumina) was done at the SNP&SEQ Technology platform (Uppsala). Cases and controls were randomized on the arrays.

DNA methylation analysis

Methylation profiles for every cell type were analyzed individually in R using the Minfi (23) and ChAMP package (24) following the pipeline according to Marabita *el al* (25). Briefly, type 1 and type 2 probes were normalized using quantile normalization and BMIQ. Sex of the samples was confirmed using the GetSex function from the Minfi package and the cell type identity was confirmed using the cell type deconvolution method from Minfi based on the Houseman algorithm (26). The following probes were filtered out: i) probes not passing the detection p-value cutoff of 0.01, ii) probes with known SNPs, and iii) X and Y chromosome probes. Batch effects were identified using principal component analysis (PCA) and corrected using ComBat from the SVA package (27). The loading methylation profiles was performed in this manner for each cohort used in this study. Differentially methylated positions (DMPs) were determined with linear

modeling using the limma package (28) in a model that included age and sex as covariates. The influence of treatment has been investigated using both PCA and covariate regression and potential confounding effects of the treatment status in this cohort have been excluded. Differences were calculated between RRMS and healthy controls, RRMS and SPMS, SPMS and healthy controls. In addition, eBayes was used to find differences in at least one of the comparisons.

Non-parametric combination methodology (omicsNPC)

In order to increase statistical power by using multiple layers of evidence, we applied the non-parametric combination (NPC)(29-31) methodology as implemented in the omicsNPC (32) function of the STATegRa R package. The omicsNPC procedure combines results from a series of statistical tests in order to produce a single global p-value that summarizes evidence from all tests. In this study, our goal was to identify probes whose differential methylation is detected in multiple cell types.

In short, for each probe i the limma results from the individual cell types $j = 1, \dots, n$ were combined by omicsNPC using the Liptak-Stouffer function: $T_i = \sum_j \Phi^{-1}\left(1 - \lambda_i^j\right)$, where Φ^{-1} is the normal inverse cumulative distribution function, and λ_i^j is the p-value corresponding to probe i and cell type j. The global statistic T_i is then transformed in a global p-value p_i by using a permutation approach. Notably, permutations are performed by randomly re-arranging the patients' status information (RRMS, SPMS, healthy control) across all cell types in a coordinated way. In this way the association between each measurement and the patients' status is disrupted, while the correlation structure across measurements from different cell types is left unaltered and accounted for. Neglecting such correlations would possibly lead to false positive associations. By using the Liptak-Stouffer function, significant global p-values are produced for probes that are differentially methylated, even mildly, in multiple cell types, thus supported by multiple evidence.

Finally, in our analysis omicsNPC was run with 10.000 permutations on all probes with a nominal p-value < 0.05 in individual cell type analyses. Statistical significance for the omicsNPC results was defined as Liptak p-value < 0.0001 and Liptak FDR < 0.2.

Clustering

Significant omicsNPC probes were individually transformed from beta values to cell specific Z-scores by subtracting the mean and dividing by the standard deviation. The matrix was clustered into distinct groups using MClust (33). Based on the direction of change within each cluster, the clusters were merged and assigned to one of the following groups: Unk, MS and SP.

Functional annotations

Genes associated with DMP probe IDs from the Illumina manifest were uploaded to Ingenuity Pathway Analysis (IPA) database (Qiagen) and core expression analysis was performed to identify affected canonical pathways and functional annotations. Immune tissues, including primary immune cells and cell lines, were used. Right-tailed Fisher's exact test was used to calculate a p-value determining the probability that each biological function assigned to that data set is due to chance alone. Canonical pathways/functional annotations were grouped into clusters by calculating the similarity of pathways/annotations using the relative risk (RR) of each pathway appearing with each pathway based on the genes enriched within the pathway. Only pathways representing a minimum of 5 differentially methylated genes were selected for functional exploration. RR scores were clustered into groups using kmeans. Genes associated with DMPs with absolute $\Delta\beta > 5\%$, p-value < 0.001 were used for pathway analysis of changes identified in individual cell types. Genes associated with DMPs defined as significant in omicsNPC and clustered in specific groups were used for canonical pathways and functional annotation analysis of shared changes. Over-

representation analysis (ORA) as implemented in webgestalt (34) was used to identify cluster labels. ORA was also used in the analysis of the whole blood data genes and visualized using REVIGO (35) .

Meta-analysis

Meta-analysis of $CD14^{+}$ cells for the comparison of RRMS and healthy controls in the discovery and validation cohort on the 377 607 probes shared between 450K and EPIC platforms was conducted with a random effects model (REML) using the metap and metafor R-package (36). Directionality was determined using the Metal pipeline.

Overlap with disease-associated loci and meQTLs

To determine if there is an overlap of the top-ranked shared DMPs from omicsNPC with MS-associated genetic variants, we used a set of 234 recently reported MS-associated SNPs (7). The Genomic Association Test (GAT) tool (37) was applied to estimate the significance of the overlap between MS-associated SNPs and omicsNPC DMPs. The analysis was run for bins of 2 kb windows based on the average distance between SNPs and CpGs taken from GeMes (38). For a comparison, the overlap was tested for other diseases matched for the number of SNPs with MS as well as for common control SNPs which were matched in the CpG probe density for the bins run.

Furthermore, omicsNPC probes were investigated for potential meQTLs using the Blueprint data (39). OmicsNPC probes were extracted from the full Blueprint dataset comprising monocytes and T cells. MeQTLs were considered significant if Bonferroni-corrected p-value < 0.05.

