ASSOCIATION STUDIES ARTICLE

A non-synonymous single-nucleotide polymorphism associated with multiple sclerosis risk affects the EVI5 interactome

Alessandro Didonna[1,*], Noriko Isobe[1], Stacy J. Caillier[1], Kathy H. Li[2], Alma L. Burlingame[2], Stephen L. Hauser[1], Sergio E. Baranzini[1], Nikolaos A. Patsopoulos[3,4,5] and Jorge R. Oksenberg[1]

[1]Department of Neurology, [2]Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, CA 94158, USA, [3]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA, [4]Program in Translational NeuroPsychiatric Genomics, Department of Neurology, Institute for the Neurosciences and [5]Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02142, USA

*To whom correspondence should be addressed at: Department of Neurology, University of California at San Francisco, 675 Nelson Rising Lane, San Francisco, CA 94158, USA. Tel: +1 4155027211; Fax: +1 4154765229; Email: alessandro.didonna@ucsf.edu

Abstract

Despite recent progress in the characterization of genetic loci associated with multiple sclerosis (MS) risk, the ubiquitous linkage disequilibrium operating across the genome has stalled efforts to distinguish causative variants from proxy single-nucleotide polymorphisms (SNPs). Here, we have identified through fine mapping and meta-analysis *EVI5* as the most plausible disease risk gene within the 1p22.1 locus. We further show that an exonic SNP associated with risk induces changes in superficial hydrophobicity patterns of the coiled-coil domain of EVI5, which, in turns, affects the EVI5 interactome. Immunoprecipitation of wild-type and mutated EVI5 followed by mass spectrometry generated a roster of disease-specific interactors functionally linked to lipid metabolism. Among the exclusive binding partners of the risk variant, we describe the novel interaction with sphingosine 1-phosphate lyase (SGPL1)—a key enzyme for the creation of the sphingosine-1 phosphate gradient, which is relevant to the pathogenic process and therapeutic management of MS.

Introduction

Multiple sclerosis (MS) is a disease of the central nervous system (CNS) with autoimmune etiology and a distinctive pathological signature consisting of focal lymphocytic infiltration, microglial activation, demyelination and axonal degeneration (1). The interruption of myelinated tracts and neuronal loss result in a variety of neurological symptoms including weakness, gait instability, ataxia and cognitive deficits in the advanced stages of the disease (2). The cause of MS is still largely unknown but several lines of evidence suggest that the interplay between environmental and inherited factors contributes to the risk of developing the disease (3).

From a genetic standpoint, MS clusters within the complex disease class, a group of common disorders characterized by modest disease risk heritability and cumulative effects of a large number of allelic variants (4). The human leukocyte antigen DRB1 (*HLA-DRB1*) gene in the major histocompatibility complex (MHC) locus on chromosome 6p21.3, in particular the

DRB1*15:01 allele, is the most prominent genetic factor associated with MS susceptibility and has been confirmed in nearly all studied populations (5). A number of genome-wide association studies (GWAS) and meta-analyses have been conducted in MS and to date, 110 genomic regions outside the MHC have been firmly associated with increased risk (6). However, the pervasive genome-wide linkage disequilibrium (LD) has hindered attempts to distinguish the real causative variants from surrogate proxy single-nucleotide polymorphisms (SNPs) within each locus.

Among the 110 non-MHC regions, the locus on chromosome 1p22.1 has been found significantly associated with MS susceptibility in multiple studies across different ancestral groups. The first GWAS reported by the International Multiple Sclerosis Genetic Consortium (IMSGC) included 12 360 individuals of European ancestry (combined discovery and replication datasets) and found 5 significant SNPs within this locus tagging the ribosomal protein L5 (*RPL5*) gene (rs6604026, $P = 7.94 \times 10^{-6}$, OR = 1.15), the ecotropic viral integration site 5 (*EVI5*) gene (rs10735781, $P = 3.35 \times 10^{-4}$, OR = 1.11 and rs6680578, $P = 5.00 \times 10^{-4}$, OR = 1.11) and the family with sequence similarity 69 (*FAM69*) gene (rs7536563, $P = 9.12 \times 10^{-5}$, OR = 1.12 and rs11164838, $P = 1.91 \times 10^{-4}$, OR = 1.11) (7). The two SNPs tagging *EVI5* were subsequently replicated in a case–control study on 240 individuals from a genetically isolated Dutch population (rs10735781, $P = 0.01$, OR = 2.01 and rs6680578, $P = 0.01$, OR = 1.9) and in a study with Canadian multi-case families (rs10735781, $P = 0.03$, OR = 1.15 and rs6680578, $P = 0.04$, OR = 1.15) (8). The same two SNPs were also replicated in a case–control study that included 1574 African American individuals (rs10735781, $P = 0.006$, OR = 1.233 and rs6680578, $P = 0.025$, OR = 1.185) with the former showing the strongest association outside the MHC region of the 18 tested independent markers (9). Another SNP within EVI5 was found associated with MS in 1706 individuals from a case–control study in a Spanish dataset (rs11805321, $P = 0.08$, OR = 1.29) (10). Finally, the results from the latest GWAS performed by the IMSGC involving 38 662 subjects of European descent confirmed the strongest association in the locus to be within the *EVI5* gene (rs11810217, $P = 5.85 \times 10^{-15}$, OR = 1.15) (11). The finding was later replicated on 38 589 individuals using the ImmunoChip custom genotyping array that mapped the peak of association in the *EVI5* 3′ untranslated region (3′ UTR) (rs41286801, $P = 1.4 \times 10^{-26}$, OR = 1.19) (6).

Despite the overwhelming evidence from population genetic data converging on *EVI5* as the candidate gene in the locus to be responsible for disease risk, the only available functional data implicate the neighboring gene growth factor independent 1 (*GFI1*), as the likely disease gene rather than *EVI5*. The MS-associated SNP rs11804321, which is located in the last intron of *EVI5*, was shown to overlap with an insulator element that modulates *GFI1* expression via the transcriptional repressor CCCTC-binding factor (12). To further clarify the contribution of the *EVI5* locus to MS pathology, we fine-mapped the risk association within the locus by performing a locus meta-analysis of all published case–control GWASs and functionally linked the disease-associated polymorphism to lipid metabolism.

