ORIGINAL ARTICLE

**IL12A, MPHOSPH9/CDK2AP1 and RGS1 are novel multiple sclerosis susceptibility loci**

The International Multiple Sclerosis Genetics Consortium (IMSGC)*

A recent meta-analysis identified seven single-nucleotide polymorphisms (SNPs) with suggestive evidence of association with multiple sclerosis (MS). We report an analysis of these polymorphisms in a replication study that includes 8,085 cases and 7,777 controls. A meta-analysis across the replication collections and a joint analysis with the discovery data set were performed. The possible functional consequences of the validated susceptibility loci were explored using RNA expression data. For all of the tested SNPs, the effect observed in the replication phase involved the same allele and the same direction of effect observed in the discovery phase. Three loci exceeded genome-wide significance in the joint analysis: RGS1 (P value = 3.55 x 10^{-5}), IL12A (P = 3.08 x 10^{-6}) and MPHOSPH9/CDK2AP1 (P = 3.96 x 10^{-5}). The RGS1 risk allele is shared with celiac disease (CD), and the IL12A risk allele seems to be protective for celiac disease. Within the MPHOSPH9/CDK2AP1 locus, the risk allele correlates with diminished RNA expression of the cell cycle regulator CDK2AP1; this effect is seen in both lymphoblastic cell lines (P = 1.18 x 10^{-5}) and in peripheral blood mononuclear cells from subjects with MS (P = 0.01). Thus, we report three new MS susceptibility loci, including a novel inflammatory disease locus that could affect autoreactive cell proliferation.

**Keywords:** multiple sclerosis; susceptibility loci; eQTL

Introduction

Multiple sclerosis (MS) is a chronic inflammatory disease of the central nervous system associated with demyelination, which is thought to have an underlying autoimmune etiology. Increasing evidence indicates that disease onset is associated with environmental factors interacting with an underlying genetic susceptibility. The past 3 years have seen the discovery of susceptibility loci (such as CD58, CLEC16A, IL2RA, IL7RA) outside of the Major Histocompatibility Complex that exceed the generally accepted threshold of genome-wide significance (P value < 5 x 10^{-8}) and have been validated in at least two independent publications. Several other loci have met this level of significance but await further replication.

Recently, we reported the results of a meta-analysis of genome-wide association studies that included 2624 MS subjects and 7220 control subjects. The replication phase of this study was performed in an independent set of 2214 MS subjects and 2116 controls, and the combined analysis led to the validation of three new loci that met genome-wide significance: TNFRSF1A, IRF8 and CD6. Furthermore, this analysis highlighted seven loci with suggestive evidence of association to MS: CXCR4 (rs882300, combined P value = 1.37 x 10^{-5}), IL12A (rs4680534, P value = 5.58 x 10^{-5}), MPHOSPH9/CDK2AP1 (rs1790100, P value = 7.21 x 10^{-7}), OLIG3/TNFAIP3 (rs9321619, P value = 1.71 x 10^{-6}), PTGER4 (rs6896969, P value = 2.40 x 10^{-5}), RGS1 (rs2760524, P value = 9.77 x 10^{-5}) and ZMIZ1 (rs1250540, P value = 1.59 x 10^{-6}).

In this study, we report an attempt to validate these seven putative MS susceptibility loci by genotyping the implicated polymorphisms in new samples and then combining the results with those of the published meta-analysis and its replication effort. Subsequently, we go on to explore the possible functional repercussions of three newly confirmed MS susceptibility loci.

Results

We conducted a replication study typing the single-nucleotide polymorphisms (SNPs) rs882300, rs4680534, rs1790100, rs9321619, rs6896969, rs2760524 and rs1250540 in the sample sets described in Table 1. Detailed clinical information on the included subjects is available in Supplementary Table 1. This replication effort is an extension of the formerly reported effort and involved 8085 MS patients and 7777 healthy controls. These numbers therefore include the replication sample collections from the United Kingdom and the United States of America that were used previously. The data from each sample collection were merged using a fixed-effect meta-analysis strategy, based on the observed and the expected allele dosage.