Results

Disease-associated DNA methylation patterns in four immune cell types

We have profiled DNA methylation in $CD4^+$ T cells (n=31), $CD8^+$ T cells (n=28), $CD14^+$ monocytes (n=35) and $CD19^+$ B cells (n=27) from MS patients and healthy controls (HC) (Table 1, Table S1) using Infinium HumanMethylation450 arrays (450K). All four cell types have been implicated in the pathogenesis of MS (19-21). After adjustment for confounders, we found 1 511, 666, and 30 significant differentially methylated positions (DMPs, adj. p-value < 0.05) in $CD19^+$, CD14+ and CD8+ cells, respectively, between RRMS, SPMS and HC individuals (Fig. 1A, Table S2). B cells displayed more differences between RRMS and HC (3 904 DMPs, abs. $\Delta\beta > 5\%$, adj. p-value < 0.05) compared to any other cell type (0, 1, 124 in $CD4^+$, $CD8^+$, $CD14^+$ cells, respectively). In total, ~70% (2 662/3 904) of DMPs between RRMS and HC in $CD19^+$ B cells displayed hypomethylation in RRMS (Table S2), which was also reflected on the level of the most variable DMPs (Fig. 1A). The opposite pattern was observed in $CD14^+$ monocytes (Fig. 1B), which bear the second highest number of significant differences, where ~90% (110/124) of DMPs between RRMS and HC (abs. $\Delta\beta > 5\%$, adj. p-value < 0.05) displayed hypermethylation in RRMS (Table S2). The significant methylation changes identified in B cells and monocytes were particularly enriched in open sea regions and depleted from TSS1500, 5' UTRs and shores (Fig. S1). Unlike B cells and monocytes, T cells displayed very little methylation difference between RRMS, SPMS and HC. Only one CpG was significant between RRMS and HC in $CD8^+$ T cells (Table S2) although the most variable DMPs displayed predominant hypermethylation in RRMS, which is consistent with previous findings (13). Notably, none of the CpGs in $CD4^+$ T cells passed the significance threshold (Fig. 1A, Table S2), despite previous reports (12, 16).

In order to identify biological functions that are affected by the differences in methylation patterns, we performed functional IPA analysis on genes associated with DMPs identified in the different cell types. We focused on candidate differences between RRMS and HC (abs. $\Delta\beta > 5\%$, unadjusted p-value < 0.001) as

these groups had similar size in all four cell types (Table 1). IPA analysis revealed over-representation of immune-related processes, with an enrichment of genes involved in antigen presentation, OX40 signaling, T helper cell differentiation, T lymphocyte apoptosis, and B cell development, among others, and biological functions reflecting immune cell migration and inflammatory response (Table S3). Interestingly, the majority of canonical pathways and biological functions overlapped between the four cell types (Fig. 2A), implying that similar functions may be affected by methylation changes in $CD4^+$, $CD8^+$, $CD14^+$ and $CD19^+$ cells in RRMS patients compared to controls. This is further supported by the strong correlation of changes ($\Delta\beta$) between cells types (Fig. 2B), i.e. a large fraction of CpGs exhibited the same direction of the change between RRMS and HC in all four cell types.

These findings indicate that in addition to cell type-specific effects there is a substantial fraction of DNA methylation changes that may be shared across the immune cell types implicated in MS pathogenesis.

Combining multiple immune cell types increases power to identify disease-associated DNA methylation patterns

In order to increase statistical power by using multiple layers of evidence, i.e. from $CD4^+$, $CD8^+$, $CD14^+$ and $CD19^+$ cells, we applied the non-parametric combination methodology as implemented in the omicsNPC function (32) (Fig. 3A). This stepwise approach builds on permutations of the moderated F-statistics from all probes passing nominal p-value < 0.05 in any of the comparisons (clinical groups or cell types), which were combined using the Liptak-Stouffer function. This function requires support from most of the individual analyses in order to provide a significant overall p-value (i.e., probes with a low p-value only in one single cell type are unlikely to achieve a significant overall p-value). OmicsNPC analysis for different combinations of the cell types resulted in 1 976 DMPs for all four cell types, 1 273 DMPs for lymphocytes ($CD4^+$, $CD8^+$ and $CD19^+$ cells), 423 DMPs for T cells ($CD4^+$ and $CD8^+$ cells) and 2 782 DMPs for cells with the antigen-presenting potential ($CD14^+$ and $CD19^+$ cells) (Fig. 3B, Table S4).

Interestingly, the directionality of the significant omicsNPC DMPs was shared across different cells types significantly more than expected by chance (χ2-test p-value < 0.05). For example, the majority of the omicsNPC hypomethylated DMPs in the $CD19^+$ B cells were found hypomethylated in the three other cell types as well (Fig. 4A). Furthermore, the shared directionality was also seen when comparing the T-statistics from Liptak significant probes, which displayed a high correlation between cell types within each comparison (Fig. 4B). Overall, omicsNPC methodology allowed robust identification of a substantial number of DMPs with evidence of a DNA methylation change across multiple cell types.

To address whether omicsNPC increases the discovery power, we used a validation cohort comprising methylation data from $CD14^+$ cells isolated from RRMS (n = 14) and HC (n = 11) generated using EPIC arrays. After selecting for the shared probes between the two Illumina platforms (n = 377 607, Table S5), we performed a random effect meta-analysis between $CD14^+$ methylation profiles from the two cohorts. As expected, the meta-analysis resulted in the identification of a larger number of DMPs between RRMS and HC individuals (Table S5). Comparison with omicsNPC showed that most of the additional DMPs identified in $CD14^+$ after conducting a meta-analysis of the two cohorts ranked in the top of the omicsNPC DMPs, however these probes did not rank in the top of the DMPs identified in the original 450K $CD14^+$ cohort alone (Fig. 4C). Among top 10 000 ranked DMPs, up to 27% of the top ranking omicsNPC DMPs were also top ranking in the meta-analysis, especially when considering omicsNPC comparisons containing $CD14^+$ cells (e.g. shared between $CD14^+$ and $CD19^+$, or shared across all four cells types).

Collectively, these data indicate that the omicsNPC methodology combines evidence from distinct yet disease-relevant cell types to increase the discovery power.

Co-localization of disease-associated omicsNPC CpGs with MS susceptibility loci

We examined the possible mechanisms underlying omicsNPC DMPs with evidence in $CD4^+$, $CD8^+$, $CD14^+$ and $CD19^+$ cells. Given the overlap of individuals between the cell type cohorts, ranging from 45% to 79%

between two cell types, shared methylation changes could reflect genetically-controlled methylation changes known as meQTLs (38, 39). Thus, we compared our omicsNPC DMPs from the combination of all four cell types with meQTL data detected in naïve $CD4^+$ T cells and $CD14^+$ monocytes from 197 individuals. Of 1 976 omicsNPC DMPs, only 261 (13.2%) displayed significant (Bonferroni adj.p < 0.05) meQTLs with the same SNPs in both cell types.