Results

Fine mapping of the *EVI5* locus

Genotypes from 13 799 cases and 24 252 controls representing 13 published MS datasets supplemented with imputed SNPs were included in the meta-analysis of the 1p22.1 locus (Supplementary Material, Table S1). A total of 3517 unique SNPs with minor allele frequency of at least 1% were analyzed in all
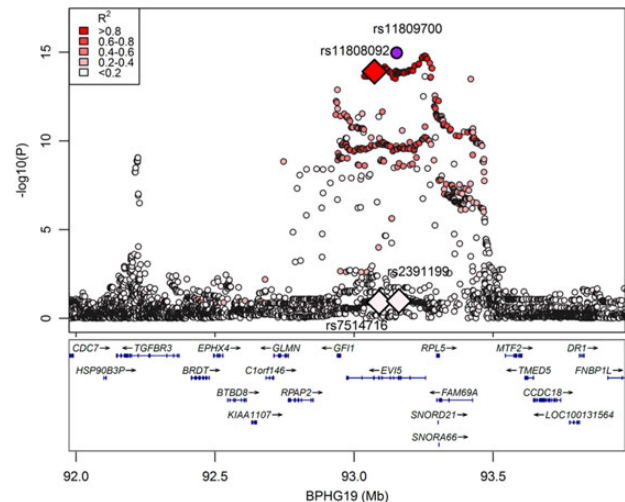
Figure 1. Meta-analysis results of the *EVI5* locus. The association *P*-values derived from meta-analysis of all reported MS case–control studies in European ancestry populations for the SNPs at 1p22.1 locus (chr1:91,975,464-93,975,464) are plotted. X-axis displays genomic positions based on hg19 and Y-axis shows $-\log_{10}$ (*P*-value) (fixed effects). Top SNP (rs11809700) is shown in purple and locates to the ninth intron of *EVI5* gene. The other SNPs are colored by the strength of LD (r^2) with the top SNP according to the European population dataset from the 1000 Genome Project. The three exonic SNPs in the *EVI5* gene are highlighted in diamonds as well. Their nomenclature and position are based on hg19 and dbSNP Build 137. Gene schematics were drawn using LocusZoom.

strata. Results show that the top SNP in the region locates in the ninth intron of *EVI5* (rs11809700, $P = 1.09 \times 10^{-15}$, OR = 1.17 in both fixed-/random-effects models) (Fig. 1). Thus, we were able to confirm the strongest association of the locus to be within *EVI5*, consistent with all previous association studies. Functional annotation for the top SNPs does not report any expression quantitative trait loci (eQTL) effects related to any gene. Similar results were obtained after extending the analysis to all the non-coding SNPs that are in linkage with the top one ($r^2 > 0.8$) according to a European reference dataset from the 1000 Genomes Project (13). Thus, we moved our investigation to exonic SNPs as possible candidates of causal variants for MS risk. Three exonic SNPs are reported within the *EVI5* gene; two SNPs (rs2391199 and rs11808092) promote the non-synonymous substitutions c.1006A>G [p.Iso336Val] and c.1836G>T [p.Gln612His], respectively, whereas the third SNP (rs7514716) encodes for the synonymous substitution c.1699A>G [p.Gln563QGln] without effects on the primary sequence of the protein. The analysis of LD patterns across the EVI5 sequence shows that rs2391199 [p.Iso336Val] is very weakly associated with the top associate intron 9 SNP rs11809700 ($r^2 = 0.057$, $P = 0.1046$) whereas rs11808092 [p.Gln612His] is in almost complete LD ($r^2 = 0.932$, $P = 1.23 \times 10^{-14}$). Importantly, the top SNP in our study is not in strong LD ($r^2 = 0.554$) with the SNP affecting *GFI1* expression (rs11804321) (12) suggesting that they are probably targeting two different biological effects. This hypothesis is also supported by the observation that after sequential conditional analysis with three independent SNPs in the locus (Table 1), the remaining top SNP (rs58394161) is in relatively strong LD with the rs11804321 SNP ($r^2 = 0.802$, $P = 1.64 \times 10^{-11}$). We also compared the genetic association results within the locus with a recently published African American MS genetic dataset (14). *EVI5* emerged as the candidate disease gene in the locus (Supplementary Material, Table S2). However, this dataset could not confirm the significance of rs11808092, most likely because of the lack of

Table 1. Meta-analysis results for associated SNPs in the *EVI5* region

Number of SNPs in model	SNP	Chr	Position	Genes	Function	RA	NRA	Fixed OR	P	Random OR	P	Het. P
0	rs11809700	1	93152635	EVI5	Intronic	T	C	1.17	1.09E−15	1.17	1.09E−15	9.14E−01
1	rs12129174	1	92222850	TGFBR3	Intronic	C	T	1.16	3.73E−10	1.16	6.72E−07	1.07E−01
2	rs1415069	1	93426869	FAM69A	Intronic	G	C	1.10	5.78E−05	1.10	5.78E−05	5.36E−01
3	rs58394161	1	92939959	GFI1	Downstream	C	T	1.11	1.23E−03	1.11	1.23E−03	4.83E−01

Positions are shown in reference to human genome 19 assembly.
Chr, chromosome; Het., heterogeneity test; NRA, non-risk allele; OR, odds ratio; RA, risk allele.

statistical power owing to the small sample size (803 cases and 1516 controls) and the low allele frequency of the exonic SNP (Supplementary Material, Table S3).

EVI5 structural analysis

EVI5 is an evolutionary conserved gene encoding in humans an 810 amino acid protein that is expressed in a variety of tissues ([15]). Structurally, it belongs to the Tre-2/Bub2/Cdc16 (TBC) protein family owing to the presence of a TBC domain in its N-terminus (160aa–371aa for human sequence) ([16]). In addition, it contains in its C-terminal half a large coiled-coil (CC) stretch (404aa–714aa) that is homologous to CC regions found in myosin and in ATPases of the structural maintenance of chromosomes family ([17]). Several lines of evidence suggest that EVI5 modulates cell cycle at multiple levels. During the early stages of mitosis, it regulates cyclin accumulation by stabilizing the anaphase and promoting complex inhibitor Emi1 ([18]). In late mitosis, it is required for the completion of cytokinesis ([19]). Moreover, EVI5 binds and modulates Rab11 functions through the GTPase-activating protein (GAP) catalytic activity of the TBC domain ([20],[21]).

