Title
Genetic associations with brain cortical thickness in multiple sclerosis

Authors
Matsushita, T
Madireddy, L
Sprenger, T
et al.

Publication Date
2015

DOI
10.1111/gbb.12190

Peer reviewed

Genes, Brain and Behavior
Official publication of the International Behavioural and Neural Genetics Society

Genes, Brain and Behavior (2015) ***14:*** *217–227* doi: 10.1111/gbb.12190

Genetic associations with brain cortical thickness in multiple sclerosis

T. Matsushita[†], L. Madireddy[†], T. Sprenger[‡], P. Khankhanian[†], S. Magon[‡], Y. Naegelin[‡], E. Caverzasi[†], R. L. P. Lindberg[‡], L. Kappos[‡], S. L. Hauser[†], J. R. Oksenberg[†], R. Henry[†], D. Pelletier[§] and S. E. Baranzini[*,†]

[†]Department of Neurology, University of California, San Francisco, CA, USA , [‡]Department of Neurology, University Hospital, Basel, Switzerland , and [§]Department of Neurology, Yale University, New Haven, CT, USA
**Corresponding author: S. E. Baranzini, PhD, Department of Neurology, University of California, San Francisco, 675 Nelson Rising Lane, Suite NS-215, San Francisco, CA, USA. E-mail: sergio.baranzini@ucsf.edu*

Multiple sclerosis (MS) is characterized by temporal and spatial dissemination of demyelinating lesions in the central nervous system. Associated neurodegenerative changes contributing to disability have been recognized even at early disease stages. Recent studies show the importance of gray matter damage for the accrual of clinical disability rather than white matter where demyelination is easily visualized by magnetic resonance imaging (MRI). The susceptibility to MS is influenced by genetic risk, but genetic factors associated with the disability are not known. We used MRI data to determine cortical thickness in 557 MS cases and 75 controls and in another cohort of 219 cases. We identified nine areas showing different thickness between cases and controls (regions of interest, ROI) (eight of them were negatively correlated with Kurtzke's expanded disability status scale, EDSS) and conducted genome-wide association studies (GWAS) in 464 and 211 cases available from the two data sets. No marker exceeded genome-wide significance in the discovery cohort. We next combined nominal statistical evidence of association with physical evidence of interaction from a curated human protein interaction network, and searched for subnetworks enriched with nominally associated genes and for commonalities between the two data sets. This network-based pathway analysis of GWAS detected gene sets involved in glutamate signaling, neural development and an adjustment of intracellular calcium concentration. We report here for the first time gene sets associated with cortical thinning of MS. These genes are potentially correlated with disability of MS.

Keywords: Brain, cortical thickness, data integration, genetics, GWAS, MRI, multiple sclerosis, networks, protein interactome, SNP

Received 13 October 2014, revised 25 November 2014, accepted for publication 27 November 2014

Multiple sclerosis (MS) (OMIM: 126200) is the most prevalent cause of non-traumatic neurologic disability in young adults in Europe and North America (Trapp & Nave 2008), and is characterized by temporal and spatial dissemination of demyelinating lesions in the central nervous system (CNS) (Hauser & Oksenberg 2006). Most patients with MS have a relapsing remitting course with progressive neurological disability being common at the late stages of the disease. Recent studies, however, show that the neurodegenerative changes contributing to the sustained clinical deficits manifest even at an early phase in the natural history of the disease.

On the basis of demyelination patterns found by magnetic resonance imaging (MRI) and post-mortem studies, searches for the etiology of MS have been largely confined to the white matter (WM) (Noseworthy *et al.* 2000). However, advances in neuroimaging and neuropathological techniques have enabled the localization of damage to the gray matter (GM) as well, manifested as both focal demyelination and more generalized cortical tissue thinning (Bo *et al.* 2003; Calabrese *et al.* 2007; Geurts *et al.* 2005; Kutzelnigg *et al.* 2005; Nakamura & Fisher 2009; Patenaude *et al.* 2011). Furthermore, both cortical lesions and volume loss appear to occur already at early stages of MS (Calabrese *et al.* 2007; Chard *et al.* 2002; Dalton *et al.* 2004; De Stefano *et al.* 2003; Giorgio *et al.* 2011; Henry *et al.* 2008; Tiberio *et al.* 2005) and are more correlated with physical disability and cognitive impairment than WM lesion load (Calabrese *et al.* 2010b; Chen *et al.* 2004; Fisniku *et al.* 2008). Finally, GM damage could serve as a potential prognostic factor for disease progression (Calabrese *et al.* 2010b, 2011; Filippi *et al.* 2010). Altogether, these findings suggest that identifying the causes of GM pathology is a valid strategy to advance our understanding of the neurodegenerative processes underlying MS.

Susceptibility to MS is thought to be influenced by both environmental and genetic factors (Hauser & Oksenberg 2006). In addition to the well-documented effect of certain human leukocyte antigen (HLA) alleles, genome-wide association studies (GWAS) carried out on large sample sizes have shown that DNA polymorphisms at more than 100 non-HLA loci influence MS risk (Beecham *et al.* 2013; IMSGC *et al.* 2011; Patsopoulos *et al.* 2011). However, with the notable exception of the HLA system, genetic factors associated with clinical phenotype or quantitative neuroimaging outcomes have not been yet identified (Beecham *et al.* 2013; IMSGC *et al.* 2011; Okuda *et al.* 2009; Tur *et al.* 2014; Yates *et al.*

Table 1: Demographic features

	UCSF		Basel	*P*-value	
	Case (557)	Control (75)	Case (218)	Case vs. control in UCSF	UCSF case vs. Basel case
Female (%)	68.8	65.3	71.9	0.6	0.55
Age (mean ± SD)	43.1 ± 9.8	40.6 ± 10.4	43.7 ± 10.9	0.15	0.5
Onset age (mean ± SD)	34.6 ± 9.4	N/A	32.2 ± 9.6	N/A	0.0019
Disease duration	8.5 ± 8.8	N/A	11.7 ± 8.9	N/A	2.80E-08
Disease course (CIS/RR/SP/PP/PR)	90/398/49/18/2	N/A	9/148/38/10/8	N/A	6.40E-06
EDSS (mean ± SD)	1.9 ± 1.6	N/A	3.0 ± 1.7	N/A	<2.2E-16

2014). On the other hand, brain structure is also influenced by genetic factors. For example, comparison of monozygotic and dizygotic twins highlighted the genetic influence on cortical thickness and even with standardized measures of intellectual ability (Eyler *et al.* 2011, 2012; Joshi *et al.* 2011; Kochunov *et al.* 2011; Kremen *et al.* 2010; Rimol *et al.* 2010). One study reported that siblings with MS had similar patterns of cortical thinning and lesion numbers (Calabrese *et al.* 2012). Although small and not yet replicated, these observations suggest that genetic factors may influence GM pathology in MS.

The aim of this study is to identify genes correlated with GM pathology measured as brain cortical thinning in two large MS cohorts. We conducted a GWAS using cortical thickness as the outcome variable in both groups independently and conducted a pathway analysis using protein interaction networks (PINs). Loci identified through this effort will shed light on a novel aspect of genetic influence on MS, which may have an effect on the clinical course.

Materials and methods

Subjects

A total of 851 subjects evaluated at the Multiple Sclerosis Center of the University of California, San Francisco (referred to as UCSF) ($n = 557$ cases and 75 healthy controls) and at the University Hospital Basel, Switzerland (referred to as Basel) ($n = 219$) were recruited to participate in a prospective study of phenotype–genotype biomarker associations using identical diagnostic and inclusion/exclusion criteria (Baranzini *et al.* 2009b). Individuals affected with all disease types were included: clinically isolated syndrome (CIS), relapsing remitting MS (RRMS), secondary progressive MS (SPMS), primary progressive MS (PPMS) and progressive relapsing MS (PRMS) (Table 1). CIS was defined as the first well-established neurological event lasting ≥48 h, involving the optic nerve, spinal cord, brainstem or cerebellum. In CIS patients, the presence of two or more hyperintense lesions on a T2-weighted (T2w) MRI sequence was also required for enrollment into the study. The diagnosis of RRMS was made utilizing the revised McDonald's criteria (Polman *et al.* 2005). Secondary progressive MS was defined by 6 months of worsening neurological disability not explained by clinical relapse. Primary progressive MS was defined both by progressive clinical worsening for more than 12 months from symptom onset without any relapses, and by abnormal cerebrospinal fluid as defined by the presence of ≥2 oligoclonal bands or an elevated IgG index. Individuals were excluded if they had experienced a clinical relapse or received treatment with glucocorticosteroids within the previous month of enrollment. The concomitant use of disease modifying therapies for MS was permitted. For all subjects, the expanded disability status scale (EDSS) score (Kurtzke 1983) was assessed. In this study, age of onset was defined as the first episode of focal neurological dysfunction suggestive of CNS demyelinating disease. This information was obtained via individual recall and verified through review of medical records.

