

ORIGINAL ARTICLE

Assessment of the genetic contribution to brain magnetic resonance imaging lesion load and atrophy measures in multiple sclerosis patients

Ferdinando Clarelli¹ | Maria Assunta Rocca^{2,3,4} | Silvia Santoro¹ | Ermelinda De Meo^{2,3,4} | Laura Ferre^{1,2,4,5} | Melissa Sorosina¹  | Filippo Martinelli Boneschi^{6,7}  | Federica Esposito^{1,2,5}  | Massimo Filippi^{2,3,4,5,8} 

¹Laboratory of Human Genetics of Neurological Disorders, Institute of Experimental Neurology, IRCCS San Raffaele Scientific Institute, Milan, Italy

²Neurology Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy

³Neuroimaging Research Unit, Division of Neuroscience, IRCCS San Raffaele Scientific Institute, Milan, Italy

⁴Vita-Salute San Raffaele University, Milan, Italy

⁵Neurorehabilitation Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy

⁶Department of Pathophysiology and Transplantation, Dino Ferrari Centre, Neuroscience Section, University of Milan, Milan, Italy

⁷Neurology Unit and MS Centre, Foundation IRCCS Ca, Granda Ospedale Maggiore Policlinico, Milan, Italy

⁸Neurophysiology Service, IRCCS San Raffaele Scientific Institute, Milan, Italy

Correspondence

Massimo Filippi, Neurology Unit, IRCCS San Raffaele Scientific Institute, Via Olgettina 60, 20132 Milan, Italy.
Email: filippi.massimo@hsr.it

Abstract

Background and purpose: Multiple sclerosis (MS) susceptibility is influenced by genetics; however, little is known about genetic determinants of disease expression. We aimed at assessing genetic factors influencing quantitative neuroimaging measures in two cohorts of progressive MS (PMS) and relapsing–remitting MS (RRMS) patients.

Methods: Ninety-nine PMS and 214 RRMS patients underwent a 3-T brain magnetic resonance imaging (MRI) scan, with the measurement of five MRI metrics including T2 lesion volumes and measures of white matter, grey matter, deep grey matter, and hippocampal volumes. A candidate pathway strategy was adopted; gene set analysis was carried out to estimate cumulative contribution of genes to MRI phenotypes, adjusting for relevant confounders, followed by single nucleotide polymorphism (SNP) regression analysis.

Results: Seventeen Kyoto Encyclopedia of Genes and Genomes pathways and 42 Gene Ontology (GO) terms were tested. We additionally included in the analysis genes with enriched expression in brain cells. Gene set analysis revealed a differential pattern of association across the two cohorts, with processes related to sodium homeostasis being associated with grey matter volume in PMS ($p = 0.002$), whereas inflammatory-related GO terms such as adaptive immune response and regulation of inflammatory response appeared to be associated with T2 lesion volume in RRMS ($p = 0.004$ and $p = 0.008$, respectively). As for SNPs, the rs7104613^T mapping to *SPON1* gene was associated with reduced deep grey matter volume ($\beta = -0.731$, $p = 3.2 \times 10^{-7}$) in PMS, whereas we found evidence of association between white matter volume and rs740948^A mapping to *SEMA3A* gene ($\beta = 22.04$, $p = 5.5 \times 10^{-6}$) in RRMS.

Conclusions: Our data suggest a different pattern of associations between MRI metrics and functional processes across MS disease courses, suggesting different phenomena implicated in MS.

KEYWORDS

genetic association study, magnetic resonance imaging, MRI, multiple sclerosis

INTRODUCTION

Multiple sclerosis (MS) is a chronic disease of the central nervous system (CNS), characterized by inflammation, demyelination, and axonal/neuronal loss [1]. Its precise aetiology is unknown, although it is widely accepted that it implicates an interplay between genetic, environmental, and lifestyle factors. Genome-wide association studies have provided a significant contribution to the investigation of the MS genetic component, revealing a highly polygenic architecture, with an ever-increasing number of common single nucleotide polymorphisms (SNPs) associated with MS risk. The recent effort of the International Multiple Sclerosis Genetics Consortium (IMSGC) provided robust evidence for association of more than 200 loci [2]. Nonetheless, the majority of MS genetic studies have focused on susceptibility, and little is known about the genetic determinants influencing disease expression and severity. The investigation of quantitative endophenotypes in genetic association studies may represent an important step towards a more comprehensive assessment of the pathogenetic mechanisms underlying complex diseases [3], because they are more precise and believed to be closer to their biological substrates. Magnetic resonance imaging (MRI) measures have an essential role in MS diagnosis, and advances in the MRI field enable a finer characterization of MS expression. In this context, conventional imaging data such as T2-weighted measures, currently used in clinical practice, provide a global, rough estimation of the inflammatory–demyelinating burden of MS but are poor predictors of clinical status and disability accumulation; on the other hand, atrophy measures, which summarize neurodegenerative processes such as neuro-axonal loss, better correlate with clinical disability. In particular, grey matter (GM) atrophy [4] occurs early in the disease course and is more relevant than whole-brain atrophy in explaining and predicting physical disability and cognitive dysfunction [5–7]. In addition, an association between atrophy of selected GM areas and clinically related dysfunction has been demonstrated. For instance, hippocampal atrophy has been related to memory impairment [8]. Analysis of such measures may thus allow capturing a more comprehensive and advanced perspective of molecular contribution to MS expression and its relationship with disability. However, only a few studies have been conducted in this field [9–12].

Here, we explored the genetic influence on structural MRI measures in patients with progressive MS (PMS) and relapsing–remitting MS (RRMS), taking advantage of a well-characterized monocentric cohort of 320 patients enrolled at San Raffaele Hospital (OSR). We investigated five MRI measures, spanning from T2 lesion volume (T2LV) to white matter volume (WMV), GM volume (GMV), and regional GMV (deep GM and hippocampal volumes). Given the modest sample size, we adopted an expression quantitative trait locus (eQTL)-informed gene set approach to constrain analyses to plausible biological processes pointing to inflammation and neurodegeneration, which are the two pathological hallmarks of MS.

METHODS

Enrolled patients

A cohort of 320 Italian patients affected by MS according to McDonald criteria [13] was studied at OSR. Among them, 34 had primary progressive MS (PPMS), 67 had secondary progressive MS (SPMS), and 219 subjects had RRMS. For the purpose of the present study, PPMS and SPMS subjects were grouped into the so-called PMS cohort, as also suggested by the 2013 revision of the MS clinical courses [14]. The study was approved by OSR institutional ethics committee (protocol number: 107/INT/2018), and written informed consent was obtained from patients before study enrollment.

Image acquisition and analysis

Using a 3.0-T scanner (Intera, Philips Medical Systems), the following brain images were acquired from all subjects: (i) dual-echo turbo spin-echo; and (ii) three-dimensional (3D) T1-weighted fast field echo. Details on sequence geometry are provided in Methods S1.