Moreover, genetic influences from the MS susceptibility loci could provide another biological explanation for the observed shared methylation changes in functionally distinct cell types. To test this hypothesis, we investigated the co-localization of omicsNPC probes and MS-associated genetic loci (7). Significant CpGs from different omicsNPC combinations were tested for being enriched in genetic regions associated with MS as well as other inflammatory and non-inflammatory diseases that have their genetic architecture similar to that of MS. In total, 234 MS-associated SNPs were taken from the most recent association study (7), while SNPs associated with other diseases were taken from the GWAS catalog (https://www.ebi.ac.uk/gwas/). There was a significant co-localization of omicsNPC CpGs with the MS associated SNPs, while no overlap could be found with the SNPs that associate with asthma, bone mineral density, major depressive disorder, psoriasis, as well as common control SNPs (Fig. 4D).

These data imply that at least a fraction of methylation changes that are shared across distinct immune cell types may be driven by genetic variants that predispose for MS development.

DNA methylation patterns at omicsNPC DMPs implicate functionally distinct pathways during MS progression

To explore if DNA methylation patterns can inform about distinct MS features, we first performed unbiased clustering of individuals based on z-score transformed omicsNPC CpGs derived from a combination of all four cells types. The optimal clustering revealed a grouping of individuals into seven different clusters (Fig. 5A, Table S6). Based on average methylation levels for each cell type, MS status and stage, these seven

clusters were further assigned to three biological groups. The first group (SP) comprised changes that primarily related to the SPMS stage, i.e. most of the differences were related to the SPMS vs HC or RRMS. This group comprised two clusters with 370 hypermethylated and 319 hypomethylated omicsNPC CpGs in SPMS compared to both HC and RRMS (Fig. 5A). The second group (MS) reflected MS-specific changes present in both RRMS and SPMS compared to HC and comprised four clusters, one cluster of 341 hypomethylated omicsNPC CpGs and three clusters with 338, 156 and 8 hypermethylated omicsNPC CpGs in RRMS, and to a lesser extent SPMS, compared to HC (Fig. 5A). The third group (Unk, from unknown) comprised one cluster of 444 omicsNPC CpGs where the differences could not be unambiguously attributed to a specific clinical group and did not always share directionality across cells types (Fig. 5A).

Because the average age of SPMS patients is higher than the average age of RRMS and HC individuals, we investigated if differences in age could have resulted in the identification of changes specific for SPMS, although the age was used as a covariate in our analysis. OmicsNPC probes displayed a minimal correlation with age in different cell types with e.g. < 5% of 1 976 probes showing correlation with age (Spearman r > 0.4) in at least one cell type (Fig. S2A). Similarly, there was a limited overlap of omicsNPC DMPs with known age-related DMPs identified in sorted cells (Fig. S2B) (40) and whole blood from a large longitudinal twin cohort (Fig. S2C) (41). In contrast to omicsNPC DMPs, previously reported age-related DMPs (40, 41) correlated significantly with age also in our cohort (data not shown).

We then investigated the functional relevance of the changes that associated with the three biological groups using IPA. The enriched canonical pathways and functional annotations were grouped together based on the RR clustering (see Methods). The pathway analysis was based on 174 unique genes associated with differentially methylated CpGs, of which only 11 (6%) were shared between the groups, indicating that very specific functions are affected by methylation in different clinical groups. Although there was occasionally overlap between pathway labels between different groups, the RR analysis

demonstrated that the majority of the genes comprising those pathways did not overlap. We observed seven distinct clusters of canonical pathways that differed between the MS and SP groups (Fig. 5B, Table S6). The pathways related to MS in general encompass genes implicated in signaling downstream T- and B-cell receptors as well as in T cell activation. As expected, these pathways comprise many molecules involved in signaling in immune cells such as *IL1RL2*, *GNG4/7*, *IRAK2*, *MAPK14*, *NCOR2*, *PLCB2*, *PTPRJ/O*, *PRKZC*, *RUNX3*, *SMAD9*, *STAT5A* (Table S6, *PRKCZ* is shown in Fig. 5D). The canonical pathways associated to SP include genes involved in cAMP-mediated signaling, NO signaling, metabolism, respiratory burst and phagocytosis. The Unk group showed enrichment of genes related to actin cytoskeleton. Annotation of biological processes revealed three major clusters, each specific for a clinical group, supporting the functional specificity of methylation changes in clinical groups. While the SP-specific functions included development and activation of predominantly myeloid cells, more general MS functions included chemotaxis of both myeloid and lymphoid cells (Fig. 5C, Table S7). For the Unk group the functions included cell-to-cell signaling and interaction, inflammatory response, cell morphology and function of APCs. Several examples of DMPs are shown in Figure 5D.

Surprisingly, many genes in the SP group, despite being detected in immune cells, have previously been involved in neurodevelopmental and/or neurodegenerative functions. They include *APAF1*, *ASIC2*, *BAIAP2*, *CALB2*, *CDH23*, *CLDN14*, *CR1*, *CX3CR1*, *GAB1/2*, *GLI3*, *GNAO1*, *GRID2*, *GRIN1*, *GRM2*, *ITPR2/3*, *JAK2*, *MAPK10*, *NTN1/GN1*, *TGFBR1* and *TUBB2A/6* (Table S6, *GNAO1, JAK2, CALB2* and *GLI3* are shown in Fig. 5D). Indeed, in addition to immune-related processes and functions, the SP group displayed an enrichment of changes in genes implicated in neuronal functions, such as “Axonal Guidance Signaling”, “CREB Signaling in Neurons”, “eNOS Signaling” and “Synaptic Long Term Potentiation/Depression”. To examine this association of SP changes in blood with neurodegenerative functions, we analyzed DNA methylation data from whole blood in an independent cohort (n = 275) (11). Gene Ontology analyses revealed shared (Fig.

368 6A-B) and distinct (Fig. 6C) pathways and biological functions associated to RRMS and SPMS patients in
369 comparison to healthy individuals and confirmed enrichment of neuronal processes in SPMS patients,
370 specifically.