The control group consisted of genetically unrelated individuals, primarily spouses/partners, friends and other volunteers. Control subjects were of northern European ancestry and matched as a group, proportionally with cases according to age (±5 years) and sex. A family history or current diagnosis of MS as well as a relation to another case or control subject in the cohort were sufficient reasons for exclusion. The protocols were approved by the Committee on Human Research in San Francisco and the ethics committee in Basel. Informed consent was obtained from all participants.

Image acquisition

Brain MRI scans were performed on all subjects upon entry into the study, and analyses were carried out without knowledge of disease subtype, duration or treatment history. Magnetic resonance imaging images at UCSF were acquired using an eight-channel phased array coil in reception and a body coil in transmission on a 3T GE Excite scanner (GE Healthcare Technologies, Waukesha, WI, USA). For the UCSF cohort, MRI examination included axial dual-echo spin echo sequences (TE at 20 and 90 milliseconds, TR = 2000 milliseconds, $512 \times 512 \times 44$ matrix, $240 \times 240 \times 132$ mm^3 FOV, slice thickness = 3 mm, interleaved). A high-resolution inversion recovery spoiled gradient-echo T1-weighted isotropic, volumetric sequence (3D IRSPGR, $1 \times 1 \times 1$ mm^3, 180 slices) was also performed (TE/TR/TI = 2/7/400 milliseconds, flip angle = 15°, $256 \times 256 \times 180$ matrix, $240 \times 240 \times 180$ mm^3 FOV, NEX = 1). In Basel, MRI scanning was performed using a 1.5 T Magnetom Avanto scanner (Siemens Medical Solutions, Erlangen, Germany). High-resolution T1-weighted MPRAGE images were acquired in sagittal plane (TR/TI/TE = 2080/1100/3.0 milliseconds; $\alpha = 15°$, 160 slices, isotropic voxel of $1 \times 1 \times 1$ mm). Axial 3 mm proton density-weighted (PDw) and T2w images were also acquired to determine the WM lesion load (double spin echo: TR/TE1/TE2 = 3980/14/108 milliseconds; 40 slices with an in-plane resolution of 1×1 mm).

Brain lesions were identified by medical doctors specializing in MS after analyzing high-resolution T1-weighted, T2w and PDw images. To create T1 lesion masks, lesions were outlined based on semi-automated thresholding with manual editing utilizing in-house software in San Francisco and using Amira 3.1.1 (Mercury Computer Systems Inc., Hillsboro, OR, USA) in Basel (Blum *et al.* 2002). Both intra- and inter-observer variability analyses were performed to ensure the accuracy of data acquired. T2-lesion volumes from regions of interest (ROI) selections were calculated by multiplying the area of the lesion by the slice thickness and the number of slices penetrated using in-house software.

Measurement of cortical thickness

FreeSurfer Image Analysis Suite v5.2.0 (http://surfer.nmr.mgh.harvard.edu/) (Dale *et al.* 1999; Fischl *et al.* 1999) was used for cortical reconstruction, volumetric segmentation, subcortical (Fischl *et al.* 2004) and cortical parcellation (Desikan-Kylliany Atlas). Volumes were normalized for intracranial volume (Desikan *et al.* 2006).

Segmentation results were assessed by an experienced operator to ensure accuracy and manually edited when needed. Both WM and juxtacortical lesions were filled before running FreeSurfer segmentation of the cortical thicknesses.

Two cases were removed for having more than two times the interquartile range of median cortical thickness across all cortical regions. Thickness values were confined to a range of the mean ± 3 SD in each region.

Defining cortices of interest for GWAS

In order to define ROI for the GWAS, cortical thickness adjusted by age, sex and a volumetric scaling factor (used for standardizing head size) in each of the 34 regions parcellated by FreeSurfer was compared between cases and controls in UCSF using a linear regression model. The Bonferroni method was used to adjust significance for multiple hypotheses testing in each hemisphere. ROI were defined as cortex areas showing a statistically significant ($P_{corr} < 0.05$) difference between cases and controls in either hemisphere (corrected for 34 tests) and a nominal *P*-value of 0.05 in the contralateral hemisphere. The cortical thickness in the ROI was independently measured in patients from the Basel cohort in order to confirm the genetic associations observed. This analysis was performed using the R statistical package.

SNP genotyping

DNA was isolated from blood cells using standard protocols. Genotyping of subjects was performed at the Illumina facilities using the Sentrix® HumanHap550 BeadChip as described elsewhere (Baranzini *et al.* 2009b). Single nucleotide polymorphisms (SNPs) that did not satisfy the following quality control criteria were excluded: genotype call rate <95%, significant deviation from Hardy–Weinberg equilibrium $P < 0.001$ and minor allele frequency (MAF) <0.005.

GWAS with cortical thickness as outcome variable

For the MS patients, allelic effects for cortical thickness were calculated based on a linear regression model including age at examination, sex, disease course, a volumetric scaling factor and T2 lesion load as covariates. Disease course was aggregated into two categories: non-progressive MS (CIS + RRMS) and progressive MS (SP + PP + PRMS). The genetic effects for thickness were tested using minor allele dominant/recessive/additive linear regression models and the most significant *P*-value among the three models was selected for each SNP. To keep minor allele homozygous sample size, MAF threshold was set at 0.15 in the recessive model. All calculations were performed in PLINK v1.07 (Purcell *et al.* 2007).

Computing gene-wise P*-values and association blocks*

The first step in any pathway analysis is to compute gene-level *P*-values. To that end, we used versatile gene-based association study (VEGAS v0.7.30), a previously described method that takes into account gene sizes and patterns of linkage disequilibrium (LD) (Liu *et al.* 2010). Versatile gene-based association study assigns SNPs to each of 17 787 autosomal genes according to positions on the UCSC Genome Browser (hg18 assembly). To capture regulatory regions and SNPs in LD, gene boundaries are defined as 50 kb beyond the 5′ and 3′ untranslated regulatory regions of each gene. Versatile gene-based association study factors in LD patterns between markers within a gene by using Monte–Carlo simulations from the multivariate normal distribution on the basis of the LD structure of a set of reference individuals (the HapMap2 CEU population). An empirical *P*-value of 0 from 10^6 VEGAS simulations can be interpreted as $P < 10^{-6}$, which exceeds a Bonferroni-corrected threshold of $P < 2.8 \times 10^{-6}$ (=0.05/17 787; this threshold is likely to be conservative given the overlap between genes).

We defined association blocks as those groups of sequential genes with a *P*-value <0.05. A block_id was assigned to each association block along the genome for each region. To evaluate the likelihood that these nominally significant genes were replicated in two data sets by chance, a hypergeometric test (one-sided Fisher's exact test) was conducted in each region.

PIN-based pathway analysis

We downloaded the entire iRefIndex database (Razick *et al.* 2008), a collection of 15 human PIN data sets from different sources, and computed the union data set. This set comprised more than 400 000 interactions among ~25 000 proteins. However, many of these interactions were either predicted or backed up by a single experiment (i.e. a single publication). In order to minimize the rate of false positives, we then filtered this large network to keep only those interactions that were described in at least two independent publications. This resulted in a network of 8960 proteins (nodes) and 27 724 interactions (edges). We used this high-confidence network for all subsequent analyses. The network was uploaded into Cytoscape 3.0.1 and annotated with genomic position, gene-wise *P*-value and block_id (loaded as node attributes) for all the analyzed regions.