Using a local thresholding segmentation technique (Jim 6, Xinapse Systems), T2-hyperintense lesion volumes were measured on T2-weighted scans. Normalized WMV and GMV were measured on 3D T1-weighted scans using SIENAX, after T1-hypointense lesion refilling [12].

Automatic segmentation of thalamus, caudate, putamen, pallidum, hippocampus, amygdala, and accumbens was performed on lesion-filled 3D T1-weighted scans using the FMRIB Integrated Registration and Segmentation Tool (<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FIRST>) software. The volume of these structures was multiplied by the head-normalization factor derived from SIENAX. The normalized volume of deep GM nuclei (the sum of the thalamus, caudate, putamen, pallidum, amygdala, and accumbens [DeepGMV]) was also obtained and analyzed together with hippocampus volume (HippV). The hippocampus was analyzed separately given its well-known involvement not only in neurodegenerative processes but also in synaptic plasticity and neurogenesis processes [15], which might potentially contribute to functional preservation and restoration.

Genotyping and imputation

The 320 MS patients had been genotyped on different Illumina array platforms available at the Laboratory of Human Genetics of Neurological Disorders at OSR: Infinium Omni2.5 ($n = 30$), Infinium OmniExpress ($n = 257$), and HumanOmni1-Quad ($n = 33$). Array-specific quality controls were performed using PLINK 1.9 [16] (Methods S1). The merged dataset was prephased with the SHAPEIT2 tool [17], and missing genotypes were imputed with

Minimac3 [18] on the 1000 Genomes Project Phase 3 ALL panel. Imputed markers were discarded if they were rare (minor allele frequency [MAF] < 1%) or of poor imputation quality ($r^2 < 0.3$); to limit the number of tests, we finally retained only variants genotyped on Illumina Infinium OmniExpress platform ($n = 696,061$) and the ones retrieved as cis-eQTL (see next paragraph).

Selection of candidate gene sets

Due to the modest sample size, we adopted a candidate pathway strategy, by jointly testing association of SNPs assigned to genes participating in a set.

We revised key topics in literature with specific focus on processes deemed relevant to neurodegeneration and neuroinflammation, and finally we mapped these concepts to gene sets derived from pathways resources. More precisely, pathways representing inflammatory processes such as T-cell differentiation and activation, B- and T-cell receptor signalling, nuclear factor κ B, tumor necrosis factor α , and chemokine signalling, cytokine production, antigen processing and presentation, and microglia activation were selected. As for the neurodegenerative domain, we examined key concepts related to synapses (glutamatergic and γ -aminobutyric acidergic synapse, synaptic transmission), mitochondrial dysfunction and oxidative stress, ion channel dysregulation, and homeostasis (calcium/sodium/potassium/iron/copper) [19–24]. The mentioned topics were mapped onto Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO; biological process hierarchy) pathway collections, derived from Molecular Signatures Database v6.2 (www.gsea-msigdb.org, release July 2018), with the constraint of a pathway size more than 10 and less than 300. These lower and upper limits were chosen to avoid excessive pathway testing and analysis of extremely broad sets.

Additionally, we investigated the contribution to MRI measures of genes with enriched expression in major brain cells, leveraging results from the study of McKenzie et al. [25], in which a meta-analysis of two human brain cell-specific signatures was conducted [26,27]. We hence retrieved the top 100 meta-analysed enriched genes in neurons, astrocytes, oligodendrocytes, microglia, and endothelial cells, as reported in McKenzie et al. [25]. As regarding assignment of SNPs to genes, we summarized information in genes by including a narrow flanking window of 2 kb, to minimize overlap between nearby genes. Moreover, we included in the analyses also SNPs that affect gene regulation in cis (cis-eQTL) in brain tissues and whole blood, according to GTEx v7 (<https://gtexportal.org/home/datasets>; see Methods S1).

Statistical analysis

We tested association between quantitative MRI metrics and the selected gene sets using a multilocus test of enrichment with a competitive test, which tests whether the genes in a gene set are

more strongly associated with MRI outcomes than all other genes. We used the procedure implemented in the multilocus test of enrichment (MTE) method [28] (Methods S1).

We then performed single-SNP analysis on markers assigned to genes annotated to the selected pathways by fitting linear regression models with PLINK 1.9 [16], assuming additive effects. For T2LV measurements, due to its skewed distribution, we used log-transformed values.

We considered sex, age, disease duration at MRI examination, and being under treatment (yes/no) as covariates. Variable selection with backward elimination and minimization of Akaike information criterion (AIC) was performed using the MASS R package, to identify covariates to be included for each metric, in both gene set and single-SNP analysis.

To account for population substructure, principal component analysis (PCA) was conducted with the Eigenstrat tool [29], estimating eigenvectors on a pruned subset of SNPs in low linkage disequilibrium (pairwise $r^2 < 0.2$ in a window of 100 SNPs). The first three eigenvectors were included.

To account for dependencies among SNPs, we adjusted for multiple testing by first deriving the number of independent tests M_{eff} and then calculating the correspondingly inflated Bonferroni adjustment by the M_{eff} factor, using the Genetic Error Calculator algorithm [30].

The lead SNPs for the five MRI metrics and their proxies ($r^2 > 0.8$) were then queried for functional and regulatory annotations with HaploReg v4.1 [31].

RESULTS

Gene set analysis

MRI data for 320 patients were available; of these, 313 were analyzed after removal of seven subjects who were estimated as outliers at PCA; the outliers did not show specific features in terms of ethnicity, clinico-demographic characteristics, and disease severity. Table 1 summarizes demographic, clinical, and MRI features for PMS ($n = 99$) and RRMS ($n = 214$) subjects. The correlogram in Figure S1 shows the expected inverse correlation between T2LV and atrophy measures and a moderate correlation between GMV and DeepGMV, whereas WMV was poorly correlated with GM volumetric data.

The final collection of investigated gene sets comprised 42 GO terms, 17 KEGG pathways, and five brain cell-specific sets of enriched genes; the complete collection with annotated genes is reported in gene matrix transposed format in Table S1. There was a remarkable heterogeneity in gene set size, in terms of both genes and assigned SNPs, with a median number of 60 annotated genes and 5170 SNPs per set (Figures S2 and S3).