Discussion

We investigated genome-wide DNA methylation in four immune cell types, implicated in the pathogenesis of MS (19-21), from healthy individuals and MS patients in the RRMS and SPMS stage. The most prominent changes were detected in B cells, while no significant changes could be detected in T cells. However, we observed evidence of shared DNA methylation changes across different cell types and we developed a non-parametric framework to detect such changes, thus increasing the power to identify disease-associated differences that can cluster individuals into distinct functional groups and uncover known and novel pathways in MS pathogenesis.

Several studies investigated DNA methylation in immune cells sorted from MS patients using the same Illumina array-based methodology, with negligible overlap between the findings (11-14, 16, 18). Likewise, our cell type-specific analyses demonstrated none and one significant DMP in $CD4^+$ and $CD8^+$ T cells between RRMS and controls, respectively, and no overlap with previous findings (12-14, 16). However, we corroborated previously reported overall higher methylation in $CD8^+$ T cells of RRMS patients (13, 42). Moreover, the top most variable positions, particularly in $CD4^+$ T cells, segregated MS patients from controls, suggesting a lack of power to identify true underlying differences. Indeed, a recent meta-analysis in $CD4^+$ and $CD8^+$ T cells demonstrated two significant DMRs mapping to *HLA-DRB1* and *SLFN12* genes (42), the same genes that displayed changes in multiple cell types in our study (Table S8). Nevertheless, difficulties to identify methylation changes in T cells suggest that future analysis need to be carried out in a sub-population of T cells, i.e. more relevant rare pathogenic sub-types. Our analysis further suggests that most methylation changes can be detected in bulk B cells and monocytes, although their reproducibility remains to be tested in independent cohorts.

Despite limited power to detect significant differences, the functional immunological pathways associated with the top candidate DMPs appeared enriched in all cell types and there was a significant correlation in

394 DNA methylation changes between different cell types. We, therefore, hypothesized that we can increase
395 the power to identify disease-relevant changes by combining evidence from multiple cell types. For that
396 purpose, we extended the application of the non-parametric combination methodology in the omicsNPC
397 (32) and applied the Liptak-Stouffer function to combine p-values. This function produces significant DMPs
398 supported by multiple evidence, i.e. probes that are differentially methylated, even mildly, in multiple cell
399 types. Indeed, omicsNPC approach enabled identification of more DMPs than analyses in the individual
400 cell types. Furthermore, the directionality of the change for the omicsNPC DMPs was also often shared
401 among the different cell types. As the omicsNPC pipeline uses absolute values for effect size, the shared
402 directionality was not enforced by the methodology but was a result of the analysis. This suggests that the
403 approach increases the discovery power, which we confirmed using an independent cohort of $CD14^+$ cells
404 from RRMS and HC. In addition, a number of omicsNPC significant DMPs, e.g. in *SAMD11* (multiple CpGs),
405 *HLA* class II locus (multiple CpGs) *CASZ*, *TMEM48* and *FSCN2* genes displayed at least nominal significance
406 with the same directionality of the change in previous independent studies of $CD4^+$ and $CD8^+$ T cells or
407 $CD19^+$ B cells (Table S8).

408 MeQTLs could provide one explanation for the shared methylation changes in cells with distinct functions,
409 and it was recently suggested that changes detected in case-control cohorts largely reflect meQTL effects
410 (43). Therefore, we compared our omicsNPC DMPs with significant meQTL detected in $CD4^+$ T cells and
411 monocytes (39). In total, 261/1 976 (13%) omicsNPC DMPs were shown to be genetically regulated by the
412 same SNP in both cell types, implying a potential that some of these CpGs are identified due to a varying
413 genetic background between patients and controls, but these are only a minor fraction of the identified
414 changes. On the other hand, we discovered significant co-localization of the omicsNPC CpGs with the loci
415 involved in MS susceptibility (7). This suggests that shared methylation changes may be under the
416 regulation of disease-predisposing genetic factors in immune cells involved in the pathogenesis.

Accordingly, we have recently shown that methylation in the *HLA-DRB1*, the major genetic risk factor, and the most reproducible methylation differences across studies, can mediate the risk of developing MS (11). Moreover, no significant co-localization with genetic factors of other tested inflammatory (psoriasis) or neuropsychiatric diseases (MDD) was observed, indicating that the identified shared changes show specificity for MS. Another explanation for shared changes could be exposure to the same environmental conditions, be it external (e.g. infections, smoking, sun exposure, vitamin D levels) or internal (e.g. chronic inflammation). However, the relative contribution of these mechanisms warrants further studies.

Interestingly, only methylation changes in B cells, which displayed by far the largest number of methylation changes compared to other cell types, did not display significant co-localization with MS risk loci. As genetic studies of susceptibility typically address factors of disease initiation, this may suggest the involvement of B cells in events other than triggering MS or that the contribution of B cells to MS susceptibility might be conveyed in a non-genetic manner to a greater extent. In that regard, B cells are the primary targets of Epstein Barr virus (EBV) infection, which is one of the major environmental factors associated with susceptibility to develop MS (8). It has been shown that B cell immortalization by EBV results in hypomethylation that affects promoters of proliferative genes and a large part of the B cell genome (44, 45). We have also observed that nearly 70% of DMPs in B cells of RRMS patients display hypomethylation. These findings are also of interest in context of recent experimental observations regarding the non-redundant role of memory B cells in activating memory T cells in an antigen-specific manner, as well as the remarkable efficacy of B cell depleting therapies, both of which supports the notion of an important role for B cells in sustaining inflammatory activity in MS (46, 47).

The omicsNPC DMPs clustered individuals into distinct groups with one group corresponding to changes occurring in MS patients in general and another group comprising changes that are more specific for the SPMS stage. Functional annotation analysis implicated signaling pathways downstream T- and B-cell

receptors and T cell activation to be affected by methylation changes in MS patients in general. These pathways agree with well-recognized role of adaptive immunity in triggering MS (4-7). SPMS patients in the MS group often showed changes that were intermediary, i.e. less pronounced than in the RRMS stage but not at the level of healthy individuals, which may also reflect age-related decline in the adaptive immunity in older individuals (48). The genes affected by methylation changes in the MS group are often involved in signaling in adaptive immune cells including transcription factors *RUNX3* (Runt Related Transcription Factor 3) and *STAT5A* (Signal Transducer And Activator Of Transcription 5A) that are critical for differentiation of cytotoxic T cells (49), and balance between regulatory and effector functions (50), respectively.