For each region and data set, we computed significant first-order interactions by filtering the main network so as to keep only those genes (i.e. their encoded proteins) with VEGAS *P*-values <0.05. Then, the number of resulting nodes and edges and the size of the largest connected component were computed within Cytoscape. To evaluate the likelihood that these numbers were obtained by chance (as a consequence of the sheer number of interactions), we computed 1000 simulations by assigning *P*-values at random from the same network and creating subnetworks of similar size. These simulations were used as background for estimating the significance of the subnetworks obtained with the real gene-wise *P*-values. We then used the program (plugin) jActiveModules to conduct searches of subnetworks enriched with (but not necessarily composed of) genes with significant *P*-values (Ideker *et al.* 2002). jActiveModules starts by converting each gene *P*-value into a *Z* score by using the inverse normal cumulative distribution function. Then, it produces an aggregate *Z* score (Z_A) for an entire subnetwork A of k genes by summing the Z_i over all genes in the subnetwork $Z_A = 1/\sqrt{k} \sum_{i \in A} Z_i$. *Z* score is later corrected for network size and topology and an S_A value is produced. As in previous studies, we took an $S_A > 3$ as evidence of a biologically active subnetwork (Baranzini *et al.* 2009a; Ideker *et al.* 2002; IMSGC 2013). To evaluate the likelihood that these enriched networks were replicated in the two data sets by chance, a hypergeometric test (one-sided Fisher's exact test) was performed in each region and the significance was estimated by a permutation test, in which we iteratively shuffled gene-wise *P*-values, mapped the shuffled *P*-values to the PIN, extracted the enriched subnetworks for 50 times and then we calculated an enrichment between any two out of the 50 independent data sets. Based on an enrichment distribution of the 1225 pair-wise combinations, we calculated the fraction of permutations that showed a more extreme value than the enrichment observed in each original data set

Functional annotation enrichment analysis

The reported biological significance of gene sets was evaluated by gene ontology (GO) analysis (biological process FAT set) by using the Cytoscape plug-in ClueGO v2.1.5 (Bindea *et al.* 2009) with the following parameters: statistical test = enrichment (right-sided hypergeometric test), correction method = Bonferroni step down, minimum level = 3, maximum level = 8, number of genes = 3, minimum percentage = 5.0 and Kappa score threshold = 0.3.

Results

Here, we present a gene-level association study of cortical thickness on 557 MS patients and 75 controls of European ancestry studied at UCSF. An independent cohort of 219 patients studied in Basel was used as replication. The complete workflow of our strategy is depicted in Fig. S1. Through this approach, we merge statistical evidence of association

Table 2: Comparing adjusted thickness between cases and controls

		Left		Right	
	Region	Estimate	*P*	Estimate	*P*
Frontal lobe	Frontal pole	3.05E-03	9.34E-01	−1.57E-02	6.59E-01
	Medial orbital frontal	3.94E-02	6.90E-02	3.91E-02	6.05E-02
	Lateral orbital frontal	−3.28E-02	6.65E-02	−2.14E-02	2.28E-01
	Pars orbitalis	−1.66E-02	4.91E-01	1.65E-02	5.03E-01
	Pars triangularis	−1.66E-02	4.91E-01	1.65E-02	5.03E-01
	Pars opercularis	−1.82E-02	2.31E-01	−8.49E-03	6.05E-01
	Superior frontal*	−4.50E-02	1.27E-03	−3.00E-02	3.90E-02
	Rostral middle frontal	−9.69E-03	4.38E-01	−3.18E-02	1.70E-02
	Caudal middle frontal	−3.43E-02	1.87E-02	−1.26E-02	4.35E-01
	Paracentral*	−4.31E-02	1.09E-02	−6.22E-02	1.73E-04
	Precentral	−3.77E-02	6.88E-03	−1.36E-02	3.19E-01
Parietal lobe	Precuneus*	−6.86E-02	7.78E-06	−5.94E-02	1.90E-04
	Postcentral	−1.49E-02	2.91E-01	−1.25E-02	3.77E-01
	Superior parietal*	−4.56E-02	9.39E-04	−4.39E-02	2.51E-03
	Supramarginal*	−5.16E-02	3.80E-04	−1.97E-02	1.68E-01
	IP*	−7.20E-02	6.97E-06	−4.47E-02	6.44E-03
Temporal lobe	Transverse temporal	3.42E-02	1.42E-01	6.50E-02	1.00E-02
	Superior temporal	2.82E-04	9.87E-01	−2.18E-02	2.07E-01
	Banks of superior temporal sulcus*	−9.03E-02	1.49E-06	−8.14E-02	3.03E-05
	Middle temporal	−1.58E-02	3.70E-01	−2.22E-02	1.75E-01
	Inferior temporal	−2.96E-02	1.00E-01	−2.66E-02	1.44E-01
	FUS*	−6.59E-02	2.00E-04	−4.56E-02	1.38E-02
	Temporal pole	9.87E-02	8.56E-02	1.50E-01	7.78E-03
Occipital lobe	Cuneus	−1.80E-02	2.80E-01	−1.05E-02	5.23E-01
	Pericalcarine	3.06E-03	8.53E-01	−9.25E-03	5.87E-01
	Lateral occipital	−2.68E-02	8.61E-02	−3.37E-02	2.75E-02
	Lingual	−4.20E-02	6.81E-03	−2.41E-02	1.01E-01
Limbic system	Rostral anterior cingulate	−3.64E-03	9.15E-01	7.13E-02	3.08E-02
	Caudal anterior cingulate	5.36E-02	1.44E-01	4.69E-02	1.32E-01
	Posteriorcingulate	−1.66E-02	3.85E-01	−4.07E-02	4.48E-02
	Isthmus cingulate*	−1.93E-01	2.11E-12	−1.47E-01	1.96E-08
	Parahippocampal	2.03E-02	5.79E-01	5.01E-02	1.18E-01
	Entorhinal	−1.33E-01	6.69E-03	−8.48E-03	8.74E-01
Insula	Insula	−3.17E-02	1.19E-01	−4.64E-02	2.17E-02

Estimate describes the difference of adjusted mean thickness between cases and controls.
*ROI. *P*-value corrected by Bonferroni method <0.05 in one hemisphere and nominal *P*-value < 0.05 in the contralateral hemisphere.

and physical evidence of interaction at the protein level to identify associated loci and highlight functional pathways involved in cortical thickness of MS.

Defining ROI for GWAS

Cortical thickness was determined in all cases and controls at UCSF for 34 regions bilaterally using FreeSurfer (see *Methods and methods*). We first selected ROI for GWAS by comparing the thickness of each region between MS and age- and sex-matched controls. Demographic features are summarized in Table 1. Cortical thickness was independently measured in 219 MS patients from the Basel cohort. Cases enrolled in Basel had significantly longer disease duration and more disability than cases in UCSF, and the frequency of progressive MS was also higher in Basel than in UCSF (Table 1). One individual who had more than two times the interquartile range of median cortical thickness was excluded from the Basel data set. Differences in the mean adjusted cortical thickness between cases and controls were statistically significant in nine regions (Table 2). Eight of the nine regions showed evidence for bilateral pathology (Bonferroni *P*-value <0.05 in one hemisphere and nominal *P*-value <0.05 in the contralateral hemisphere); the remaining region (supramarginal cortex) was significant on the left hemisphere but did not reach significance on the right hemisphere. These nine regions were defined as ROI. As expected, the mean thickness in cases was lower (i.e. thinner) than controls for all ROI (Fig. S2). Based on the bilateral pattern of thinning observed for most ROI, we used an average thickness of right and left hemispheres as a trait for GWAS. We also found that cortical thickness adjusted by age, sex and head size was negatively correlated with EDSS in eight of the nine ROI (with the exception of banks of superior temporal sulcus) in both UCSF and Basel data sets (Table S1).