The results of our meta-analysis of the replication samples are described in Table 2. We also conducted a
Table 1  Sample data sets used in the replication study

<table>
<thead>
<tr>
<th>Data sets</th>
<th>Number of cases</th>
<th>Number of controls</th>
<th>Total</th>
<th>Genotyping platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK-1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>831</td>
<td>1030</td>
<td>1861</td>
<td>Sequenom iPLEX Gold&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>USA-1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1383</td>
<td>1086</td>
<td>2469</td>
<td>Sequenom iPLEX Gold&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Finland</td>
<td>799</td>
<td>1082</td>
<td>1881</td>
<td>Sequenom iPLEX Gold&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Germany</td>
<td>937</td>
<td>920</td>
<td>1857</td>
<td>Sequenom iPLEX Gold&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Italy</td>
<td>830</td>
<td>642</td>
<td>1472</td>
<td>TaqMan (7900 Sequence Detection System)&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Sweden</td>
<td>2103</td>
<td>1733</td>
<td>3836</td>
<td>Sequenom iPLEX Gold&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>UK-2</td>
<td>590</td>
<td>749</td>
<td>1339</td>
<td>Sequenom iPLEX Gold&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>USA-2</td>
<td>612</td>
<td>535</td>
<td>1147</td>
<td>Sequenom iPLEX Gold&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Total</td>
<td>8085</td>
<td>7777</td>
<td>15862</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>These collections correspond to the ones used in the replication phase of the published genome-wide meta-analysis study.<sup>7</sup>

<sup>b</sup>SEQUENOM Inc, San Diego, CA, USA.

<sup>c</sup>Applied Biosystems, Foster City, CA, USA.

Table 2  Results of the replication and joint analyses

<table>
<thead>
<tr>
<th>Chr</th>
<th>SNP</th>
<th>Locus</th>
<th>BP</th>
<th>A1</th>
<th>A2</th>
<th>MAF cases</th>
<th>MAF controls</th>
<th>OR</th>
<th>95%-CI</th>
<th>P&lt;sub&gt;replication&lt;/sub&gt;</th>
<th>P&lt;sub&gt;discovery&lt;/sub&gt;</th>
<th>P&lt;sub&gt;joint&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>rs2760524</td>
<td>RGS1</td>
<td>190797171</td>
<td>A</td>
<td>G</td>
<td>0.150</td>
<td>0.169</td>
<td>0.87 (0.81-0.92)</td>
<td>3.20 × 10−4</td>
<td>1.07 × 10−4</td>
<td>3.55 × 10−9</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>rs4680534</td>
<td>IL12A</td>
<td>161181639</td>
<td>C</td>
<td>T</td>
<td>0.363</td>
<td>0.338</td>
<td>1.11 (1.05-1.16)</td>
<td>2.81 × 10−5</td>
<td>6.80 × 10−5</td>
<td>3.08 × 10−8</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>rs1790100</td>
<td>MPHOSPH9/CDK2AP1</td>
<td>122222678</td>
<td>G</td>
<td>T</td>
<td>0.240</td>
<td>0.218</td>
<td>1.10 (1.04-1.16)</td>
<td>3.95 × 10−4</td>
<td>2.74 × 10−4</td>
<td>3.96 × 10−6</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>rs882300</td>
<td>CXCR4</td>
<td>136692725</td>
<td>T</td>
<td>C</td>
<td>0.431</td>
<td>0.460</td>
<td>0.91 (0.87-0.96)</td>
<td>9.29 × 10−5</td>
<td>5.17 × 10−4</td>
<td>4.34 × 10−7</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>rs6896969</td>
<td>PTGER4</td>
<td>40460183</td>
<td>A</td>
<td>C</td>
<td>0.383</td>
<td>0.400</td>
<td>0.94 (0.88-1.00)</td>
<td>4.11 × 10−5</td>
<td>1.44 × 10−7</td>
<td>2.80 × 10−4</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>rs1250540</td>
<td>ZMIZ1</td>
<td>8076013</td>
<td>G</td>
<td>A</td>
<td>0.374</td>
<td>0.372</td>
<td>1.05 (0.99-1.11)</td>
<td>1.38 × 10−4</td>
<td>9.89 × 10−6</td>
<td>2.31 × 10−4</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>rs9321619</td>
<td>OLG3/TNFAIP3</td>
<td>137916101</td>
<td>G</td>
<td>A</td>
<td>0.464</td>
<td>0.478</td>
<td>0.94 (0.90-0.99)</td>
<td>2.29 × 10−2</td>
<td>9.34 × 10−4</td>
<td>2.85 × 10−4</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: A1, minor allele; A2, major allele; BP, physical position of the SNP in build 36; chr, chromosome; MAF, minor allele frequency; OR, odds ratio; P<sub>discovery</sub>, P-value from the previous genome-wide meta-analysis study<sup>7</sup> (discovery phase); P<sub>replication</sub>, P-value obtained combining the results from the discovery and replication phase; P<sub>joint</sub>, P-value obtained pooling together the results from all the replication samples; SNP, single-nucleotide polymorphism.