We integrated the set of $n = 696,061$ SNPs with eSNPs from GTEx on the relevant tissues (Methods S1), yielding a total of 1,390,486 SNPs. Upon assignment of SNPs to the selected 64 candidate gene

TABLE 1 Demographic and clinical features and MRI metrics for the analyzed cohorts

Demographic and clinical features	PMS, <i>n</i> = 99	RRMS, <i>n</i> = 214
Gender, female:male, <i>n</i>	53:46	141:73
Age, years, mean (range)	50.2 (27.3–70.5)	38.6 (18.9–63.3)
Disease duration, years, mean (range)	17.7 (2.8–43)	11.1 (0–32)
EDSS, median (range)	6 (3.5–9)	1.8 (0–6)
Brain MRI metrics, ml, mean (range)		
T2LV	15.3 (0.5–62.2)	7.5 (0–62.8)
GMV	628.7 (433.9–765.6)	699.8 (438.9–854.7)
WMV	798.8 (645.0–970.9)	818.5 (591.1–950.9)
HippV	4.4 (2.7–5.9)	4.8 (2.9–7.1)
DeepGMV	3.7 (2.5–4.6)	4.0 (2.8–4.9)

Abbreviations: DeepGMV, deep grey matter volume; EDSS, Expanded Disability Status Scale; GMV, grey matter volume; HippV, hippocampus volume; MRI, magnetic resonance imaging; PMS, progressive multiple sclerosis; RRMS, relapsing–remitting multiple sclerosis; T2LV, T2-weighted lesion volume; WMV, white matter volume.

sets and selection of common variants (MAF > 5%), the total number of investigated markers was 189,400.

Gene set tests of enrichment were conducted with the MAGMA method, incorporating into models the three principal components and covariates detected in MRI outcome-specific AIC selection for the two courses (Figure S4). Age and disease duration variables were moderately correlated ($r < 0.48$) in the two cohorts.

A listing of the two top associated pathways and brain cell-specific sets is reported in Table 2. In both courses, it is apparent from the heatmap representation in Figure 1 that there is a clustering of association levels for GMV, HippV, and DeepGMV, whereas T2LV seems to display a pattern of association that is more related to WMV.

Twelve pathways were found to be nominally significant in the PMS cohort, with GO Cellular Sodium Ion Homeostasis being the top pathway associated with GMV ($p = 0.002$). It is noteworthy that this pathway was also associated with T2LV ($p = 0.006$) and DeepGMV ($p = 0.01$). The second most significant pathway in PMS was GO Cytokine Production Involved in Immune Response, which was associated with GMV ($p = 0.003$). In the RRMS cohort, 13 nominally significant pathways were identified, with GO Adaptive Immune Response being the top associated gene set related to T2LV ($p = 0.004$). Additionally, in RRMS we observed an association for processes related to inflammatory response (GO Regulation of Inflammatory Response with T2LV, $p = 0.008$) and apoptotic process (GO Regulation of Neuron Apoptotic Process with WMV, $p = 0.008$). Regarding the apoptotic process, it was involved also in PMS, with GO Neuron Apoptotic Process being associated with DeepGMV ($p = 0.027$).

TABLE 2 List of gene sets associated with the MRI metrics

MRI metric	Pathway	<i>p</i>
PMS		
T2LV	GO Cellular Sodium Ion Homeostasis	0.006
	GO Copper Ion Transport	0.013
WMV	GO Regulation of Neuron Projection Regeneration	0.015
	GO Regulation of Axonogenesis	0.020
GMV	GO Cellular Sodium Ion Homeostasis	0.002
	GO Cytokine Production Involved in Immune Response	0.003
DeepGMV	GO Cellular Sodium Ion Homeostasis	0.011
	GO Neuron Apoptotic Process	0.027
HippV	GO Response to Axon Injury	0.039
	GO Cytokine Production Involved in Immune Response	0.040
DeepGMV	Endothelial	0.035
RRMS		
T2LV	GO Adaptive Immune Response	0.004
	GO Regulation of Inflammatory Response	0.008
WMV	GO Regulation of Neuron Apoptotic Process	0.008
	GO Cellular Response to Reactive Nitrogen Species	0.011
GMV	GO Response to Axon Injury	0.038
	GO Microglial Cell Activation	0.042
HippV	KEGG Antigen Processing AND Presentation	0.020
	GO Central Nervous System Neuron Axonogenesis	0.032
WMV	Oligodendrocytes	0.008
DeepGMV	Microglia	0.017
T2LV		0.036
GMV		0.057

Note: The top two pathways for each MRI metric are reported. No pathways were deemed significant in DeepGMV for the RRMS cohort. Abbreviations: DeepGMV, deep grey matter volume; GMV, grey matter volume; GO, Gene Ontology; HippV, hippocampus volume; KEGG, Kyoto Encyclopedia of Genes and Genomes; MRI, magnetic resonance imaging; PMS, progressive multiple sclerosis; RRMS, relapsing–remitting multiple sclerosis; T2LV, T2-weighted lesion volume; WMV, white matter volume.

As for brain cell-specific enriched gene sets, in PMS the only set showing association was the one of endothelial-enriched genes in relation to DeepGMV ($p = 0.034$), whereas in RRMS the oligodendrocyte-specific gene set was found to be the most associated with WMV ($p = 0.008$). Of note, in this cohort microglia-specific genes showed mild association with three of five phenotypes ($p = 0.036$, $p = 0.056$, and $p = 0.017$ in T2LV, GMV, and DeepGMV, respectively); interestingly, also the GO term Microglial Cell Activation showed a trend of association with two MRI metrics in RRMS (T2LV $p = 0.053$ and GMV $p = 0.042$).