Functional annotation of the SPMS-specific group, on the other hand, suggested the involvement of myeloid cells and functions such as NO signaling, metabolism and phagocytosis, adding to the increasing evidence of the involvement of these mechanisms in the disease progression (51). However, the most surprising finding is a significant enrichment of pathways linked to neurological processes specifically in the SP group. We confirmed these distinctive changes in the SPMS stage on the pathway level in an independent cohort. The finding is interesting in light of the MS paradigm that proposes that exhaustion of CNS reserves, caused by inflammation, likely represent a breaking point to enter progressive stage of disease (51). Several of these genes have been linked to neurodegenerative processes but also severity and progression of MS, including *ASIC2* (Acid Sensing Ion Channel Subunit 2)(52), *CALB2* (Calbindin 2)(53), *CERK* (Ceramide Kinase)(54), *CR1* (Complement C3b/C4b Receptor 1)(55, 56), *CX3CR1* (Fractalkine Receptor)(57), *GRIN1* (Glutamate Ionotropic Receptor NMDA Type Subunit 1)(58), *LRP1* (LDL Receptor Related Protein 1)(59), *NTN1* (Netrin1)(60) and *TNFRSF1A* (TNF Receptor Superfamily Member 1A)(61). Interestingly, some of these neuronal genes have been shown to play key roles outside the CNS, particularly in immune cells. This is the case for example of the neurotransmitter glutamate signaling,

displaying DNA methylation changes at several genes encoding receptor (GluR) subunits (GRID2, GRIN1, GRM2) and downstream signaling molecules (e.g. ITPR2/3, PRKC/A, CAMK2D, PRKAR1B) in SP cluster. Compelling evidence has demonstrated that glutamate exerts potent effects on normal immune cells and in the context of MS, e.g. affecting T cells activation, adhesion and migration, either directly through GluRs expressed at the surface of immune cells (62, 63) or indirectly via glutamate-dependent pathways (64). Similarly, axonal guidance cues such as netrins (NTN1, NTNG1 genes in our cohort), recently found altered in sera of MS patients(65), have been shown to affect crucial cellular functions of both innate and adaptive immune cells(66-68). This neuronal pattern in immune cells suggests that processes occurring in the brain might imprint an overlapping molecular signature on the peripheral immune cells. Such brain signature can occur when the immune cells infiltrate the CNS, as seen in the case of stroke (69), or via unknown mechanisms, as suggested in other CNS pathologies (70, 71). Another explanation implies external factors causing overlapping signatures between the tissues. One possibility is that chronic inflammation in MS causes age acceleration, as low-grade inflammation is one of the factors suggested to cause aging (72). This process may result in overlapping molecular signatures between the tissues and lead to exhaustion of the CNS functions as suggested in other CNS abnormalities (73, 74) and disorders (75). While the functional relevance of DNA methylation changes remains to be studied, this is the first report of immune cells exhibiting a unique molecular signature indicative of processes in the brain during the progressive stage of disease.

Conclusion

We demonstrate that four distinct immune cell types from MS patients share functionally relevant DNA methylation changes compared to healthy individuals. Owing to a gain of discovery power, omicsNPC methodology allows detection of such changes in complex diseases and further enables the identification

485 of discrete changes in MS patients, in general, and the SPMS stage, in particular. The findings provide new

486 insights into the putative mechanisms underlying MS pathogenesis and progression.

Abbreviations

Bone mineral density (BMD)

Central nervous system (CNS)

Differentially methylated position (DMP)

Epstein Barr virus (EBV)

Genomic association testing (GAT)

Healthy controls (HC)

Ingenuity Pathway Analysis (IPA)

Major depressive disorder (MDD)

Multiple Sclerosis (MS)

Non-parametric combination (NPC)

Over representation analysis (ORA)

Peripheral blood mononuclear cell (PBMC)

Principal component analysis (PCA)

Relapsing-remitting Multiple Sclerosis (RRMS)

Random effects model (REML)

Secondary progressive Multiple Sclerosis (SPMS)

Single nucleotide polymorphism (SNP)

5-methylcytosine (5mC)

Declarations

Acknowledgements

The authors would like to thank Dr. Y. Liu and Dr. A. Feinberg for provision of a part of the raw data and Dr. H. Morikawa for his input during data analysis.

Funding

This study was supported by grants from the Swedish Research Council (MJ, FP, JT), the Swedish Association for Persons with Neurological Disabilities (MJ), the Swedish Brain Foundation (MJ, JT), the Stockholm County Council - ALF project (MJ, FP), AstraZeneca - AstraZeneca-Science for Life Laboratory collaboration (MJ, FP, JT), StratNeuro (JT), STATEGRA FP7 (JT), the European Research Council (FP/2007-2013) / ERC Grant Agreement n. 617393; CAUSALPATH (IT, VL) and Karolinska Institute's funds (MJ). L. Kular is supported by a fellowship from the Margaretha af Ugglas Foundation.

Availability of data and materials

The Illumina 450K array data from $CD4^+$ T cells, $CD8^+$ T cells, $CD14^+$ monocytes and whole blood are available in the Gene Expression Omnibus (GEO) database under accession number GSEXX (deposition pending), GSEXX (deposition pending), GSE43976 and GSE106648, respectively. The Illumina 450K array data from $CD19^+$ B cells will be made available from the corresponding author upon request.

Authors' contributions

EE, JT, DGC and MJ conceived and designed the study. EE analyzed the data. NK, VL, IT developed, provided statistical guidance and description of the original omicsNPC framework. EE, SJ and DGC further optimized the omicsNPC framework. EE developed and implemented RR analysis. EE, LK, SR, FP and MJ generated the data. FP recruited study subjects. EE, LK and MJ contributed data interpretation. EE, MJ and LK wrote the manuscript. All authors read, provided input and approved the final manuscript.

Ethics approval and consent to participate

All study subjects provided written informed consent and the study was approved by the regional ethics committee. The research in this study conformed to the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

FP has received research grants from Biogen, Genzyme, Merck KGaA and Novartis, and fees for serving as Chair of DMC in clinical trials with Parexel. Other authors declare that they have no competing interests.

538 **Table 1.** Characteristics of Multiple Sclerosis (MS) patients and healthy controls used for 450K methylation
539 analysis.