SNP-level GWAS

Genotypes ($n = 553\,139$) for 464 (out of 557) cases from UCSF, and 211 (out of 219) cases from Basel obtained with the Illumina HumanHap550 BeadChip were available for this analysis. After quality control (see *Materials and methods*), GWAS was carried out on 550 067 SNPs in 464 cases from UCSF and 211 cases from Basel using left–right average thickness at each ROI as phenotype (nine GWAS in total). To this end, a linear regression model was used for each data set including age at examination, sex, volumetric scaling factor, T2 lesion load and disease course (non-progressive or progressive MS) as covariates. After correcting for the number of GWAS performed, no marker was statistically significant at genome-wide level in either data set. We explored the possibility of replication at the SNP level by selecting the top 100 SNPs for each ROI in one data set (UCSF) and examined the corresponding GWAS in the Basel data set. None of the top 100 SNPs chosen was replicated for any ROI.

Nominal gene-level associations with cortical thickness

Previous work by us and others suggest that gene-level analysis afford greater power by virtue of aggregation of variants into functionally relevant units (IMSGC 2013; IMSGC *et al.* 2010; Liu *et al.* 2010; Purcell *et al.* 2009). Thus, we first converted SNP-level *P*-values into gene-level *P*-values and subsequently conducted a PIN-based pathway analysis. We started by computing gene-wise *P*-values using the software VEGAS, which aggregates evidence of association for all SNPs mapping to a given gene taking into account relative genomic position, number of SNPs within a gene and LD patterns for the appropriate ethnic background. Because our main hypothesis states that even modestly associated genes might participate in biologically plausible pathways, we considered all nominally significant genes (VEGAS $P < 0.05$). As in previous analyses, we observed that the distribution of nominally significant genes was not random across the genome, but it followed patterns resembling LD blocks. We then defined 'association blocks' as stretches of contiguous genes showing $P < 0.05$. This resulted in an average of 1286 blocks across the nine GWAS (range: 1249–1334) for the UCSF data set and 1273 blocks (range: 1213–1330) for the Basel data set. An average of 2762 genes were nominally significant in the UCSF data set (range: 2629–2913) and 2654 (range: 2462–2778) in the Basel data set. While most of these genes may represent false positives (given the low threshold of significance), we computed the overlap between the two data sets and compared against a background distribution generated with 1000 permutations of randomly selected genes from the same distribution, as a measure of replication. We found that the GWAS for fusiform (FUS) and inferior parietal (IP) ROI resulted in higher than expected (45% and 51%, respectively), statistically significant (one-sided Fisher's exact test, $P = 0.0388$ and 0.0107, respectively) overlap (Table S2). It is noteworthy that, unlike the SNP-level analysis, the gene-level analysis was able to suggest global gene-level replication across sites, at least for two ROI. A representative Manhattan plot visualization of the GWAS on IP at the gene level shows similar peaks (arrowheads) in both data sets (Fig. 1). The remaining plots are available in Fig. S3.

Fourteen genes in UCSF and 57 genes in Basel data set were nominally significant in all nine ROI (Table S3), and

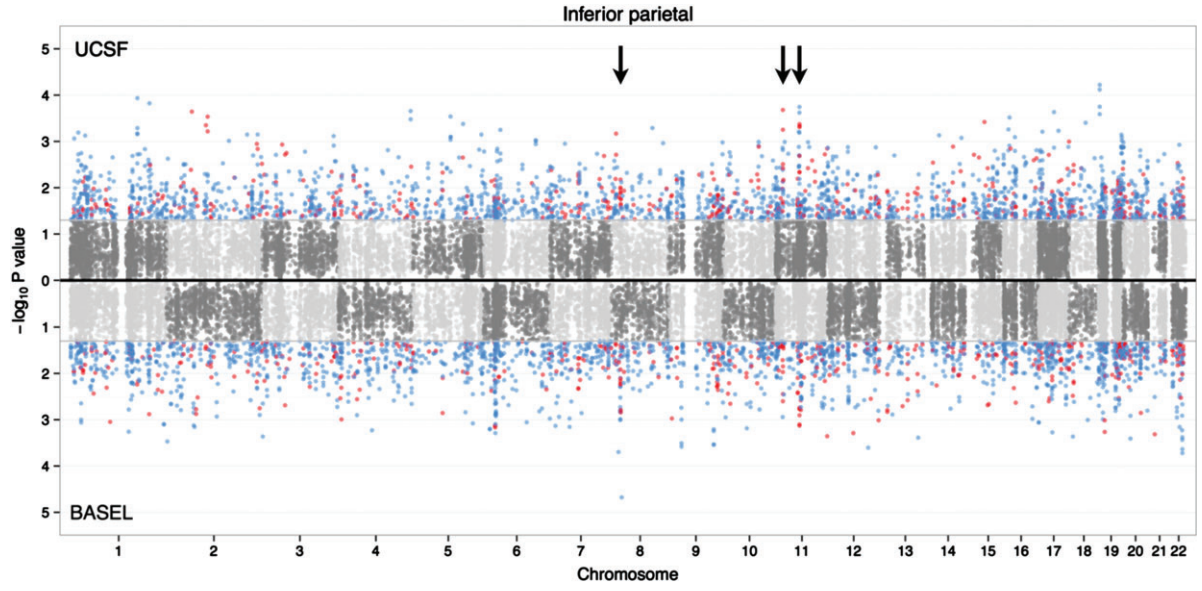

Figure 1: Double Manhattan plot. A Manhattan plot showing the gene-level *P*-values of both GWAS for thickness of IP cortex. Gene-level *P*-values from the GWAS in UCSF data set are displayed at the top, and those corresponding to the GWAS in Basel data set are at the bottom. Genes that are not nominally significant ($P > 0.05$) are displayed as gray points and nominally significant genes are displayed as blue points. Genes that are nominally significant in both two data sets are displayed as red points.

two genes (*RYR2* and *CDH13*) were replicated in both data sets (55.5-fold enrichment, one-sided Fisher's exact test, $P = 0.00090$).

Pathway analysis using PINs

In order to identify loci associated with cortical thickness by combining statistical evidence of gene association and physical evidence of interaction of their respective gene products, we assigned gene-wise association *P*-values to a curated human PIN data set consisting of 8960 proteins (nodes) and 27 724 interactions (edges) in each GWAS data set. All subsequent experiments were performed with Cytoscape, an open-source and extensible tool for network visualization and analysis (Shannon *et al.* 2003). We first extracted the nodes with *P*-values <0.05 and computed the number of nodes and edges in the resulting subnetworks for the UCSF and Basel data sets, respectively (Table S4) (we refer to these as first-order interaction networks, as all nodes are at least nominally significant). We next computed the overlap between genes (proteins) in each subnetwork between the UCSF and Basel data sets, as a measure of replication. In both data sets, the subnetworks in most of the regions were more connected than would be expected by chance, as shown by a simulation experiment in which 1000 networks of similar size were extracted from the same PIN at random (Fig. 2a). In each case, subnetworks were composed of a large connected component and several smaller networks or isolated nodes (singletons). We also tested whether the size of the main component was higher than what would be expected by chance (given the number of edges in the first-order network) (Fig. 2b). Using these two indices of network complexity, networks identified in seven of the nine ROI were beyond the 75 percentile of at least one index in random networks in either UCSF or Basel data set. In general, nominally significant genes were more connected and had a larger main connected component than expected. Our results suggest that the properties of these networks are consistent with those displayed by other biologically relevant networks.

Given the small-world topology of the human protein interactome, it is possible that a few highly connected nodes (hubs) bring together several associated genes, even though the hubs themselves are not associated. To explore this possibility, we conducted searches for subnetworks enriched with significant genes by using jActiveModules, a Cytoscape plugin based on a greedy heuristic algorithm with internal cross-validation.(Ideker *et al.* 2002) Between 23 and 36 significant and minimally overlapping (<20%) modules were identified in each region for both data sets (Table 3). We next computed the union of all modules for each cortical region within each data set, resulting in a single connected network of 1763.8 nodes on average (range: 1592–1963) for the UCSF data set and another of 1898 nodes on average (range: 1707–2087) for the Basel data set. We next computed the nodes in common between both data sets for each cortical region, resulting in an average overlap of 698.6 nodes (range: 610–784).