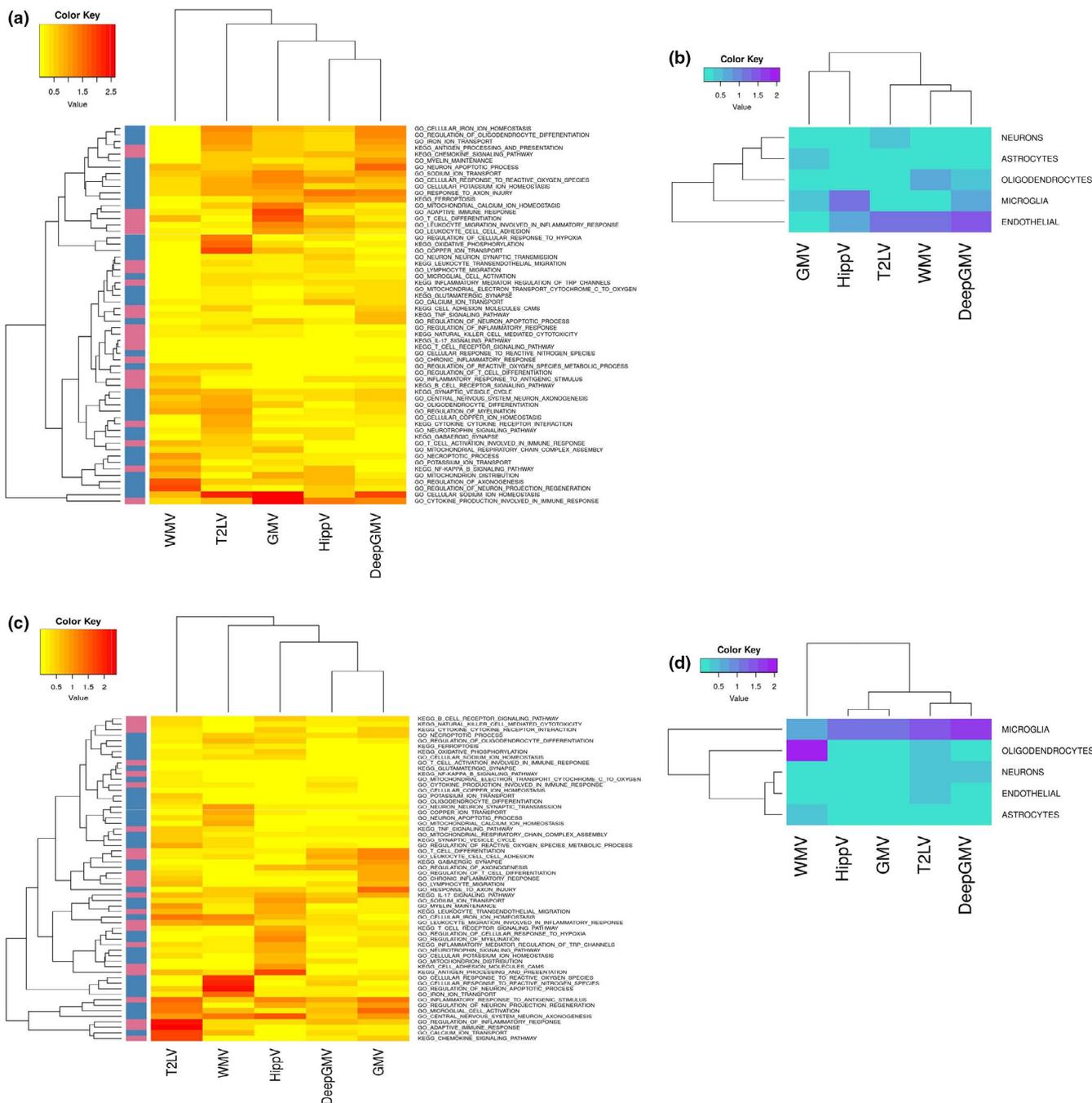


FIGURE 1 Heatmap representation of gene set analysis. Heatmaps illustrate the association level for the Kyoto Encyclopedia of Genes and Genomes pathway, Gene Ontology terms, and brain cell-specific genes from competitive tests with the multimarker analysis of genomic annotation procedure. The colormap representing level of association in terms of $-\log_{10}(p)$ is reported. Dendrograms representing hierarchical clustering with complete linkage of association p -values across the five magnetic resonance imaging metrics and the gene sets are reported. Blue indicates pathways deemed to represent neurodegenerative domain. Violet indicates pathways deemed to represent inflammatory and immune domains. (a, b) Pathways and brain cell-enriched genes for the progressive multiple sclerosis cohort. (c, d) Pathways and brain cell-enriched genes for the relapsing-remitting multiple sclerosis cohort. DeepGMV, deep grey matter volume; GMV, grey matter volume; HippV, hippocampus volume; T2LV, T2 lesion volume; WMV, white matter volume [Colour figure can be viewed at wileyonlinelibrary.com]

Single SNP analysis

In single SNPs association analyses, upon evaluation of the correlation pattern among SNPs [30], we estimated the effective number of independent tests $M_{\text{eff}} = 53,301$, with a “tagging ratio”

$R = M_{\text{eff}}/189,400=0.28$; this ratio is a measure of the extent to which the analyzed markers are nonredundant. We correspondingly set the Bonferroni threshold at $\alpha = 9.4 \times 10^{-7}$.

Also in the SNP-level analysis, we observed a differential pattern of association for the two cohorts. Top SNPs associated at

$p < 10^{-5}$ with at least one of the five MRI measures are reported in Table 3, with their clustering across the MRI measures plotted in Figure 2. Specifically, in PMS we found evidence of association for the rs7104613^T in gene *SPON1*, which was associated with a reduction of DeepGMV ($\beta = -0.73$, $p = 3.21 \times 10^{-7}$, Figure 3) and was the only marker withstanding multiple testing correction according to our estimated number of effective tests. For RRMS, clustering of phenotypes revealed hierarchically grouped associations for the GM volumetric measures driven by variants in *SEMA3A* and *GRIN2B* loci, with evidence of association between rs740948^A in *SEMA3A* and higher WMV ($\beta = 22.04$, $p = 5.5 \times 10^{-6}$), whereas the variant rs7970177^T in *GRIN2B* is associated with decreased volume of deep GM ($\beta = -0.29$, $p = 6.51 \times 10^{-6}$). Furthermore, we found an association between the SNP rs9837011^T in *KNCMB2* and higher T2LV ($\beta = 0.44$, $p = 4.707 \times 10^{-6}$).

We next evaluated in our data the major finding of a previous MRI genetics study in MS [11], involving a similarly sized cohort of 326 patients (262 RRMS, 64 SPMS) and investigating genetic

correlations of glutamate-related genes with seven MRI phenotypes related to brain and lesion volumes. The top-selected variant was rs3859123 in *GRIN2A*, whose G allele was reported to be associated with decreased brain volume and WMV. We confirmed the same direction of effect in our RRMS subjects, with a borderline association of rs3859123 with WMV ($\beta = -9.39$, $p = 0.09$).

DISCUSSION

In this study, we sought to correlate genetic variation derived from candidate pathways with structural MRI measures of inflammation and neurodegeneration, by separately analyzing PMS and RRMS courses to better delineate the contribution of genetics in these pathological processes.

In PMS, the gene set analysis pointed to the role of sodium homeostasis. The importance of sodium imbalance has been increasingly recognized in recent years as a possible contributor to MS pathogenesis

TABLE 3 List of SNPs associated with at least one MRI metric at $p < 10^{-5}$

	SNP	Chr	Position	A1	A2	MAF	Context	Mapped gene	Associated metric	Beta	p	Enhancer histone marks
PMS	rs7104613	11	14079931	T	C	0.051	Intronic	<i>SPON1</i>	DeepGMV	-0.73	3.21E-07	Cortex-derived primary cultured neurospheres; hippocampus, middle; substantia nigra; anterior caudate; cingulate gyrus; inferior temporal lobe; angular gyrus; dorsolateral prefrontal cortex; fetal brain, female; fetal brain, male
	rs681751	3	53458934	T	G	0.096	Intergenic	<i>PRKCD</i>	HippV	-0.69	9.96E-06	
RRMS	rs9837011	3	178267549	T	C	0.167	Intronic	<i>KCNMB2</i>	T2LV	0.44	4.70E-06	
	rs740948	7	83599335	A	G	0.411	Intronic	<i>SEMA3A</i>	WMV	22.04	5.50E-06	Fetal brain, female; fetal brain, male
	rs7970177	12	13738988	T	C	0.094	Intronic	<i>GRIN2B</i>	DeepGMV	-0.29	6.51E-06	Brain, cingulate gyrus; brain, inferior temporal lobe; fetal brain, male

Note: Position: evaluated on genome assembly GRCh37 (hg19). A1: minor allele, on which additive coding was based for regression models. A2: major allele. Enhancer histone marks: evaluated on HaploReg v4.1 tool.