The two non-synonymous SNPs are located within the functional domains of EVI5—rs2391199 (I336V) in the TBC domain and rs11808092 (Q612H) in the CC domain, respectively. We modeled the tertiary structure of the two domains using the crystal structures of TBC1D1 RabGAP domain and tropomyosin as templates (Supplementary Material, Fig. S1A and B). As expected, given the evolutionary constraint across the regions, homology modeling did not highlight any alteration in the spatial conformation of the two motifs upon amino acidic substitutions. We were able instead to detect changes in the superficial hydrophobicity patterns of the CC domain owing to the exchange of a polar uncharged amino acid (Q) with a charged one (H) (Supplementary Material, Fig. S1B). On the contrary, no differences were identified for the TBC domain as either I or V belongs to the class of hydrophobic amino acids (Supplementary Material, Fig. S1A). Considering that CC domains mediate highly specific homo- and heteromeric protein–protein interactions ([22]), we hypothesized that the Q612H polymorphism, which is strongly associated with MS risk, might affect the EVI5 CC-dependent interactome.

Proteomic analysis of EVI5 interactome

To test this hypothesis, we employed immunoprecipitation (IP) assays followed by mass spectrometry, which has become the method of choice to identify protein–protein interactions ([23]). The human EVI5 coding sequence was cloned as a FLAG-HA-fusion protein into the pSGL5 vector, and the two amino acidic substitutions were introduced independently by site-directed mutagenesis. The three constructs (designated WT-EVI5, I336V-EVI5 and Q612H-EVI5) were overexpressed into HeLa cells (Supplementary Material, Fig. S2), and the EVI5 complexes were immunoprecipitated and processed for mass spectrometry discovery. After filtering out the background, we obtained 29 (for WT-EVI5), 27 (for I336V-EVI5) and 46 (for Q612H-EVI5) putative interactors (Supplementary Material, Table S4), including 22 common interactors among the 3 variants. One protein is exclusive to each WT-EVI5 and I336V-EVI5 variants, whereas 16 are exclusive to Q612H-EVI5 (Fig. 2A), confirming our prediction that the Q612H substitution in the CC domain influenced a change in EVI5 interactome. Despite the different interactomes driven by each of the EVI5 alternative alleles, the overall properties of their resulting local networks remained largely unchanged. This was confirmed by a detailed analysis of 11 topological properties, including connectivity, centrality and path lengths (Supplementary Material, Fig. S3). This suggests—as expected—that the higher-level functions of the EVI5 interactome are not affected by these common polymorphisms.

Pathway analysis and validation

To gain mechanistic insights into the effects of the Q612H substitution on EVI5 function, we performed gene ontology (GO) analysis on the common and Q612H-EVI5-specific sets of putative interactors (Supplementary Material, Table S5). A significant enrichment in GO terms related to lipid metabolism was observed for the 16 Q612H-EVI5 specific, but not for the common binders (Fig. 2B), suggesting that the rs11808092 SNP might promote a possible gain-of-function association with lipid biosynthesis.

Remarkably, sphingosine 1-phosphate lyase (SGPL1)—a key enzyme for the creation of the sphingosine 1-phosphate (S1P) gradient—was the only putative Q612H-EVI5 binding partner that was never detected in the other variants' pull-downs in all the three independent proteomic experiments we performed. Thus, SGPL1 represented a very suitable candidate to be independently validated by western blot. After FLAG-immunoprecipitation of EVI5 complexes, samples were separated by SDS–PAGE and probed with an antibody specific for SGPL1. Only in the Q612H-EVI5 lane, a band was detected at 60 kDa, which is compatible with the molecular weight of SGPL1 (Fig. 2C). We also performed the reciprocal CoIP experiment by pulling down endogenous SGPL1 and probing immunoprecipitated samples with an antibody against EVI5. In this experiment, we were able to detect a signal only in cells overexpressing Q612H-EVI5 and not the other two variants (Fig. 2D). Additionally, we carried out as a positive control, CoIP experiments on the common interactor Disks large homolog 1 (DLG1), showing that the protein can be detected after specific immunoprecipitation form cells transfected with any of the three EVI5 variants (Supplementary Material, Fig. S4).
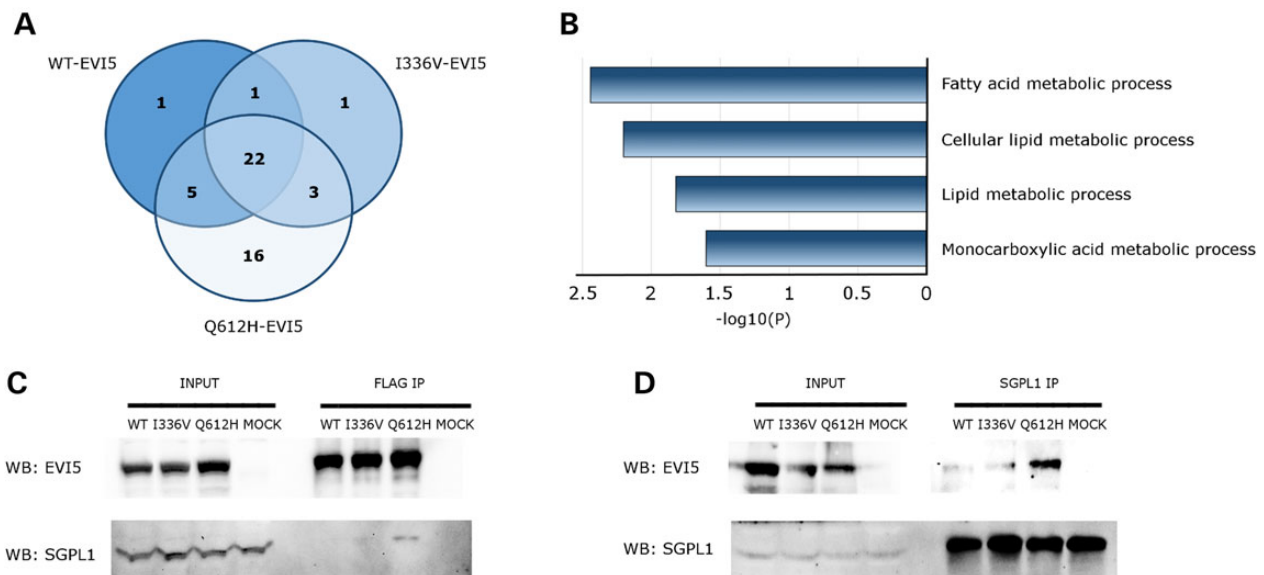
Figure 2. SGPL1 interacts with the MS-associated Q612H-EVI5 variant. (**A**) Venn diagram showing the overlap among the sets of interactors for WT-EVI5, I336V-EVI5 and Q612H-EVI5 variants. Only the proteins seen in three independent pull-down experiments and represented by at least two unique peptides were included in the final list of binders for each variant. (**B**) All the significant GO terms (biological process) related the Q612H-EVI5 exclusive binding partners, which are showed in the graph. A significant enrichment in terms connected to lipid metabolism was identified. (**C**) Immunoprecipitation experiment to confirm the interaction between SGPL1 and Q612H-EVI5. HeLa cell lysates overexpressing the different EVI5 variants or the empty vector were incubated with FLAG-beads, and immunoprecipitated samples were probed with anti-SGPL1 antibody by western blot. A signal for SGPL1 can be appreciated only in the Q612H-EVI5 lane. (**D**) Reciprocal co-immunoprecipitation confirms the interaction. HeLa cell lysates overexpressing the different EVI5 variants were incubated with beads conjugated to an SGPL1 antibody. After immunoprecipitation, samples were probed with an antibody specific for EVI5 by western blot. A positive signal can be seen only in the Q612H-EVI5 lane. Images are representative of three independent experiments.