We then considered only replicated genes and further searched for subnetworks in order to detect networks robustly associated with cortical thickness. In each region, two to four enriched subnetworks composed by 9–92 nodes were identified. In each subnetwork, we extracted genes that had nominally significant *P*-values in both UCSF and Basel data sets (Table S5). These genes are of highest importance to our approach because they had significant *P*-values in both data sets and because they were identified as components of significant networks in both data sets. Among these genes, *NPAS3* (Neuronal PAS domain protein 3 is a brain-enriched transcription factor belonging to the bHLH-PAS superfamily of transcription factors), *OPCML* and *HS3ST4* were frequently observed across all regions. They were observed in eight, seven and seven regions, respectively.

Biological significance of associated genes in MS

To explore the biological significance of the genes constituting the significant pathways, we conducted a GO analysis by using Cytoscape plug-in, ClueGO, for significant subnetworks identified in both data sets. GO analysis (biological process) identified up to nine significantly enriched categories in all regions except FUS cortex (Figs 3 and S4, Table S6). The category 'negative regulation of neuron differentiation' was significantly enriched in IP (enrichment: 73.5, $P_{corr} = 1.73 \times 10^{-5}$), paracentral (enrichment: 82.8, $P_{corr} = 1.51 \times 10^{-6}$), superior parietal (enrichment: 32.3, $P_{corr} = 8.74 \times 10^{-7}$) and terms related to glutamate biology were identified in

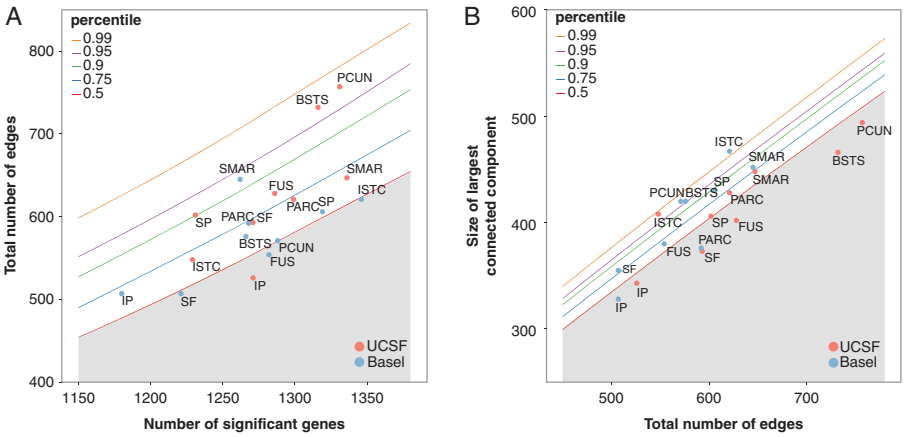

Figure 2: Statistical significance of identified modules. The number of connections among significant genes is evaluated in the background of 1000 random simulations. (a) The total number of edges is plotted as a function of the number of significant genes for each region. (b) The size of the largest connected component is plotted as a function of the total number of edges. The colored lines represent the 50th (red), 75th (blue), 90th (green), 95th (purple) and 99th (orange) percentiles obtained through simulations with random gene sets of similar size.

Table 3: Summary of enriched network modules in each ROI

	Enriched networks		Size of union net					
	UCSF	Basel	UCSF	Basel	Intersection nodes	Enrichment	Pemutation *P*-value	Enriched networks in intersection
Banks superior temporal sulcus	30	26	1809	1847	701	3.31	3.76E-01	2
FUS	25	28	1632	1829	642	3.35	3.20E-01	2
IP	27	25	1834	1758	714	3.71	3.84E-02	3
Isthmus cingulate	27	32	1701	2012	732	3.53	1.36E-01	3
Paracentral	32	36	1963	2087	784	2.91	9.34E-01	2
Precuneus	25	32	1592	1994	663	3.24	5.26E-01	3
Superior frontal	30	31	1867	1934	733	3.17	6.33E-01	4
Superior parietal	26	23	1659	1707	610	3.29	4.24E-01	3
Supramarginal	26	30	1818	1916	709	3.14	6.70E-01	3

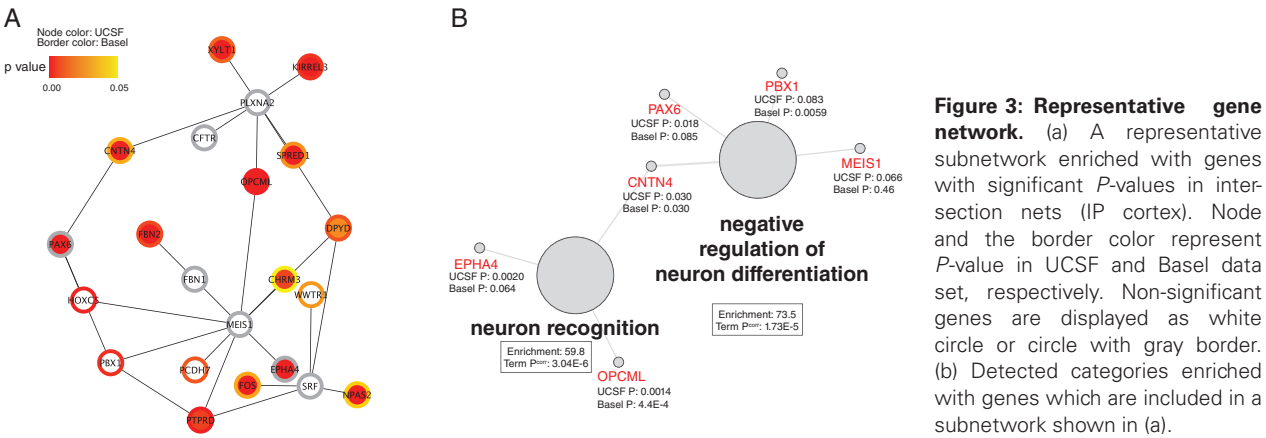

Figure 3: Representative gene network. (a) A representative subnetwork enriched with genes with significant *P*-values in intersection nets (IP cortex). Node and the border color represent *P*-value in UCSF and Basel data set, respectively. Non-significant genes are displayed as white circle or circle with gray border. (b) Detected categories enriched with genes which are included in a subnetwork shown in (a).

banks of superior temporal sulcus ('glutamate secretion', enrichment: 49.5, $P_{corr} = 5.06 \times 10^{-5}$) and superior frontal ('glutamate receptor signaling pathway', enrichment: 21.5, $P_{corr} = 5.18 \times 10^{-4}$). Terms associated with calcium channel were also identified in two regions (superior frontal, enrichment: 41.9, $P_{corr} = 8.41 \times 10^{-5}$, supramarginal, enrichment: 40.0, $P_{corr} = 7.84 \times 10^{-4}$). The most significant GO categories in individual regions were 'SMAD protein [The SMAD proteins are homologs of both the Drosophila protein, mothers against decapentaplegic (MAD) and the Caenorhabditis elegans protein SMA (from gene sma for small body size)] signal transduction' (IP, enrichment: 302.8, $P_{corr} = 4.42 \times 10^{-9}$), 'phosphatidylinositol 3 kinase activity' (isthmus of cingulate gyrus, enrichment 350.4, $P_{corr} = 6.81 \times 10^{-9}$) 'fluid transport' (paracentral, enrichment: 71.6, $P_{corr} = 1.97 \times 10^{-5}$), positive regulation of 'protein kinase B signaling cascade' (precuneus, enrichment: 30.4, $P_{corr} = 3.49 \times 10^{-5}$) and 'positive regulation of phosphoprotein phosphatase activity' (superior frontal, enrichment: 90.8, $P_{corr} = 1.02 \times 10^{-4}$).

Discussion

We identified nine cortical regions in which the average thickness of cases and controls was significantly different and conducted a GWAS within MS patients to identify genetic variants that could associate with this phenotype. The thinner regions were distributed mainly along the parietal lobe, with portions of frontal and temporal lobes also involved. This pattern is somewhat different than reported in previous studies, in which the frontal and temporal lobes were predominantly involved (Achiron *et al.* 2013; Calabrese *et al.* 2010a; Sailer *et al.* 2003). The isthmus of the cingulate gyrus and the banks of superior temporal sulcus were significantly thinner in CIS cases than in controls. When comparing established MS (SP, PP and PRMS) and controls, eight of the nine ROI (except for superior parietal cortex in RRMS) were similarly involved. The finding that these regions were reproducibly identified even in different disease courses underscores the importance of these ROI in MS. Interestingly, cortical thickness in almost all ROI correlated with EDSS, suggesting the relevance of these regions in the development of disability in MS. Intriguingly, eight of the nine ROI (except banks of superior temporal sulcus) were previously reported as highly heritable based on twin studies (Joshi *et al.* 2011; Kremen *et al.* 2010; Rimol *et al.* 2010), indicating a genetic effect to cortical thinning in MS.