Abbreviations: DeepGMV, deep grey matter volume; HippV, hippocampus volume; MAF, minor allele frequency; MRI, magnetic resonance imaging; PMS, progressive multiple sclerosis; RRMS, relapsing–remitting multiple sclerosis; SNP, single nucleotide polymorphism; T2LV, T2-weighted lesion volume; WMV, white matter volume.

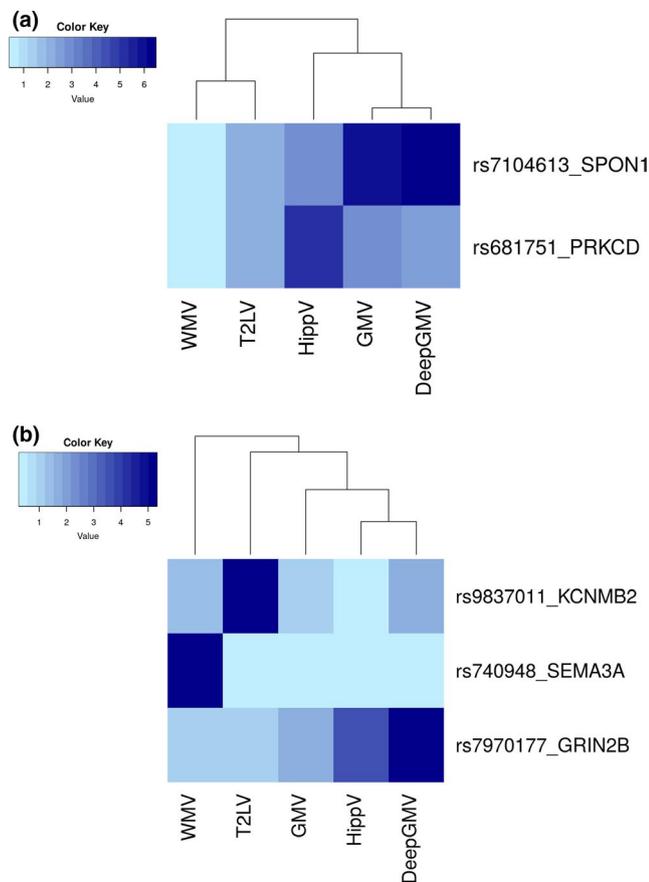


FIGURE 2 Top-associated single nucleotide polymorphisms (SNPs) in progressive multiple sclerosis (PMS) and relapsing–remitting multiple sclerosis (RRMS). The heatmaps illustrate the association level for SNPs detected as significant in at least one of the five magnetic resonance imaging (MRI) phenotypes at $p < 10^{-5}$. The colormap represents level of association in terms of $-\log_{10}(p)$. On the right, each SNP rs ID is reported with the correspondingly assigned gene symbol. A dendrogram representing hierarchical clustering with complete linkage of association p -values across the five MRI metrics is reported. (a) PMS cohort; (b) RRMS cohort. DeepGMV, deep grey matter volume; GMV, grey matter volume; HippV, hippocampus volume; T2LV, T2 lesion volume; WMV, white matter volume [Colour figure can be viewed at wileyonlinelibrary.com]

and to the mechanisms of axonal degeneration, which are primarily responsible for permanent neurological disability. Moreover, high sodium concentrations have been observed using imaging techniques in lesions, cortical GM, and deep GM [32], as a possible result of an increment of the cation in the intracellular space. Also, it has been found that in normal-appearing WM and deep GM sodium concentrations were higher in PMS subjects than in RRMS patients [33].

As for RRMS, we identified the term GO Adaptive Immune Response as the most associated gene set in relation to T2LV; the primary role of the adaptive arm of the immune system has long been recognized for MS pathogenesis [34], and has also been confirmed by the first large-scale international genome-wide association study [35]. Moreover, analyses of T-cell receptors showed that substantial fractions of CD4⁺ and CD8⁺ T cells were isolated from MS lesions [36].

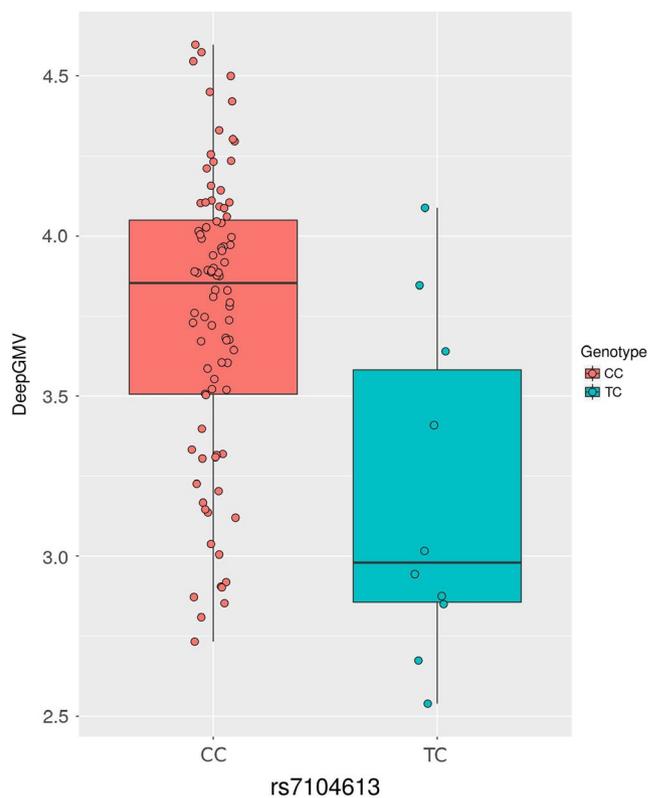


FIGURE 3 Distribution of deep grey matter volume (DeepGMV) values for rs7104613 alleles in SPON1 in the progressive multiple sclerosis (PMS) cohort. The boxplot illustrates the distribution of DeepGMV values for the wild type and heterozygous genotypes CC and TC for the single nucleotide polymorphism (SNP) rs7104613 in the PMS cohort. Lower and upper borders correspond to 25th and 75th percentiles. The upper (lower) whisker extends from the border to the largest (smallest) value no further than $1.5 \times$ interquartile range. The SNP is an intronic variant located in the SPON1 gene, with minor allele T having an allele frequency = 0.08 (1000 Genomes project). The variant was found to be associated in a linear regression model applied to DeepGMV in the PMS cohort at $p = 3.2 \times 10^{-7}$, $\beta = -0.73$ [Colour figure can be viewed at wileyonlinelibrary.com]

Finally, gene set analysis on brain cell-specific genes identified microglia-enriched genes as associated with multiple MRI measures in RRMS patients. Compelling evidence has been given for the implication of CNS-resident innate immune cells as mediators of the inflammatory response. Microglia secrete many proinflammatory and anti-inflammatory cytokines, and patterns of demyelination are associated with their activation, which in turn can induce the production of oxygen and nitric oxide radicals and impact on oxidative stress and on mitochondrial injury. In MS brains, microglia activation is not restricted to lesions but has also been detected in the normal-appearing WM and GM [37].