Discussion

In this study, we tried to test the hypothesis that an MS-associated exonic SNP in the *EVI5* gene affects the complex interactions between its protein product and other cellular proteins. Immunoprecipitation followed by proteomic profiling allowed us to define groups of common and selective interactors to EVI5 allelic variants. Among the pool of 22 common EVI5-binding partners, we found the structurally related EVI5-like protein (EVI5L), which is not surprising as EVI5 is known to form dimers through the CC domain (24). Although the previously reported binder tubulin (24) is not in the final list of binders, tubulin-β-6 was detected as a binding partner in two independent experiments. Additionally, two microtubules-associated proteins (cytoskeleton-associated protein 4 and kinesin-like protein KIF26A) were also identified. On the contrary, we did not detect the other validated interactor rab11 (20). However, we found nine proteins associated with intracellular vesicles; this is an interesting finding because vesicle trafficking is one of the main roles of rab11 proteins (25,26). Moreover, the protein rab10 was identified in two independent pull-downs. This other member of the Rab family was shown to interact with EVI5L *in vitro* (27) and was found juxtaposed to EVI5 in vesicles (28). The stringency of our IP protocol and analytical pipeline might explain the discrepancies with previous data although we were still able to capture the biological functions related to these proteins through other interactors.

It is noteworthy that we identified the HLA class I, A-68 alpha chain as a binder of WT-EVI5 and I336V-EVI5, though it did not pass quality control filters for the Q612H susceptibility variant. This finding is consistent with previous studies on HeLa cells showing they carry the *HLA-A*68:02* allele (29). HLA class I molecules principally present short peptides generated from degradation of

cytosolic proteins to CD8+ T cells (30). However, the fact that a tagging epitope was used for pull-down experiments suggests that we captured not the expected self-peptide-HLA class I interaction but a possible regulatory interaction with the full-length EVI5. This second type of binding might occur through the HLA-A cytoplasmic domain as the rest of the molecule faces the endoplasmic reticulum (ER) lumen or the extracellular environment. The roles of the cytoplasmic tail have been linked to endocytosis, internalization, degradation and ER retrieval of HLA-A proteins (31). Thus, the interaction with EVI5 might modulate the immune response through the regulation of HLA-A turnover. Particularly with regards to MS, several HLA class I alleles have been associated with MS protection although the molecular basis of their activity has not been fully elucidated (32–34). Moreover, a genetic interaction between EVI5 and the MS-associated *HLA-DRB1* haplotype has been reported in both Europeans and African Americans, corroborating the idea of a possible functional role for EVI5 in the immune response (9,35).

With regards of the risk variant-specific interactors, we were able to highlight a possible functional correlation with lipid biosynthesis through pathway analysis. This finding is particularly relevant in the context of MS pathogenesis as alterations in lipid metabolism have been observed in MS patients. Specifically, a shift in lipid composition from a higher phospholipid and lower sphingolipid content was detected in the normal appearing white and gray matter (36). Aberrant lipid peroxidation may occur in brain tissues of MS patients as well (37). Additionally, there is evidence of increased polyunsaturated lipid content and phosphatidylserine linked to myelin fluidity and vesiculation in the murine MS model experimental autoimmune encephalomyelitis (EAE) (38).

Lipid metabolism plays an additional role in MS and other autoimmune diseases by modulating immune cell trafficking.

In particular, the S1P pathway has been shown to regulate the egress of T and B cells from lymphoid organs after immune activation (39). An S1P gradient exists between the circulation and peripheral tissues, and a marked difference in S1P concentrations represents the driving force for lymphocyte migration (40). Such gradient is generated through SGPL1, which catalyzes the irreversible decomposition of S1P by a retro-aldol fragmentation that yields hexadecanaldehyde and phosphoethanolamine (41). SGPL1 activity counterbalances S1P synthesis via sphingosine kinases and maintains S1P at low concentrations in most organs (42). Pharmacologic inhibition of SGPL1 catalytic activity was shown to disrupt the S1P gradient and sequester lymphocytes within lymphoid organs (40). The immunosuppressant effects of SGPL1 inhibition are also sufficient to prevent an autoimmune response as both genetic and pharmacologic ablation of SGPL1 fully protects mice from developing EAE (43,44).

We identified SGPL1 as the most robust selective interactor of Q612H-EVI5 because it was the only protein co-eluted exclusively with the risk variant in all our proteomic experiments. Conversely, the other Q612H-EVI5 interactors were observed at least once in the pull-downs with the two non-risk variants, suggesting that the Q612H substitution in the CC domain may promote both qualitative and subtle quantitative differences in protein–protein interactions. Considering that the binding of two proteins is mediated by a set of non-covalent electrostatic interactions, it is likely that a change in surface charge could affect the strength of such interaction.

In the long term, the implications of SGPL1–EVI5 interaction may be broad considering that the S1P axis has been one of the preferred druggable targets for therapeutic intervention against MS. For instance, fingolimod, the first oral disease-modifying drug recently approved by the FDA, prevents lymphocytic infiltration in the CNS by acting as an S1P receptor agonist (45). The possibility that the rs11808092 SNP could modulate responsiveness to fingolimod treatment is suggestive, and it will be evaluated as soon as the number of treated patients will reach enough power for association studies.