Possibly due to the limited size of the cohorts analyzed, no marker exceeded a genome-wide significant *P*-value

($P < 9.42 \times 10^{-8}$), and no replication was observed between the two data sets at the SNP level. However, markers detected in GWAS explain a limited fraction of the heritable component of common diseases and traits and a sizable proportion of risk alleles are still being missed under an assumption that each marker has independent effect (Bodmer & Bonilla 2008).

Recently, novel approaches to GWAS analysis have been proposed. These strategies focus on the combined effects of several loci, acknowledging that each may make a small contribution to the overall phenotype, and potentially providing valuable insights into the genetic basis of common disease (IMSGC *et al.* 2010; Purcell *et al.* 2009). By analyzing gene-wise *P*-values, we report replication of more nominally significant genes ($P < 0.05$) than expected by chance in two data sets (IP and FUS) (Table S2). Also, while discrepancies existed between studies (probably due to the limited size of the smaller data set), the list of nominally significant genes in each cohort contain more shared associations than expected (Table S3). However, when genes arranged in interaction networks were compared between the two cohorts, higher replication rates were observed (Table 3). This strategy produces comparable results to the more established approach extending a fixed genetic distance from the lead SNP to the next recombination hotspot and maximizes the potential of finding *bona fide* associations(IMSGC 2013). In fact, proteins encoded by nominally associated genes were more connected in the PIN than what would be expected by chance in most of the regions (Fig. 2). Thus, analyzing nominal gene-level significance and studying genes in the context of biological networks appears a reasonable approach for these data sets.

Of the 194 genes that were nominally significant in both data sets, 53 were observed in more than one region. This finding correlates well with a previous report that both global and regionally specific genetic factors influence cortical surface areas (Eyler *et al.* 2011). Variation within *NPAS3* was a robust finding in eight of the nine ROI. *NPAS3* encodes a member of the basic helix-loop-helix and PAS [The PAS domain is a protein domain contained in many signaling proteins where it functions as a signal sensor. The PAS domain was named after the three proteins in which it was first discovered: Per (period circadian protein); Arnt (aryl hydrocarbon receptor nuclear translocator protein); Sim (single-minded protein)] domain-containing family of transcription factors and the protein is involved in neurogenesis. Chromosomal abnormalities that affect the coding potential of this gene are associated with schizophrenia and mental retardation (Pickard *et al.* 2005) and SNPs within this domain are associated with efficacy of antipsychotic iloperidone in patients with schizophrenia (Lavedan *et al.* 2009).

Of interest, most of the genes identified in this study do not harbor variants associated with susceptibility, most of which are involved in inflammatory processes (IMSGC *et al.* 2011; Patsopoulos *et al.* 2011). Thus, our results underscore the need to differentiate genetic variation that affects susceptibility, from that affecting endophenotypes such as neural degeneration (Kutzelnigg *et al.* 2005; Wegner *et al.* 2006). In this model, inflammation is a pervasive feature in the pathology of MS, and varying degrees of neurodegeneration ensue thereafter. Results from the work presented here suggest that part of the variability in neurodegeneration might be due to genetic variation. Notable exceptions are *FOXP1* (paracentral), *SDK1* (FUS, paracentral, precuneus, superior frontal, superior parietal and supramarginal), *SLC2A4RG* (paracentral) and *WWOX* (precuneus and superior parietal). *NXPH1* had nominally significant effect on cortical thickness of IP cortex in both data sets and *EPHA4* was included in a significant pathway for superior frontal cortex. SNPs proximal to these genes have been recently reported to have a significant effect on GM density of frontal lobe in Alzheimer's disease. Interestingly, GO categories related to glutamate biology were significantly enriched in banks of superior temporal sulcus and superior frontal cortex. We previously identified a module with high relevance to glutamate biology by using GWAS for glutamate concentration in MS and reported that individuals carrying a higher number of associated alleles from genes in the module showed greater decreases in brain volume over 1 year of follow-up (Baranzini *et al.* 2010). Thus, pathways related to glutamate may influence cortical thickness as well as overall brain volume. The GO categories 'glutamate toxicity' and 'neuron differentiation' were also involved in more than one region highlighting their potential role in the neurodegenerative process seen in MS. For example, *PAX6* is a critical gene shown to regulate development of CNS and axon guidance (Georgala *et al.* 2011). It is then plausible that genes involved in neural development would influence brain cortical thickness and disability in MS.

Other significant GO categories were 'cell proliferation' and 'calcium signaling'. Multiple growth factors, their receptors (*EGFR*, *FGF12*, *IGF1R*, *IGF2*, *KIT*, *PDGFB* and *PDGFRB*) and their downstream signaling cascade (phosphatidylinositol 3-kinase cascade) were identified, all of which are related to cell proliferation or survival and overlap with neural development pathways. Ryanodine receptor 2 (RYR2), which had nominally significant *P*-values in all ROI in both data sets, mediates the release of Ca^{2+} from the endoplasmic reticulum (ER) into the cytoplasm. The ryanodine receptor-mediated calcium release from the ER causes intra-axonal calcium overload and results in axonal injury (Stirling & Stys 2010). In superior frontal cortex, genes controlling re-uptake of cytoplasmic calcium (*CHRM3* and *ADCY2*) were also involved. These findings suggest that dysregulation of intra-axonal calcium level contributes to the cortical thinning.

In summary, we report variation of genes associated with cortical thinning, and indirectly with disability in MS by GWAS based on two independent data sets. The genes identified were largely independent from those harboring variants associated with MS risk and were involved in glutamate signaling, neural development and an adjustment of intracellular calcium concentration. These results suggest that excitotoxicity and genetic vulnerability for axonal damage can poise the MS brain to initiate the cascade that will result in neurological disability. This study highlights the genetic influence on an aspect of neurodegeneration of MS and may be helpful in the search for therapeutic targets of disability in MS.

References

Achiron, A., Chapman, J., Tal, S., Bercovich, E. & Gil, H. (2013) Superior temporal gyrus thickness correlates with cognitive performance in multiple sclerosis. *Brain Struct Funct* **218**, 943–950.

Baranzini, S.E., Galwey, N.W., Wang, J., Khankhanian, P., Lindberg, R., Pelletier, D., Wu, W., Uitdehaag, B.M., Kappos, L., Polman, C.H., Matthews, P.M., Hauser, S.L., Gibson, R.A., Oksenberg, J.R. & Barnes, M.R. (2009a) Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. *Hum Mol Genet* **18**, 2078–2090.

Baranzini, S.E., Wang, J., Gibson, R.A. *et al.* (2009b) Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet* **18**, 767–778.

Baranzini, S.E., Srinivasan, R., Khankhanian, P., Okuda, D.T., Nelson, S.J., Matthews, P.M., Hauser, S.L., Oksenberg, J.R. & Pelletier, D. (2010) Genetic variation influences glutamate concentrations in brains of patients with multiple sclerosis. *Brain* **133**, 2603–2611.

Beecham, A.H., Patsopoulos, N.A., Xifara, D.K. *et al.* (2013) Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat Genet* **45**, 1353–1360.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.H., Pages, F., Trajanoski, Z. & Galon, J. (2009) ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093.

Blum, D., Yonelinas, A.P., Luks, T., Newitt, D., Oh, J., Lu, Y., Nelson, S., Goodkin, D. & Pelletier, D. (2002) Dissociating perceptual and conceptual implicit memory in multiple sclerosis patients. *Brain Cogn* **50**, 51–61.

Bo, L., Vedeler, C.A., Nyland, H.I., Trapp, B.D. & Mork, S.J. (2003) Subpial demyelination in the cerebral cortex of multiple sclerosis patients. *J Neuropathol Exp Neurol* **62**, 723–732.