As regards single variant analysis, the most significant marker in the PMS cohort was rs7104613^T in SPON1, which overlaps with an active enhancer in multiple brain regions as of H3K4me1 marks and correlates with more pronounced atrophy in deep GM. The gene encodes for a cell adhesion protein that promotes the attachment of spinal cord and sensory neuron cells and the outgrowth of neurites

in vitro. This finding correlates well with the identification of *SPON1* in a genome-wide scan of brain connectome as a gene influencing dementia severity [38]. Another variant in *SPON1* emerged in a genetic study of Alzheimer disease patients, in which disease heterogeneity was mapped from volumetric T1-weighted MRI scans [39].

Among the most prominent signals in RRMS, we highlight two intronic variants: rs740948^A mapping to *SEMA3A* (WMV) and rs7970177^T in *GRIN2B* (DeepGMV). Both variants are located in an enhancer histone marks in fetal brain according to the regulatory chromatin states from DNase and histone ChIP-Seq as reported in HaploReg data. *SEMA3A*, a member of the semaphorin family, has already been implicated in MS [40] and encodes a protein critical for normal neuronal pattern development. Moreover, the expression of semaphorin 3A as well as of semaphorin 7A and their receptors has been detected by immunostaining in MS WM lesions [41]. *GRIN2B* encodes for the N-methyl-D-aspartate (NMDA) receptor subunit type 2B, playing a role in brain development, circuit formation, and synaptic plasticity. A structural MRI study revealed regional volumetric brain abnormalities of several GM areas and glutamate system genes in pediatric obsessive-compulsive disorder patients [42], whereas a combined effect in terms of genetic score was found to be correlated with prefrontal activity in a functional MRI study in healthy adults [43]. Interestingly, our data for the RRMS cohort corroborate the most relevant finding of a similar study [11] conducted on a cohort of comparable size, confirming direction of the association between WMV and an intronic variant located in an enhancer for *GRIN2A*, encoding for subunit 2A of NMDA glutamate receptor.

Of note, the abovementioned genetic variants and genes associated with MRI metrics in PMS and RRMS did not overlap with the more than 200 loci known to be associated with MS susceptibility, according to the last paper published by the IMSGC [2].

To our knowledge, this is the first study that investigates genetic contribution to these MRI measurements in MS, especially considering the assessment of selected GM regions. Previous efforts in the MS imaging-genetic field include the study of Kalincik et al. [9] that investigated the cumulative impact of known MS susceptibility loci, in the form of a weighted genetic risk score, on MRI volumetric parameters from patients with a clinically isolated syndrome. Another study [11] used a multivariate approach to estimate the impact of SNPs mapping to genes involved in glutamatergic neurotransmission on seven MRI phenotypes. Glutamate signalling was also the enriched process of subnetworks inferred from a genome-wide screen on brain cortical thickness [10].

For the purpose of the analyses, SPMS and PPMS subjects were combined together. This approach was chosen due to the limited sample size, but was also based on observations derived from natural history studies [44], which demonstrated that the two courses start at similar age and share a similar pace of disability accumulation, which is hypothesized to be driven by a unifying age-dependent neurodegenerative mechanism [45].

The most noticeable strength of our study is the availability of a comprehensive brain MRI characterization, which allowed us to conduct an investigation of MS radiological expression, using the same

scanner and segmentation protocol on an ethnically homogeneous cohort. We sought to reduce the amount of noise in the assignment of SNPs to genes by selecting a restricted flanking window [46] and including eSNPs in disease-relevant tissues. It is widely known that the majority of risk variants for complex traits are not located in coding regions [47], and it has been found that SNPs associated with complex traits as reported in the National Human Genome Research Institute GWAS Catalog are significantly more likely to be eQTL than frequency-matched SNPs [48]. Integrating knowledge generated from reference transcriptomic repositories like GTEx has proven successful, and transcriptome-wide association studies have been proposed [49,50].

The present findings have, however, to be considered also in light of some limitations, the most evident being the sample size, with impact on statistical power and positive predictive value [51]. The most powerful setting would be a multisite study; on the other hand, concern arises about differences in MRI hardware, scanner and software settings, and data-processing procedures. Lack of independent replication is another major drawback of our findings, mainly related to it being challenging to localize genotyped cohorts of MS patients with advanced neuroimaging characterization.

Being aware of the limited sample size, we strived to maximize the likelihood of biologically meaningful findings by leveraging previous biological knowledge on most likely pathways involved in neurodegenerative and inflammatory processes, thus likely increasing the prior odds and subsequently the likelihood of signals being true positive [52]. We adopted a gene set approach, which has been suggested as a powerful alternative [53]; because pathways can be dysregulated at multiple points, it can account for genetic heterogeneity and detect signals in situations in which many SNPs in a pathway might influence the outcome with individually small effect. We are aware that the candidate approach implies a level of arbitrariness in the process of selection, although we sought to manually curate the selection of key concepts derived from the most relevant reviews in the field. Another potential shortcoming of our study is the inclusion of treated patients, which could impact on MRI measurements. We did not exclude these subjects due to sample size reasons, and we sought to partially overcome this issue by including a dichotomous treatment variable into models.

There are now increasing efforts to move MS genetic research from susceptibility and dichotomous phenotypes to more precise quantitative endophenotypes. Validation of the current findings in independent cohorts is warranted, typically within a multicenter prospective effort, aimed to assess the impact of genetic variation on imaging phenotype early in MS history.

ACKNOWLEDGEMENTS

The authors would like to thank all patients for their participation to the study and the nursing staff for their cooperation in collecting patients' blood.

CONFLICT OF INTEREST

M.A.R. received speaker's honoraria from Bayer, Biogen, Bristol Myers Squibb, Celgene, Genzyme, Merck Serono, Novartis, Roche,

and Teva, and receives research support from the MS Society of Canada and Fondazione Italiana Sclerosi Multipla. F.M.B. has received compensation for consulting services and/or speaking activities from Teva Pharmaceutical Industries, Sanofi Genzyme, Merck Serono, Biogen Idec, Roche, Medday, and Excemed, and has received research support from Merck, Teva Pharmaceutical Industries, Italian Ministry of Health, Fondazione Italiana Sclerosi Multipla, and Fondazione Cariplo. F.E. has received compensation for consulting services and/or speaking activities from Novartis, Sanofi Genzyme, Ammiral, and Merck Serono. M.F. is Editor-in-Chief of the *Journal of Neurology* and Associate Editor of *Human Brain Mapping*; received compensation for consulting services and/or speaking activities from Ammiral, Alexion, Bayer, Biogen, Celgene, Eli Lilly, Genzyme, Merck-Serono, Novartis, Roche, Sanofi, Takeda, and Teva Pharmaceutical Industries; and receives research support from Biogen Idec, Merck-Serono, Novartis, Roche, Teva Pharmaceutical Industries, Italian Ministry of Health, Fondazione Italiana Sclerosi Multipla, and ARISLA (Fondazione Italiana di Ricerca per la SLA). The other authors have nothing to disclose.