From a mechanistic point, we currently ignore how the interaction with Q612H-EVI5 affects the S1P pathway. Preliminary experiments in which we probed SGPL1 enzymatic activity using a fluorogenic SGPL1 substrate failed to show significant differences in cells overexpressing the Q612H variant (Supplementary Material, Fig. S5). Several reasons could account for this result. For instance, there is the possibility that the effects on SGPL1 activity may be small enough to fall below the resolution of our experimental system, especially considering that in polygenic diseases like MS the contribution of each risk variant to overall risk is modest (46). Moreover, while the HeLa cell line is a suitable and widely used vessel for proteomic screenings owing to high transfectability (47,48), it might represent a sub-optimal model for the functional characterization of this interaction, giving that we still ignore in which cytotype it occurs *in vivo*. Further studies employing transgenic mouse models carrying the different EVI5 variants will be required to fill this gap. Alternatively, we could speculate that other SGPL1 functions, not directly correlated with S1P degradation, might be affected. Lastly, it might also be possible that the interaction affects other EVI5 functions rather than SGPL1.

In conclusion, although we do not rule out the possibility that secondary independent signals may exist targeting different biological effects, we have proposed *EVI5* as a strong candidate disease risk gene in the 1p22.1 MS-risk locus. Moreover, we have performed the first comprehensive proteomic analysis of the EVI5 protein interactome and shown the functional effects of an SNP associated with MS risk by highlighting differences in EVI5-binding partners. Among them, we confirmed the novel interaction with the SGPL1 protein, which is relevant to the pathogenic process of MS.

Materials and Methods

Meta-analysis in the *EVI5* locus

An *EVI5* locus meta-analysis was conducted with 13 datasets of European descents. The region was defined as 2 Mb (1 Mb centromeric and 1 Mb telomeric) flanking the updated lead SNP rs41286801 reported in the ImmunoChip study (6). All datasets have been previously reported (11,49). We applied typical GWAS quality check filters (49), and then we imputed all datasets to the 1000 Genomes European Phase I panel using BEAGLE (50). Within each dataset, we tested the imputed dosages for association with affectation status using logistic regression, including the five first eigenvectors as covariates to correct for population stratification. Then, we penalized the standard error of the Odds Ratio (OR) for any excessive genomic inflation ($\lambda > 1$) (49). Next we applied both a fixed-effects model and random-effects model to meta-analyze across the 13 datasets. We tested for the presence of statistical heterogeneity using Cochran's Q. To identify statistically independent effects, a forward stepwise logistic regression was applied for the SNPs within 2 Mb from the most associated SNP. First, the primary SNP (with the most significant P-value) was included as a covariate in an association analysis for the remaining SNPs. This process was repeated until no SNPs reached the minimum level of significance ($P < 10^{-4}$). To assess the relevance of European meta-analysis findings, the association within the *EVI5* locus was also explored in the African American ImmunoChip dataset (803 cases and 1516 controls) (14). Parameters of LD were calculated using the applicable population dataset of the 1000 Genomes Project dataset (13). For all the analyses, PLINK v1.07 (51) and R 2.11 tools were used. Functional annotation of non-coding SNPs was performed using HaploReg tool (52).

Cell culture

HeLa cells (ATCC) were maintained in Modified Eagle's Medium (GIBCO/Invitrogen), supplemented with 10% v/v fetal bovine serum (GIBCO/Invitrogen) and antibiotics (100 IU/ml penicillin and 100 mg/ml streptomycin) at 37°C in a humidified atmosphere with 5% CO_2.

DNA constructs

Human full-length WT EVI5-coding sequence was PCR-amplified from the pENTR223.1 clone HsCD00081286 (DNASU repository) using the following primers: Forward 5′ CCCCCCGAATTCATGGT TACCAACAAAATGACTG 3′; Reverse 5′ CCCCCCCTCGAGTCAGA CAGTGGTTGAATACGA 3′. The PCR product was then double-digested with the restriction enzymes EcoRI and XhoI (New England Biolabs) and cloned into the pGL5S vector (10791; Addgene) that was linearized with the same enzymes. Constructs expressing the I336V-EVI5 and Q612H-EVI5 variants were obtained by introducing the point mutations 1006A>G and 1836G>T in the WT EVI5 sequence using the QuickChange Lightning Site-Directed Mutagenesis Kit (Stratagene). Individual clones were confirmed by Sanger sequencing.

Antibodies

The following antibodies were used: anti-FLAG M2 monoclonal antibody conjugated to agarose beads (A2220) from Sigma; anti-EVI5 polyclonal antibody (ab31269), anti-SGPL1 polyclonal antibody (ab105183) and anti-DLG1 polyclonal antibody (ab3437) from Abcam.

Immunoprecipitation assays

HeLa cells were cultured in 100-mm dishes at 90% confluence in preparation for transfection. About 10 µg of plasmid (either empty vector or EVI5 expressing) were transfected using Lipofectamine 2000 (Invitrogen) according to manufacturer's instructions. After 48 h, the cells were washed twice with cold phosphate-buffered saline and then lysed in 1 ml of lysis buffer (50 mM Tris–HCl, pH 7.5; 150 mM NaCl; 0.5% NP-40; protease and phosphatase inhibitor cocktails, Roche) for 10 min on ice. Lysates were spun at 2300g at 4°C for 5 min, and the supernatants were collected. 40 µl from each sample were saved as input controls.

For EVI5 immunoprecipitation, each sample was first pre-cleared with 40 µl of mouse IgG agarose slurry (Sigma) for 1 h at 4°C, rotating. Immunoprecipitation was then carried out by incubating each sample with 40 µl of anti-FLAG M2 Affinity Gel for 5 h at 4°C, rotating. After incubation, beads were washed with 500 µl of lysis buffer for six times at 4°C. Last wash was performed using lysis buffer without NP-40. EVI5 complexes were subsequently eluted from the beads using 50 µl of FLAG peptide (0.5 mg/ml) in lysis buffer without NP-40 for 30 min at 4°C, shaking. Supernatants were collected, and Laemnli buffer was added before boiling each sample at 95°C for 10 min. Input controls were prepared the same way.

For SGPL1 and DLH1 immunoprecipitation, 2.5 µg of polyclonal anti-SGPL1 or anti-DLG1 antibody were conjugated with 50 µl of Dynabeads Protein G (Invitrogen) according to manufacturer's instructions. Cell lysates were incubated with antibody-coupled beads for 5 h at 4°C, rotating. Beads were then washed with 500 µl of lysis buffer for five times at 4°C. To elute protein complexes, beads were resuspended in 40 µl of 2× Laemmli buffer and boiled at 95°C for 10 min.