Bodmer, W. & Bonilla, C. (2008) Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet* **40**, 695–701.

Calabrese, M., De Stefano, N., Atzori, M., Bernardi, V., Mattisi, I., Barachino, L., Morra, A., Rinaldi, L., Romualdi, C., Perini, P., Battistin, L. & Gallo, P. (2007) Detection of cortical inflammatory lesions by double inversion recovery magnetic resonance imaging in patients with multiple sclerosis. *Arch Neurol* **64**, 1416–1422.

Calabrese, M., Rinaldi, F., Mattisi, I., Grossi, P., Favaretto, A., Atzori, M., Bernardi, V., Barachino, L., Romualdi, C., Rinaldi, L., Perini, P. & Gallo, P. (2010a) Widespread cortical thinning characterizes patients with MS with mild cognitive impairment. *Neurology* **74**, 321–328.

Calabrese, M., Rocca, M.A., Atzori, M., Mattisi, I., Favaretto, A., Perini, P., Gallo, P. & Filippi, M. (2010b) A 3-year magnetic resonance imaging study of cortical lesions in relapse-onset multiple sclerosis. *Ann Neurol* **67**, 376–383.

Calabrese, M., Rinaldi, F., Mattisi, I., Bernardi, V., Favaretto, A., Perini, P. & Gallo, P. (2011) The predictive value of gray matter atrophy in clinically isolated syndromes. *Neurology* **77**, 257–263.

Calabrese, M., Seppi, D., Cocco, E., Poretto, V., Rinaldi, F., Perini, P. & Gallo, P. (2012) Cortical pathology in RRMS: taking a cue from four sisters. *Mult Scler Int* **2012**, Article ID 760254, doi: 10.1155/2012/760254.

Chard, D.T., Griffin, C.M., Parker, G.J., Kapoor, R., Thompson, A.J. & Miller, D.H. (2002) Brain atrophy in clinically early relapsing-remitting multiple sclerosis. *Brain* **125**, 327–337.

Chen, J.T., Narayanan, S., Collins, D.L., Smith, S.M., Matthews, P.M. & Arnold, D.L. (2004) Relating neocortical pathology to disability progression in multiple sclerosis using MRI. *Neuroimage* **23**, 1168–1175.

Dale, A.M., Fischl, B. & Sereno, M.I. (1999) Cortical surface-based analysis. I Segmentation and surface reconstruction. *Neuroimage* **9**, 179–194.

Dalton, C.M., Chard, D.T., Davies, G.R., Miszkiel, K.A., Altmann, D.R., Fernando, K., Plant, G.T., Thompson, A.J. & Miller, D.H. (2004) Early development of multiple sclerosis is associated with progressive grey matter atrophy in patients presenting with clinically isolated syndromes. *Brain* **127**, 1101–1107.

De Stefano, N., Matthews, P.M., Filippi, M., Agosta, F., De Luca, M., Bartolozzi, M.L., Guidi, L., Ghezzi, A., Montanari, E., Cifelli, A., Federico, A. & Smith, S.M. (2003) Evidence of early cortical atrophy in MS: relevance to white matter changes and disability. *Neurology* **60**, 1157–1162.

Desikan, R.S., Segonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S. & Killiany, R.J. (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968–980.

Eyler, L.T., Prom-Wormley, E., Panizzon, M.S. *et al.* (2011) Genetic and environmental contributions to regional cortical surface area in humans: a magnetic resonance imaging twin study. *Cereb Cortex* **21**, 2313–2321.

Eyler, L.T., Chen, C.H., Panizzon, M.S., Fennema-Notestine, C., Neale, M.C., Jak, A., Jernigan, T.L., Fischl, B., Franz, C.E., Lyons, M.J., Grant, M., Prom-Wormley, E., Seidman, L.J., Tsuang, M.T., Fiecas, M.J., Dale, A.M. & Kremen, W.S. (2012) A comparison of heritability maps of cortical surface area and thickness and the influence of adjustment for whole brain measures: a magnetic resonance imaging twin study. *Twin Res Hum Genet* **15**, 304–314.

Filippi, M., Rocca, M.A., Calabrese, M., Sormani, M.P., Rinaldi, F., Perini, P., Comi, G. & Gallo, P. (2010) Intracortical lesions: relevance for new MRI diagnostic criteria for multiple sclerosis. *Neurology* **75**, 1988–1994.

Fischl, B., Sereno, M.I. & Dale, A.M. (1999) Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. *Neuroimage* **9**, 195–207.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B. & Dale, A.M. (2004) Automatically parcellating the human cerebral cortex. *Cereb Cortex* **14**, 11–22.

Fisniku, L.K., Brex, P.A., Altmann, D.R., Miszkiel, K.A., Benton, C.E., Lanyon, R., Thompson, A.J. & Miller, D.H. (2008) Disability and T2 MRI lesions: a 20-year follow-up of patients with relapse onset of multiple sclerosis. *Brain* **131**, 808–817.

Georgala, P.A., Carr, C.B. & Price, D.J. (2011) The role of Pax6 in forebrain development. *Dev Neurobiol* **71**, 690–709.

Geurts, J.J., Pouwels, P.J., Uitdehaag, B.M., Polman, C.H., Barkhof, F. & Castelijns, J.A. (2005) Intracortical lesions in multiple sclerosis: improved detection with 3D double inversion-recovery MR imaging. *Radiology* **236**, 254–260.

Giorgio, A., Stromillo, M.L., Rossi, F., Battaglini, M., Hakiki, B., Portaccio, E., Federico, A., Amato, M.P. & De Stefano, N. (2011) Cortical lesions in radiologically isolated syndrome. *Neurology* **77**, 1896–1899.

Hauser, S.L. & Oksenberg, J.R. (2006) The neurobiology of multiple sclerosis: genes, inflammation, and neurodegeneration. *Neuron* **52**, 61–76.

Henry, R.G., Shieh, M., Okuda, D.T., Evangelista, A., Gorno-Tempini, M.L. & Pelletier, D. (2008) Regional grey matter atrophy in clinically isolated syndromes at presentation. *J Neurol Neurosurg Psychiatry* **79**, 1236–1244.

Ideker, T., Ozier, O., Schwikowski, B. & Siegel, A.F. (2002) Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* **18**(Suppl. 1), S233–S240.

IMSGC (2013) Network-based multiple sclerosis pathway analysis with GWAS data from 15,000 cases and 30,000 controls. *Am J Hum Genet* **92**, 854–865.

IMSGC, Bush, W.S., Sawcer, S.J., de Jager, P.L., Oksenberg, J.R., McCauley, J.L., Pericak-Vance, M.A. & Haines, J.L. (2010) Evidence for polygenic susceptibility to multiple sclerosis--the shape of things to come. *Am J Hum Genet* **86**, 621–625.

IMSGC, Sawcer, S., Hellenthal, G. *et al.* (2011) Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* **476**, 214–219.

Joshi, A.A., Lepore, N., Joshi, S.H., Lee, A.D., Barysheva, M., Stein, J.L., McMahon, K.L., Johnson, K., de Zubicaray, G.I., Martin, N.G.,