		CD4$^+$	**CD8$^+$**	**CD14$^+$**	**CD19$^+$**
Healthy controls	N (Female/Male)	11 (7/4)	14 (9/5)	13 (9/4)	10 (6/4)
	Mean age (range)	43 (28-62)	37 (20-65)	41 (28-62)	39 (28-60)
Relapsing-remitting MS	N (Female/Male)	12 (9/3)	10 (5/5)	10 (7/3)	12 (7/5)
	Mean age (range)	38 (26-57)	35 (26-44)	40 (29-57)	37 (26-57)
Secondary progressive MS	N (Female/Male)	8 (4/4)	4 (0/4)	12 (7/5)	5 (2/3)
	Mean age (range)	50 (35-63)	44 (38-50)	49 (35-60)	47 (35-56)

540

541 All four cell types were available from 11 individuals, while between any two cell types the overlap of
542 individuals ranged from the average of 45% between CD8$^+$ and CD14$^+$ cells to 79% between CD4$^+$ and CD14$^+$
543 cells (details can be found in Supplementary Table 1). Patients had not been treated within 6 months prior
544 to sample collection.

Figure 1

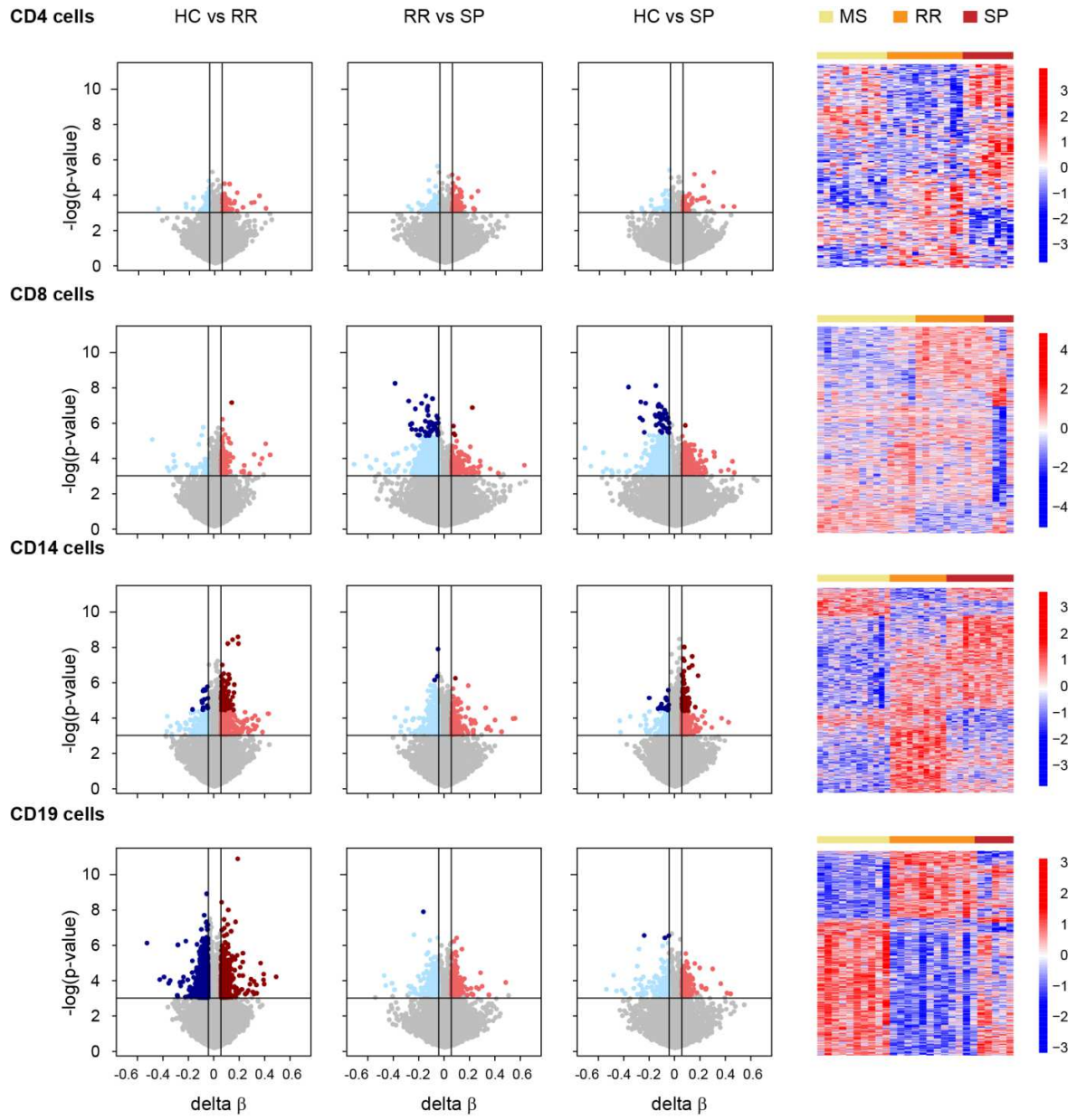


Figure 1. Methylation changes in cells sorted from peripheral blood of Multiple Sclerosis (MS) patients and healthy controls (HC). DNA methylation was measured using Illumina 450K arrays in $CD4^+$, $CD8^+$, $CD14^+$ and $CD19^+$ cells sorted from peripheral blood of untreated relapsing-remitting (RR) and secondary progressive (SP) MS patients and HC (details are provided in Table 1 and Table S1). **(A)** Volcano plots

illustrate differences in DNA methylation between RRMS, SPMS and HC. Hyper- and hypo-methylated CpGs with min 5% methylation change and p-value < 0.001 are indicated in light red and light blue, respectively, while darker red and darker blue indicate CpGs with min 5% methylation change and adj. p-value < 0.05. **(B)** Heat maps were generated using 1 000 most significant differentially methylated CpG sites between the conditions (the scale represents Z-score).