Protein identification using reversed-phase liquid chromatography electrospray tandem mass spectrometry (LC-MS/MS)

Immunoprecipitated samples were separated by SDS–PAGE on 10% gels and subsequently stained with Gelcode Blue Stain Reagent (Thermo Scientific). Each lane was cut into five pieces and proteins within each piece were subjected to in-gel tryptic digestion. The proteins were first reduced with 10 mM dithiothreitol (Sigma) at 56°C for 1 h, followed by alkylation with 55 mM iodoacetamide (Sigma) at RT in the dark for 45 min. The samples were then incubated overnight with 100 ng trypsin (Promega) at 37°C. The peptides formed from the digestion were extracted using 50% acetonitrile and 5% formic acid and then re-suspended in 10 µl of 0.1% formic acid in water and analyzed by on-line LC-MS/MS technique. The LC separation was performed using a NanoAcquity UPLC system (Waters) on an Easy-Spray PepMap column (75 µm × 15 cm, Thermo Scientific) whereas the MS/MS analysis was performed using an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific). During the LC separation step, 0.1% formic acid in water was used as the mobile phase A and 0.1% formic acid in acetonitrile was employed as the mobile phase B. Following the initial equilibration of the column in 98% A/2% B, 5 µl of the sample was injected. A linear gradient was started with 2% B and increased to 30% B in 27 min at a flow rate of 300 nL/min, followed by an increase of 50% B in the next 2 min. The subsequent MS analysis was performed using a top six data-dependent acquisitions. The sequence includes one survey scan in the FT mode in the Orbitrap with mass resolution of 30 000 followed by six CID scans in LTQ, focusing on the first six most intense peptide ion signals whose m/z-values were not in

the dynamically updated exclusion list and their intensities were over a threshold of 1000 counts. The analytical peak lists were generated from the raw data using an in-house software, PAVA (53). The MS/MS data were searched against the UniProt database using an in-house search engine Protein Prospector. Stringent inclusion criteria were adopted to filter out the non-specific contaminants from the list of proteins generated for each EVI5 variant. First, all the proteins in common with the mock samples were removed, then all the proteins represented by only one unique peptide were removed and lastly, only the proteins consistently detected in three independent pull-downs were included.

Western blot assays

Both input and immunoprecipitated samples were separated by SDS–PAGE on 10% gels and then transferred to nitrocellulose membranes (Immobilon) at 100 V for 30 min. Membranes were then blocked with 5% milk in Tris-buffered saline supplemented with 0.05% Tween-20 (TBS-T) for 1 h at room temperature (RT). After blocking, membranes were incubated with rabbit polyclonal antibodies against EVI5, SGPL1 or DLG1 in blocking solution (1:1000) overnight at 4°C. The day after, the membranes were washed three times with TBS-T and incubated with horseradish peroxidase (HRP)-conjugated protein A (Invitrogen) in blocking solution (1:5000) for 1 h at RT. After extensive washing, membranes were incubated with Supersignal West Dura reagent (Thermo Scientific) and the chemiluminescent signals were detected using a Molecular Imager ChemiDoc XRS System equipped with Quantity One software (Biorad).

SGPL1 activity assays

SGPL1 activity was measured using the fluorogenic S1P analog 2S-ammonio-3R-hydroxy-5-((2-oxo-2H-chromen-7-yl)oxy)pentylhydrogen phosphate (Cayman Chemical) following the protocol by Bedia *et al.* (54) with small modifications. Briefly, each EVI5 variant was overexpressed for 48 h in HeLa cells and cytosolic proteins were extracted using the same protocol as the proteomic experiments. 75 µl of cell lysate (lysis buffer for negative controls) were incubated in a black 96-well plate with 15 µl of substrate solution (2.5 µl of a 5 mM stock solution in slightly acidic methanol solubilized in 15 µl of potassium phosphate buffer 0.5 m pH 7.4; 125 µm final concentration), 25 µm Na_3VO_4 (5 µl, 0.5 mM) and 0.25 mM pyridoxal phosphate (5 µl, 5 mM) ON at 37°C in the dark. The day after, the reaction was stopped by adding 50 µl of methanol to each well and after 2 h in the dark the fluorescence was measured using a Spectramax Gemini plate reader (Molecular Devices).

Pathway analysis

The cellular component, molecular function and biological process ontologies were annotated using the AmiGO2 tool (55). Significantly enriched GO terms (experimental only) were searched for each list of interactors identified by mass spectrometry, and Bonferroni *post hoc* test was used for multiple-test correction. Significance was considered at P-values of 0.05 or less.

Homology modeling

The tertiary structures of the TBC and CC domains were predicted using the protein homology/analogy recognition engine 2 (Phyre2) tool (56). Briefly, the primary sequences—either WT or mutated—of TBC (160aa–371aa) and CC (404aa–714aa) domains

of human EVI5 were uploaded on the Phyre2 server, and modeling prediction was run in normal mode. The crystal structures of TBC1D1 RabGAP domain (PDB id: c3qyeA) and tropomyosin (PDB id: c1c1gA) were identified as the highest scoring templates (near 100% confidence and coverage) to model the TBC and the CC domains, respectively. The PDB files with atomic coordinates generated by Phyre2 for each domain were visualized using the PyMol software on the POLYVIEW-3D server (57).

Global interactome analysis

Protein interactions obtained for EVI5 and for each of the two variants were loaded as undirected graphs (networks) in Cytoscape (58) to compare the topological features of the three interactomes. Next, the first neighbors of each node were added from a high-quality human global interactome, and topological metrics were computed for each resulting network.

Supplementary Material

Supplementary Material is available at *HMG* online.

Conflict of Interest statement. None declared.

Funding

References

1. Ciccarelli, O., Barkhof, F., Bodini, B., De Stefano, N., Golay, X., Nicolay, K., Pelletier, D., Pouwels, P.J., Smith, S.A., Wheeler-Kingshott, C.A. *et al.* (2014) Pathogenesis of multiple sclerosis: insights from molecular and metabolic imaging. *Lancet Neurol.*, **13**, 807–822.

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

3. Hauser, S.L., Chan, J.R. and Oksenberg, J.R. (2013) Multiple sclerosis: Prospects and promise. *Ann. Neurol.*, **74**, 317–327.

4. Oksenberg, J.R. (2013) Decoding multiple sclerosis: an update on genomics and future directions. *Expert Rev. Neurother.*, **13**, 11–19.

5. Oksenberg, J.R. and Barcellos, L.F. (2005) Multiple sclerosis genetics: leaving no stone unturned. *Genes Immun.*, **6**, 375–387.