Wright, M.J., Toga, A.W. & Thompson, P.M. (2011) The contribution of genes to cortical thickness and volume. *Neuroreport* **22**, 101–105.
Kochunov, P., Glahn, D.C., Lancaster, J., Thompson, P.M., Kochunov, V., Rogers, B., Fox, P., Blangero, J. & Williamson, D.E. (2011) Fractional anisotropy of cerebral white matter and thickness of cortical gray matter across the lifespan. *Neuroimage* **58**, 41–49.
Kremen, W.S., Prom-Wormley, E., Panizzon, M.S., Eyler, L.T., Fischl, B., Neale, M.C., Franz, C.E., Lyons, M.J., Pacheco, J., Perry, M.E., Stevens, A., Schmitt, J.E., Grant, M.D., Seidman, L.J., Thermenos, H.W., Tsuang, M.T., Eisen, S.A., Dale, A.M. & Fennema-Notestine, C. (2010) Genetic and environmental influences on the size of specific brain regions in midlife: the VETSA MRI study. *Neuroimage* **49**, 1213–1223.
Kurtzke, J.F. (1983) Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* **33**, 1444–1452.
Kutzelnigg, A., Lucchinetti, C.F., Stadelmann, C., Bruck, W., Rauschka, H., Bergmann, M., Schmidbauer, M., Parisi, J.E. & Lassmann, H. (2005) Cortical demyelination and diffuse white matter injury in multiple sclerosis. *Brain* **128**, 2705–2712.
Lavedan, C., Licamele, L., Volpi, S., Hamilton, J., Heaton, C., Mack, K., Lannan, R., Thompson, A., Wolfgang, C.D. & Polymeropoulos, M.H. (2009) Association of the NPAS3 gene and five other loci with response to the antipsychotic iloperidone identified in a whole genome association study. *Mol Psychiatry* **14**, 804–819.
Liu, J.Z., McRae, A.F., Nyholt, D.R., Medland, S.E., Wray, N.R., Brown, K.M., Hayward, N.K., Montgomery, G.W., Visscher, P.M., Martin, N.G. & Macgregor, S. (2010) A versatile gene-based test for genome-wide association studies. *Am J Hum Genet* **87**, 139–145.
Nakamura, K. & Fisher, E. (2009) Segmentation of brain magnetic resonance images for measurement of gray matter atrophy in multiple sclerosis patients. *Neuroimage* **44**, 769–776.
Noseworthy, J.H., Lucchinetti, C., Rodriguez, M. & Weinshenker, B.G. (2000) Multiple sclerosis. *N Engl J Med* **343**, 938–952.
Okuda, D.T., Srinivasan, R., Oksenberg, J.R., Goodin, D.S., Baranzini, S.E., Beheshtian, A., Waubant, E., Zamvil, S.S., Leppert, D., Qualley, P., Lincoln, R., Gomez, R., Caillier, S., George, M., Wang, J., Nelson, S.J., Cree, B.A., Hauser, S.L. & Pelletier, D. (2009) Genotype-phenotype correlations in multiple sclerosis: HLA genes influence disease severity inferred by 1HMR spectroscopy and MRI measures. *Brain* **132**, 250–259.
Patenaude, B., Smith, S.M., Kennedy, D.N. & Jenkinson, M. (2011) A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage* **56**, 907–922.
Patsopoulos, N.A., Esposito, F., Reischl, J. *et al.* (2011) Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Ann Neurol* **70**, 897–912.
Pickard, B.S., Malloy, M.P., Porteous, D.J., Blackwood, D.H. & Muir, W.J. (2005) Disruption of a brain transcription factor, NPAS3, is associated with schizophrenia and learning disability. *Am J Med Genet B Neuropsychiatr Genet* **136B**, 26–32.
Polman, C.H., Reingold, S.C., Edan, G., Filippi, M., Hartung, H.P., Kappos, L., Lublin, F.D., Metz, L.M., McFarland, H.F., O'Connor, P.W., Sandberg-Wollheim, M., Thompson, A.J., Weinshenker, B.G. & Wolinsky, J.S. (2005) Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Ann Neurol* **58**, 840–846.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. & Sham, P.C. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559–575.
Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M., O'Donovan, M.C., Sullivan, P.F. & Sklar, P. (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752.
Razick, S., Magklaras, G. & Donaldson, I.M. (2008) iRefIndex: a consolidated protein interaction database with provenance. *BMC Bioinformatics* **9**, 405–424.
Rimol, L.M., Panizzon, M.S., Fennema-Notestine, C., Eyler, L.T., Fischl, B., Franz, C.E., Hagler, D.J., Lyons, M.J., Neale, M.C., Pacheco, J., Perry, M.E., Schmitt, J.E., Grant, M.D., Seidman, L.J., Thermenos, H.W., Tsuang, M.T., Eisen, S.A., Kremen, W.S. & Dale, A.M. (2010) Cortical thickness is influenced by regionally specific genetic factors. *Biol Psychiatry* **67**, 493–499.
Sailer, M., Fischl, B., Salat, D., Tempelmann, C., Schonfeld, M.A., Busa, E., Bodammer, N., Heinze, H.J. & Dale, A. (2003) Focal thinning of the cerebral cortex in multiple sclerosis. *Brain* **126**, 1734–1744.
Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. & Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498–2504.
Stirling, D.P. & Stys, P.K. (2010) Mechanisms of axonal injury: internodal nanocomplexes and calcium deregulation. *Trends Mol Med* **16**, 160–170.
Tiberio, M., Chard, D.T., Altmann, D.R., Davies, G., Griffin, C.M., Rashid, W., Sastre-Garriga, J., Thompson, A.J. & Miller, D.H. (2005) Gray and white matter volume changes in early RRMS: a 2-year longitudinal study. *Neurology* **64**, 1001–1007.
Trapp, B.D. & Nave, K.A. (2008) Multiple sclerosis: an immune or neurodegenerative disorder? *Annu Rev Neurosci* **31**, 247–269.
Tur, C., Ramagopalan, S., Altmann, D.R., Bodini, B., Cercignani, M., Khaleeli, Z., Miller, D.H., Thompson, A.J. & Ciccarelli, O. (2014) HLA-DRB1*15 influences the development of brain tissue damage in early PPMS. *Neurology* **83**, 1712–1718.
Wegner, C., Esiri, M.M., Chance, S.A., Palace, J. & Matthews, P.M. (2006) Neocortical neuronal, synaptic, and glial loss in multiple sclerosis. *Neurology* **67**, 960–967.
Yates, R.L., Esiri, M.M., Palace, J., Mittal, A. & DeLuca, G.C. (2014) The influence of HLA-DRB1*15 on motor cortical pathology in multiple sclerosis. *Neuropathol Appl Neurobiol*. doi: 10.1111/nan.12165. [Epub ahead of print]

Acknowledgments

This study was supported by a grant from the National Institutes of Health (R01NS26799). S.E.B. is a Harry Weaver Neuroscience fellow from the National Multiple Sclerosis Society. We thank S. Caillier, R. Guerrero, H. Mousavi and A. Santaniello for sample and database management. We thank the UCSF EPIC Study team and MS patients and controls who participated in the study. None of the authors declared any conflict of interest.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web-site:

Figure S1: Strategy of this study. In the first step, ROI were determined by comparing thickness of cortex between cases and controls in UCSF data set. Thickness of the cortex in these regions was independently measured for patients in Basel. Then, GWAS for the thickness were conducted. Next, a gene-wise *P*-value was obtained using VEGAS, and the replication of genes was tested with *P*-value <0.05 between UCSF and Basel data set by hypergeometric test (one-sided Fisher's exact test). In the following step, gene *P*-values were overlapped with a curated PIN (see *Materials and methods*) and interactions among nominally significant genes (first-order) were computed. With the goal of discovering additional associations, a heuristic search for subnetworks enriched in significant genes was conducted using jActiveModules within Cytoscape. Replication of genes included in the subnetworks was tested again in each region. All significant subnetworks within each independent data set were

merged and finally the intersection network was computed. Finally, subnetworks enriched with significant genes were extracted from the intersection network and a GO analysis was conducted for the subnetworks.

Figure S2: Comparison of cortical thickness adjusted by age, sex and head size in ROI between cases and controls. In all detected regions, thickness is significantly lower in cases than in controls.

Figure S3: Manhattan plots showing the gene-level *P*-values of both GWAS for cortical thickness in all ROI. Gene-level *P*-values from the GWAS in UCSF data set are displayed at the top, and those corresponding to the GWAS in Basel data set are at the bottom. Genes that are not nominally significant ($P > 0.05$) are displayed as gray points and nominally significant genes are displayed as blue points. Genes that are nominally significant in both two data sets are displayed as red points.

Figure S4: All detected subnetwork enriched with genes with significant *P*-values in intersection nets and categories extracted from the subnetworks by GO analysis.

Table S1: Correlation between adjusted thickness in ROI and EDSS

Table S2: Numbers of nominally significant genes on protein interaction nets in UCSF and Basel data sets

Table S3: Nominally significant genes in all ROI in both UCSF and Basel data sets

Table S4: Connectedness of first-order interaction networks

Table S5: Nominally significant genes in intersection nets

Table S6: Main categories in enriched intersection nets