AUTHOR CONTRIBUTIONS

Ferdinando Clarelli: conceptualization (equal), data curation (equal), formal analysis (equal), methodology (equal), writing—original draft (equal). **Maria Assunta Rocca:** conceptualization (equal), investigation (equal), supervision (equal), writing—review & editing (equal). **Silvia Santoro:** data curation (equal), investigation (equal), writing—review & editing (equal). **Ermelinda De Meo:** investigation (equal), writing—review & editing (equal). **Laura Ferrè:** investigation (equal), writing—review & editing (equal). **Melissa Sorosina:** investigation (equal), writing—review & editing (equal). **Filippo Martinelli Boneschi:** investigation (equal), writing—review & editing (equal). **Federica Esposito:** conceptualization (equal), investigation (equal), supervision (equal), writing—review & editing (equal). **Massimo Filippi:** conceptualization (equal), supervision (equal), writing—review & editing (equal).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Melissa Sorosina  <https://orcid.org/0000-0003-4376-2550>

Filippo Martinelli Boneschi  <https://orcid.org/0000-0002-9955-1368>

Federica Esposito  <https://orcid.org/0000-0003-2550-0805>

Massimo Filippi  <https://orcid.org/0000-0002-5485-0479>

REFERENCES

- Filippi M, Bar-Or A, Piehl F, et al. Multiple sclerosis. *Nat Rev Dis Primers*. 2018;4(1):43. <https://doi.org/10.1038/s41572-018-0041-4>
- International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science*. 2019;365(6460):aav7188. <https://doi.org/10.1126/science.aav7188>
- Glahn DC, Thompson PM, Blangero J. Neuroimaging endophenotypes: strategies for finding genes influencing brain structure and function. *Hum Brain Mapp*. 2007;28(6):488-501. <https://doi.org/10.1002/hbm.20401>
- Calabrese M, Poretto V, Favaretto A, et al. Cortical lesion load associates with progression of disability in multiple sclerosis. *Brain*. 2012;135(Pt 10):2952-2961. <https://doi.org/10.1093/brain/aws246>
- Filippi M, Preziosa P, Copetti M, et al. Gray matter damage predicts the accumulation of disability 13 years later in MS. *Neurology*. 2013;81(20):1759-1767. <https://doi.org/10.1212/01.wnl.0000435551.90824.d0>
- Damjanovic D, Valsasina P, Rocca MA, et al. Hippocampal and deep gray matter nuclei atrophy is relevant for explaining cognitive impairment in MS: a multicenter study. *Am J Neuroradiol*. 2017;38(1):18-24. <https://doi.org/10.3174/ajnr.A4952>
- De Stefano N, Airas L, Grigoriadis N, et al. Clinical relevance of brain volume measures in multiple sclerosis. *CNS Drugs*. 2014;28(2):147-156. <https://doi.org/10.1007/s40263-014-0140-z>
- Sicotte NL, Kern KC, Giesser BS, et al. Regional hippocampal atrophy in multiple sclerosis. *Brain*. 2008;131(Pt 4):1134-1141. <https://doi.org/10.1093/brain/awn030>
- Kalincik T, Guttman CR, Krasensky J, et al. Multiple sclerosis susceptibility loci do not alter clinical and MRI outcomes in clinically isolated syndrome. *Genes Immun*. 2013;14(4):244-248. <https://doi.org/10.1038/gene.2013.17>
- Matsushita T, Madireddy L, Sprenger T, et al. Genetic associations with brain cortical thickness in multiple sclerosis. *Genes Brain Behav*. 2015;14(2):217-227. <https://doi.org/10.1111/gbb.12190>
- Strijbis EM, Inkster B, Vounou M, et al. Glutamate gene polymorphisms predict brain volumes in multiple sclerosis. *Mult Scler*. 2013;19(3):281-288. <https://doi.org/10.1177/1352458512454345>
- Gourraud PA, Sdika M, Khankhanian P, et al. A genome-wide association study of brain lesion distribution in multiple sclerosis. *Brain*. 2013;136(Pt 4):1012-1024. <https://doi.org/10.1093/brain/aws363>
- Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 Revisions to the McDonald criteria. *Ann Neurol*. 2011;69(2):292-302. <https://doi.org/10.1002/ana.22366>
- Lublin FD, Reingold SC, Cohen JA, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology*. 2014;83(3):278-286. <https://doi.org/10.1212/WNL.0000000000000560>
- Rocca MA, Barkhof F, De Luca J, et al. The hippocampus in multiple sclerosis. *Lancet Neurol*. 2018;17(10):918-926. [https://doi.org/10.1016/S1474-4422\(18\)30309-0](https://doi.org/10.1016/S1474-4422(18)30309-0)
- Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7. <https://doi.org/10.1186/s13742-015-0047-847>
- Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods*. 2013;10(1):5-6. <https://doi.org/10.1038/nmeth.2307>
- Das S, Forer L, Schonherr S, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;48(10):1284-1287. <https://doi.org/10.1038/ng.3656>
- Mahad DH, Trapp BD, Lassmann H. Pathological mechanisms in progressive multiple sclerosis. *Lancet Neurol*. 2015;14(2):183-193. [https://doi.org/10.1016/S1474-4422\(14\)70256-X](https://doi.org/10.1016/S1474-4422(14)70256-X)
- Friese MA, Schattling B, Fugger L. Mechanisms of neurodegeneration and axonal dysfunction in multiple sclerosis. *Nat Rev Neurol*. 2014;10(4):225-238. <https://doi.org/10.1038/nrneuro.2014.37>
- Campbell G, Mahad DJ. Mitochondrial dysfunction and axon degeneration in progressive multiple sclerosis. *FEBS Lett*. 2018;592(7):1113-1121. <https://doi.org/10.1002/1873-3468.13013>
- Louapre C, Lubetzki C. Neurodegeneration in multiple sclerosis is a process separate from inflammation: Yes. *Mult Scler*.