6. International Multiple Sclerosis Genetics Consortium (2013) Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat. Genet.*, **45**, 1353–1360.

7. International Multiple Sclerosis Genetics Consortium (2007) Risk alleles for multiple sclerosis identified by a genomewide study. *N. Engl. J. Med.*, **357**, 851–862.

8. Hoppenbrouwers, I.A., Aulchenko, Y.S., Ebers, G.C., Ramagopalan, S.V., Oostra, B.A., van Duijn, C.M. and Hintzen, R.Q. (2008) EVI5 is a risk gene for multiple sclerosis. *Genes Immun.*, **9**, 334–337.

9. Johnson, B.A., Wang, J., Taylor, E.M., Caillier, S.J., Herbert, J., Khan, O.A., Cross, A.H., De Jager, P.L., Gourraud, P.A., Cree, B.C. *et al.* (2010) Multiple sclerosis susceptibility alleles in African Americans. *Genes Immun.*, **11**, 343–350.

10. Alcina, A., Fernandez, O., Gonzalez, J.R., Catala-Rabasa, A., Fedetz, M., Ndagire, D., Leyva, L., Guerrero, M., Arnal, C., Delgado, C. *et al.* (2010) Tag-SNP analysis of the GFI1-EVI5-RPL5-FAM69 risk locus for multiple sclerosis. *Eur. J. Hum. Genet.*, **18**, 827–831.

11. International Multiple Sclerosis Genetics Consortium and Wellcome Trust Case Control Consortium (2011) Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature*, **476**, 214–219.

12. Martin, D., Pantoja, C., Fernandez Minan, A., Valdes-Quezada, C., Molto, E., Matesanz, F., Bogdanovic, O., de la Calle-Mustienes, E., Dominguez, O., Taher, L. *et al.* (2011) Genome-wide CTCF distribution in vertebrates defines equivalent sites that aid the identification of disease-associated genes. *Nat. Struct. Mol. Biol.*, **18**, 708–714.

13. Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. *Nature*, **467**, 1061–1073.

14. Isobe, N., Madireddy, L., Khankhanian, P., Matsushita, T., Caillier, S.J., More, J.M., Gourraud, P.A., McCauley, J.L., Beecham, A.H., International Multiple Sclerosis Genetics Consortium *et al.* (2015) An ImmunoChip study of multiple sclerosis risk in African Americans. *Brain*, **138**, 1518–1530.

15. Lim, Y.S. and Tang, B.L. (2013) The Evi5 family in cellular physiology and pathology. *FEBS Lett.*, **587**, 1703–1710.

16. Frasa, M.A., Koessmeier, K.T., Ahmadian, M.R. and Braga, V. M. (2012) Illuminating the functional and structural repertoire of human TBC/RABGAPs. *Nat. Rev. Mol. Cell Biol.*, **13**, 67–73.

17. Hirano, T. (2006) At the heart of the chromosome: SMC proteins in action. *Nat. Rev. Mol. Cell Biol.*, **7**, 311–322.

18. Eldridge, A.G., Loktev, A.V., Hansen, D.V., Verschuren, E.W., Reimann, J.D. and Jackson, P.K. (2006) The evi5 oncogene regulates cyclin accumulation by stabilizing the anaphase-promoting complex inhibitor emi1. *Cell*, **124**, 367–380.

19. Faitar, S.L., Sossey-Alaoui, K., Ranalli, T.A. and Cowell, J.K. (2006) EVI5 protein associates with the INCENP-aurora B kinase-survivin chromosomal passenger complex and is involved in the completion of cytokinesis. *Exp. Cell Res.*, **312**, 2325–2335.

20. Westlake, C.J., Junutula, J.R., Simon, G.C., Pilli, M., Prekeris, R., Scheller, R.H., Jackson, P.K. and Eldridge, A.G. (2007) Identification of Rab11 as a small GTPase binding protein for the Evi5 oncogene. *Proc. Natl Acad. Sci. USA*, **104**, 1236–1241.

21. Dabbeekeh, J.T., Faitar, S.L., Dufresne, C.P. and Cowell, J.K. (2007) The EVI5 TBC domain provides the GTPase-activating protein motif for RAB11. *Oncogene*, **26**, 2804–2808.

22. Strauss, H.M. and Keller, S. (2008) Pharmacological interference with protein-protein interactions mediated by coiled-coil motifs. *Handb. Exp. Pharmacol.*, **186**, 461–482.

23. Kocher, T. and Superti-Furga, G. (2007) Mass spectrometry-based functional proteomics: from molecular machines to protein networks. *Nat. Methods*, **4**, 807–815.

24. Faitar, S.L., Dabbeekeh, J.T., Ranalli, T.A. and Cowell, J.K. (2005) EVI5 is a novel centrosomal protein that binds to alpha- and gamma-tubulin. *Genomics*, **86**, 594–605.

25. Stenmark, H. (2009) Rab GTPases as coordinators of vesicle traffic. *Nat. Rev. Mol. Cell Biol.*, **10**, 513–525.

26. Hutagalung, A.H. and Novick, P.J. (2011) Role of Rab GTPases in membrane traffic and cell physiology. *Physiol. Rev.*, **91**, 119–149.

27. Itoh, T., Satoh, M., Kanno, E. and Fukuda, M. (2006) Screening for target Rabs of TBC (Tre-2/Bub2/Cdc16) domain-containing proteins based on their Rab-binding activity. *Genes Cells*, **11**, 1023–1037.

28. Laflamme, C., Assaker, G., Ramel, D., Dorn, J.F., She, D., Maddox, P.S. and Emery, G. (2012) Evi5 promotes collective cell migration through its Rab-GAP activity. *J. Cell Biol.*, **198**, 57–67.

29. Johnson, D.R., Biedermann, B.C. and Mook-Kanamori, B. (2000) Rapid cloning of HLA class I cDNAs by locus specific PCR. *J. Immunol. Methods*, **233**, 119–129.

30. Trowsdale, J. and Knight, J.C. (2013) Major histocompatibility complex genomics and human disease. *Annu. Rev. Genomics Hum. Genet.*, **14**, 301–323.

31. Boyle, L.H., Gillingham, A.K., Munro, S. and Trowsdale, J. (2006) Selective export of HLA-F by its cytoplasmic tail. *J. Immunol.*, **176**, 6464–6472.

32. Fogdell-Hahn, A., Ligers, A., Gronning, M., Hillert, J. and Olerup, O. (2000) Multiple sclerosis: a modifying influence of HLA class I genes in an HLA class II associated autoimmune disease. *Tissue Antigens*, **55**, 140–148.