Figure 2

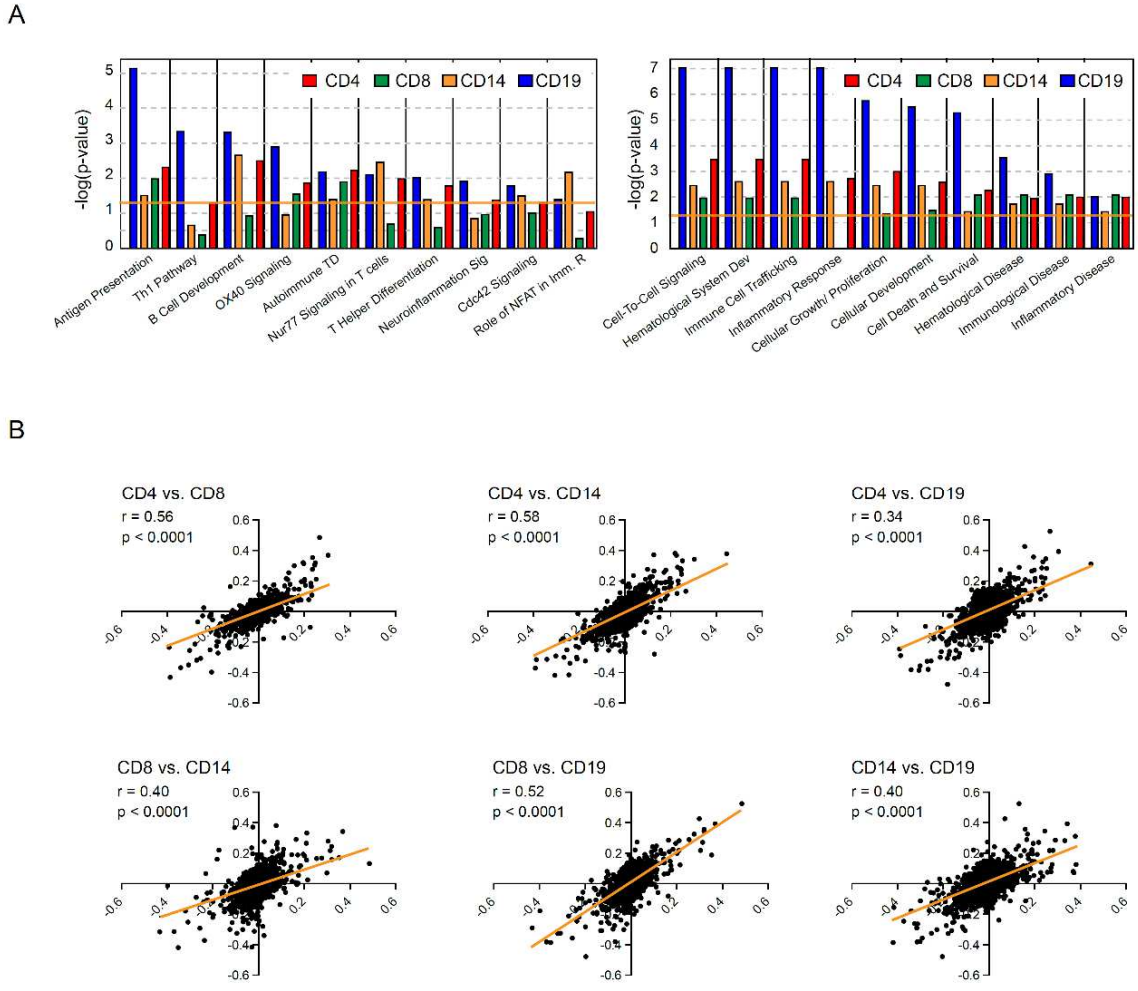
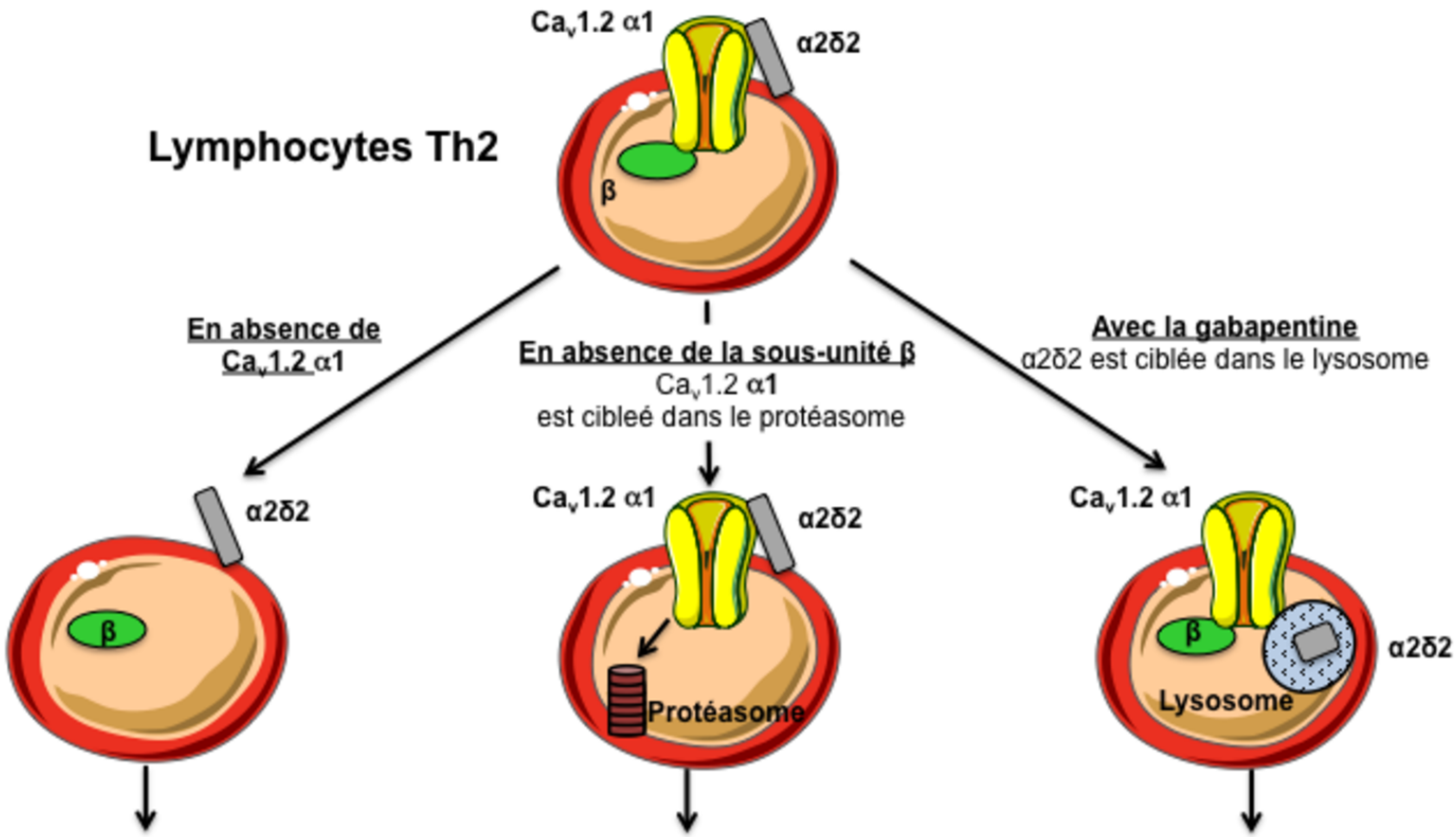