- 2015;21(13):1626-1628. <https://doi.org/10.1177/135245851558759821/13/1626>
23. Hutchinson M. Neurodegeneration in multiple sclerosis is a process separate from inflammation: No. *Mult Scler*. 2015;21(13):1628-1631. <https://doi.org/10.1177/135245851561224421/13/1628>
 24. Stys PK, Zamponi GW, van Minnen J, Geurts JJ. Will the real multiple sclerosis please stand up? *Nat Rev Neurosci*. 2012;13(7):507-514. <https://doi.org/10.1038/nrn3275>
 25. McKenzie AT, Wang M, Hauberg ME, et al. Brain cell type specific gene expression and co-expression network architectures. *Sci Rep*. 2018;8(1):8868. <https://doi.org/10.1038/s41598-018-27293-510.1038>
 26. Darmanis S, Sloan SA, Zhang Y, et al. A survey of human brain transcriptome diversity at the single cell level. *Proc Natl Acad Sci U S A*. 2015;112(23):7285-7290. <https://doi.org/10.1073/pnas.1507125112>
 27. Zhang Y, Chen K, Sloan SA, et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J Neurosci*. 2014;34(36):11929-11947. <https://doi.org/10.1523/JNEUROSCI.1860-14.2014>
 28. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol*. 2015;11(4):e1004219. <https://doi.org/10.1371/journal.pcbi.1004219>
 29. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38(8):904-909. <https://doi.org/10.1038/ng1847>
 30. Li MX, Yeung JM, Cherny SS, Sham PC. Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum Genet*. 2012;131(5):747-756. <https://doi.org/10.1007/s00439-011-1118-2>
 31. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res*. 2012;40:D930-D934. <https://doi.org/10.1093/nar/gkr917>
 32. Ciccarelli O, Barkhof F, Bodini B, et al. Pathogenesis of multiple sclerosis: insights from molecular and metabolic imaging. *Lancet Neurol*. 2014;13(8):807-822. [https://doi.org/10.1016/S1474-4422\(14\)70101-2](https://doi.org/10.1016/S1474-4422(14)70101-2)
 33. Paling D, Solanky BS, Riemer F, et al. Sodium accumulation is associated with disability and a progressive course in multiple sclerosis. *Brain*. 2013;136(Pt 7):2305-2317. <https://doi.org/10.1093/brain/awt149>
 34. Lassmann H, van Horsen J, Mahad D. Progressive multiple sclerosis: pathology and pathogenesis. *Nat Rev Neurol*. 2012;8(11):647-656. <https://doi.org/10.1038/nrneuro.2012.168>
 35. Sawcer S, Hellenthal G, Pirinen M, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature*. 2011;476(7359):214-219. <https://doi.org/10.1038/nature10251>
 36. Hemmer B, Kerschensteiner M, Korn T. Role of the innate and adaptive immune responses in the course of multiple sclerosis. *Lancet Neurol*. 2015;14(4):406-419. [https://doi.org/10.1016/S1474-4422\(14\)70305-9](https://doi.org/10.1016/S1474-4422(14)70305-9)
 37. Lassmann H. Mechanisms of white matter damage in multiple sclerosis. *Glia*. 2014;62(11):1816-1830. <https://doi.org/10.1002/glia.22597>
 38. Jahanshad N, Rajagopalan P, Hua X, et al. Genome-wide scan of healthy human connectome discovers SPON1 gene variant influencing dementia severity. *Proc Natl Acad Sci U S A*. 2013;110(12):4768-4773. <https://doi.org/10.1073/pnas.12162061101216206110>
 39. Varol E, Sotiras A, Davatzikos C. HYDRA: Revealing heterogeneity of imaging and genetic patterns through a multiple max-margin discriminative analysis framework. *NeuroImage*. 2017;145(Pt B):346-364. <https://doi.org/10.1016/j.neuroimage.2016.02.041>
 40. Farrokhi M. Sema3A and multiple sclerosis. *Gene*. 2017;615:41. <https://doi.org/10.1016/j.gene.2017.03.020>
 41. Costa C, Martinez-Saez E, Gutierrez-Franco A, et al. Expression of semaphorin 3A, semaphorin 7A and their receptors in multiple sclerosis lesions. *Mult Scler*. 2015;21(13):1632-1643. <https://doi.org/10.1177/1352458515599848>
 42. Wu K, Hanna GL, Easter P, Kennedy JL, Rosenberg DR, Arnold PD. Glutamate system genes and brain volume alterations in pediatric obsessive-compulsive disorder: a preliminary study. *Psychiatry Res*. 2013;211(3):214-220. <https://doi.org/10.1016/j.psycyhres.2012.07.003>
 43. Pergola G, Di Carlo P, Andriola I, et al. Combined effect of genetic variants in the GRIN2B coding gene (GRIN2B) on prefrontal function during working memory performance. *Psychol Med*. 2016;46(6):1135-1150. <https://doi.org/10.1017/S0033291715002639>
 44. Confavreux C, Vukusic S. Natural history of multiple sclerosis: a unifying concept. *Brain*. 2006;129(Pt 3):606-616. <https://doi.org/10.1093/brain/awl007>
 45. Scalfari A, Neuhaus A, Daumer M, Ebers GC, Muraro PA. Age and disability accumulation in multiple sclerosis. *Neurology*. 2011;77(13):1246-1252. <https://doi.org/10.1212/WNL.0b013e318230a17d>
 46. Veyrieras JB, Kudravalli S, Kim SY, et al. High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS Genet*. 2008;4(10):e1000214. <https://doi.org/10.1371/journal.pgen.1000214>
 47. Maurano MT, Humbert R, Rynes E, et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science*. 2012;337(6099):1190-1195. <https://doi.org/10.1126/science.1222794>
 48. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet*. 2010;6(4):e1000888. <https://doi.org/10.1371/journal.pgen.1000888>
 49. Gamazon ER, Wheeler HE, Shah KP, et al. A gene-based association method for mapping traits using reference transcriptome data. *Nat Genet*. 2015;47(9):1091-1098. <https://doi.org/10.1038/ng.3367>
 50. Wu C, Pan W. Integration of enhancer-promoter interactions with GWAS summary results identifies novel schizophrenia-associated genes and pathways. *Genetics*. 2018;209(3):699-709. <https://doi.org/10.1534/genetics.118.300805>
 51. Button KS, Ioannidis JP, Mokrysz C, et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*. 2013;14(5):365-376. <https://doi.org/10.1038/nrn3475>
 52. Sawcer S. The complex genetics of multiple sclerosis: pitfalls and prospects. *Brain*. 2008;131(Pt 12):3118-3131. <https://doi.org/10.1093/brain/awn081>
 53. Wang L, Jia P, Wolfinger RD, Chen X, Zhao Z. Gene set analysis of genome-wide association studies: methodological issues and perspectives. *Genomics*. 2011;98(1):1-8. <https://doi.org/10.1016/j.ygeno.2011.04.006>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Clarelli F, Rocca MA, Santoro S, et al. Assessment of the genetic contribution to brain magnetic resonance imaging lesion load and atrophy measures in multiple sclerosis patients. *Eur J Neurol*. 2021;28:2513-2522. <https://doi.org/10.1111/ene.14872>