33. Friese, M.A., Jakobsen, K.B., Friis, L., Etzensperger, R., Craner, M.J., McMahon, R.M., Jensen, L.T., Huygelen, V., Jones, E.Y., Bell, J.I. *et al.* (2008) Opposing effects of HLA class I molecules in tuning autoreactive CD8+ T cells in multiple sclerosis. *Nat. Med.*, **14**, 1227–1235.

34. Patsopoulos, N.A., Barcellos, L.F., Hintzen, R.Q., Schaefer, C., van Duijn, C.M., Noble, J.A., Raj, T., International Multiple Sclerosis Genetics Consortium, ANZgene Consortium, Gourraud, P.A. *et al.* (2013) Fine-mapping the genetic association of the major histocompatibility complex in multiple sclerosis: HLA and non-HLA effects. *PLOS Genet.*, **9**, e1003926.

35. ANZgene Consortium (2009) Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nat. Genet.*, **41**, 824–828.

36. Wheeler, D., Bandaru, V.V., Calabresi, P.A., Nath, A. and Haughey, N.J. (2008) A defect of sphingolipid metabolism modifies the properties of normal appearing white matter in multiple sclerosis. *Brain*, **131**, 3092–3102.

37. Carlson, N.G. and Rose, J.W. (2006) Antioxidants in multiple sclerosis: do they have a role in therapy? *CNS Drugs*, **20**, 433–441.

38. Ohler, B., Graf, K., Bragg, R., Lemons, T., Coe, R., Genain, C., Israelachvili, J. and Husted, C. (2004) Role of lipid interactions in autoimmune demyelination. *Biochim. Biophys. Acta*, **1688**, 10–17.

39. Cyster, J.G. and Schwab, S.R. (2012) Sphingosine-1-phosphate and lymphocyte egress from lymphoid organs. *Annu. Rev. Immunol.*, **30**, 69–94.

40. Schwab, S.R., Pereira, J.P., Matloubian, M., Xu, Y., Huang, Y. and Cyster, J.G. (2005) Lymphocyte sequestration through S1P lyase inhibition and disruption of S1P gradients. *Science*, **309**, 1735–1739.

41. Bagdanoff, J.T., Donoviel, M.S., Nouraldeen, A., Tarver, J., Fu, Q., Carlsen, M., Jessop, T.C., Zhang, H., Hazelwood, J., Nguyen, H. *et al.* (2009) Inhibition of sphingosine-1-phosphate lyase for the treatment of autoimmune disorders. *J. Med. Chem.*, **52**, 3941–3953.

42. Chi, H. (2011) Sphingosine-1-phosphate and immune regulation: trafficking and beyond. *Trends Pharmacol. Sci.*, **32**, 16–24.

43. Billich, A., Baumruker, T., Beerli, C., Bigaud, M., Bruns, C., Calzascia, T., Isken, A., Kinzel, B., Loetscher, E., Metzler, B. *et al.* (2013) Partial deficiency of sphingosine-1-phosphate lyase confers protection in experimental autoimmune encephalomyelitis. *PLOS ONE*, **8**, e59630.

44. Weiler, S., Braendlin, N., Beerli, C., Bergsdorf, C., Schubart, A., Srinivas, H., Oberhauser, B. and Billich, A. (2014) Orally active 7-substituted (4-benzylphthalazin-1-yl)-2-methylpiperazin-1-yl]nicotinonitriles as active-site inhibitors of sphingosine 1-phosphate lyase for the treatment of multiple sclerosis. *J. Med. Chem.*, **57**, 5074–5084.

45. Kappos, L., Radue, E.W., O'Connor, P., Polman, C., Hohlfeld, R., Calabresi, P., Selmaj, K., Agoropoulou, C., Leyk, M., Zhang-Auberson, L. *et al.* (2010) A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *N. Engl. J. Med.*, **362**, 387–401.

46. Didonna, A. and Oksenberg, J.R. (2015) Genetic determinants of risk and progression in multiple sclerosis. *Clin. Chim. Acta*, **449**, 16–22.

47. Holaska, J.M. and Wilson, K.L. (2007) An emerin "proteome": purification of distinct emerin-containing complexes from HeLa cells suggests molecular basis for diverse roles including gene regulation, mRNA splicing, signaling, mechanosensing, and nuclear architecture. *Biochemistry*, **46**, 8897–8908.

48. Isabelle, M., Moreel, X., Gagne, J.P., Rouleau, M., Ethier, C., Gagne, P., Hendzel, M.J. and Poirier, G.G. (2010) Investigation of PARP-1, PARP-2, and PARG interactomes by affinity-purification mass spectrometry. *Proteome Sci.*, **8**, 22.

49. Patsopoulos, N.A., Bayer Pharma MS Genetics Working Group, Steering Committees of Studies Evaluating IFNβ-1b and CCR1-Antagonist, ANZgene Consortium, GeneMSA, International Multiple Sclerosis Genetics Consortium and de Bakker, P.I.W. (2011) Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Ann. Neurol.*, **70**, 897–912.

50. Browning, S.R. and Browning, B.L. (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.*, **81**, 1084–1097.

51. 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. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.

52. Ward, L.D. and Kellis, M. (2012) HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucl. Acid Res.*, **40**, D930–D934.

53. Guan, S., Price, J.C., Prusiner, S.B., Ghaemmaghami, S. and Burlingame, A.L. (2011) A data processing pipeline for mammalian proteome dynamics studies using stable isotope metabolic labeling. *Mol. Cell Proteomics*, **10**, M111010728.

54. Bedia, C., Camacho, L., Casas, J., Abad, J.L., Delgado, A., Van Veldhoven, P.P. and Fabrias, G. (2009) Synthesis of a fluorogenic analogue of sphingosine-1-phosphate and its use to determine sphingosine-1-phosphate lyase activity. *Chembiochem*, **10**, 820–822.

55. Carbon, S., Ireland, A., Mungall, C.J., Shu, S., Marshall, B., Lewis, S. and Ami, G.O.H. and Web Presence Working, G. (2009) AmiGO: online access to ontology and annotation data. *Bioinformatics*, **25**, 288–289.

56. Kelley, L.A. and Sternberg, M.J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.*, **4**, 363–371.

57. Porollo, A. and Meller, J. (2007) Versatile annotation and publication quality visualization of protein complexes using POLYVIEW-3D. *BMC Bioinformatics*, **8**, 316.

58. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.