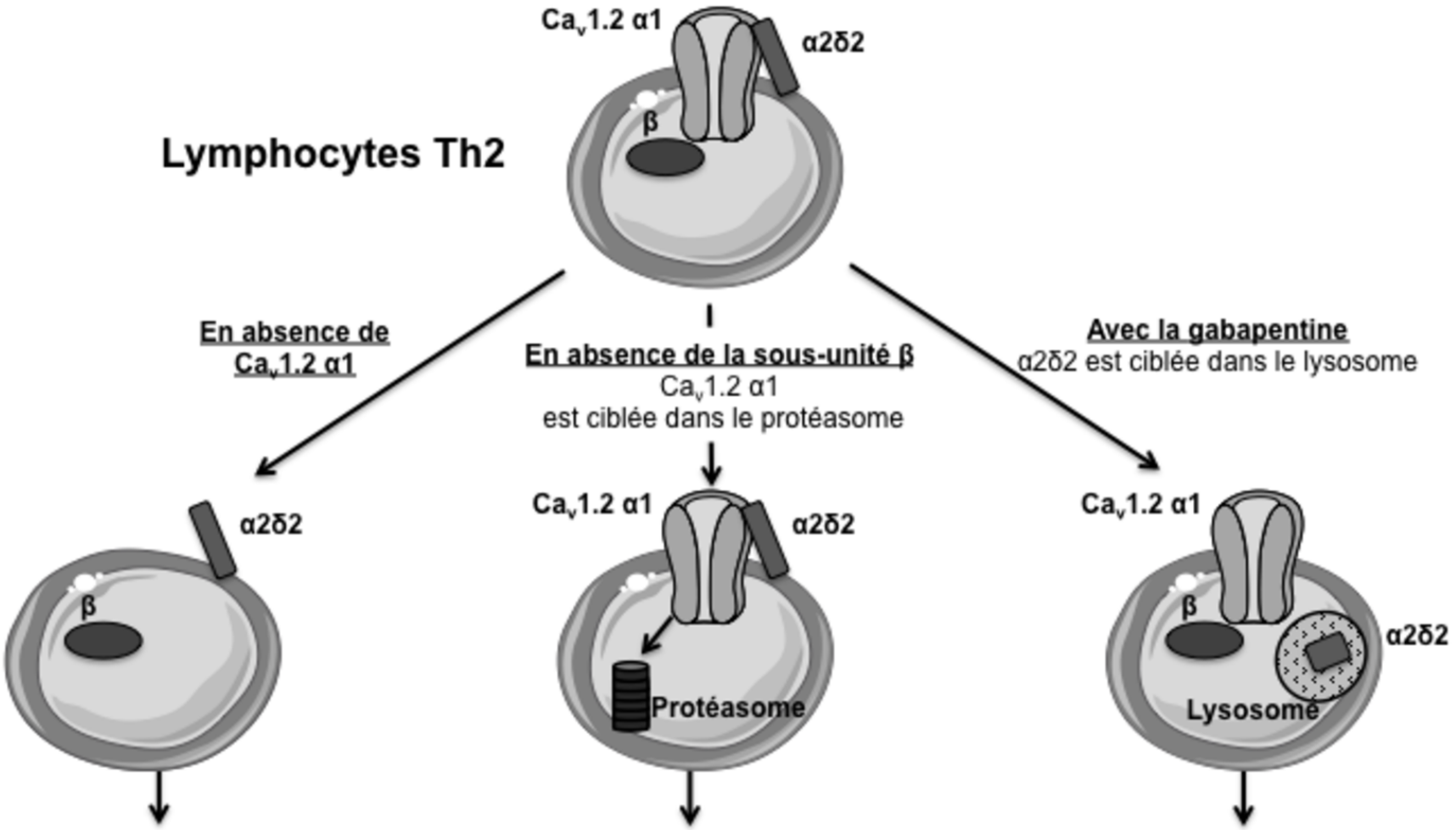
Figure 2. DNA methylation changes overlap between different cell types. (A) Selected canonical pathways and functional annotations from Ingenuity Pathway analysis (Fisher's p-value < 0.05) generated using genes associated with candidate differentially methylated positions between relapsing-remitting Multiple Sclerosis (RRMS) patients and healthy controls (HC) (absolute Δβ > 5%, p-value < 0.001) in each cell type separately. In total, 54, 87, 362, 1966 genes were used in analysis in $CD4^+$, $CD8^+$, $CD14^+$ and $CD19^+$ cells, respectively. **(B)** Correlation of effect sizes (Δβ for the RRMS-HC comparison) between cell types was

564 tested for all probes that displayed p-value < 0.001 in at least one comparison using the Spearman's rank

565 test.



- Diminution de la réponse calcique et de la production des cytokines induites par l'engagement du TCR dans les lymphocytes Th2 murins et humains, sans effet sur les lymphocytes Th1
- Inhibition de l'immunopathologie dépendante des lymphocytes Th2 au cours d'un asthme allergique



- Diminution de la réponse calcique et de la production des cytokines induites par l'engagement du TCR dans les lymphocytes Th2 murins et humains, sans effet sur les lymphocytes Th1
- Inhibition de l'immunopathologie dépendante des lymphocytes Th2 au cours d'un asthme allergique