The MAF in cases and controls was calculated as average allele frequency of the minor allele weighted for the sample size of each replication collection. All the above results are based on a fixed-effect model. The SNPs are listed in order of decreasing significance at the joint analysis.

joint analysis, combining the results from the extended replication effort with the results of the discovery phase that was conducted through a genome-wide scan<sup>7</sup> (Table 2). For each minor allele, the direction of the association in the replication analysis is consistent with the results of our previous genome-wide meta-analysis.<sup>7</sup> Furthermore, six of the seven SNPs exceed a nominal (P value < 0.05) threshold of significance in the replication analysis. This observation is unlikely to occur by chance and suggests that most, if not all, of these seven loci will eventually validate as susceptibility loci. At the conclusion of this current replication effort, we have found evidence of association at a genome-wide level of significance (P value < 5 × 10<sup>−8</sup>) in three loci: RGS1 (rs2760524, joint P value = 3.55 × 10<sup>−8</sup>), IL12A (rs4680534, joint P value = 3.08 × 10<sup>−8</sup>) and MPHOSPH9/CDK2AP1 (rs1790100, joint P value = 3.96 × 10<sup>−8</sup>). The genetic effect of these three loci in the different sample strata is shown in Figures 1–3. The plots summarizing the meta-analysis results for the remaining four loci are presented in the Supplementary Information (Supplementary Figures 1–4).

The rs2760524<sup>a</sup> susceptibility allele (odds ratio (OR) = 0.87, 95% CI = 0.81-0.92 in the replication phase) is located on chromosome 1q31 and maps 15 kb distal to the 5' end of RGS1. This locus has been recently shown to be associated with susceptibility to celiac disease (CD).<sup>6</sup> rs2760524 is in strong linkage disequilibrium (LD) with the reported susceptibility SNP for CD (rs2816316, r<sup>2</sup> = 0.86 in HapMap 60 unrelated Caucasian individuals of Northern and Western European origin (CEU) samples), and the major allele is associated with susceptibility in both diseases. Further investigations are necessary to understand which one of the two polymorphisms is the better marker and whether the same causal variant affects susceptibility to both diseases.

The rs4680534<sup>c</sup> allele associated with increased risk of MS (OR = 1.11, 95% CI = 1.05–1.16 in the replication phase) maps ~8 kb distal to the 5' end of the IL12A gene, which is located on chromosome 3q25.33-q26. For this locus there is evidence of association not only with CD but also with primary biliary cirrhosis.<sup>8,9</sup> Our SNP of interest is in modest LD (r<sup>2</sup> = 0.48 in HapMap CEU) with rs9811792, one of the susceptibility SNPs for CD; these two SNPs are located in the same LD block immediately 5' of IL12A. Interestingly, the effect of the rs4680534<sup>c</sup> allele seems to be different in these two diseases: the risk allele for MS seems to be protective in CD and vice
versa. On the other hand, rs4680534 is not in LD with the susceptibility alleles for primary biliary cirrhosis (rs6441286, \( r^2 = 0.02 \) in HapMap CEU; rs574808, \( r^2 = 0.001 \) in HapMap CEU); hence, these signals of association seem to be distinct.

The third locus that met genome-wide significance in the joint analysis contains the rs1790100 G susceptibility allele (OR = 1.10, 95% CI = 1.04–1.16 in the replication phase). This polymorphism is located in intron 12 of the MPHOSPH9 gene on chromosome 12q24.31. This region has not been previously identified as a susceptibility locus for inflammatory diseases, and little is known about it.

The remaining four loci have some evidence of replication but do not yet reach the threshold of genome-wide significance. These results could be influenced by the fact that, for three out of these four loci, the implicated polymorphisms were not typed in all of the sample collections because of technical problems (Supplementary Table 2), hence the available sample size was smaller compared with the one used to test the other polymorphisms. Nevertheless, they remain strong candidate MS susceptibility loci, particularly the two SNPs mapping in the PTGER4 and OLIG3/TNFAIP3 loci, which are already known to be associated with other autoimmune diseases.

We also explored the possible functional consequences of the seven polymorphisms by performing a quantitative trait analysis that correlates genotyping data with gene expression data. Specifically, we correlated genotype data from the HapMap phase II study with publicly available gene-expression data from lymphoblastic cell lines (LCLs) of 60 unrelated subjects of Northern and Western European origin (parents of the CEU trios) using an additive model. Only one of the seven SNPs, rs1790100 in the MPHOSPH9/CDK2AP1 locus, shows evidence of association in ‘cis’: the risk-associated allele correlates with lower CDK2AP1 RNA expression (Figure 4) in the CEU LCL data (linear regression \( P = 1.18 \times 10^{-9} \)). The CDK2AP1 gene lies ~40 kb telomeric to the 3' end of MPHOSPH9 and is within the block of LD containing rs1790100 (Supplementary Figure 5). The correlation of rs1790100 with RNA expression is specific to CDK2AP1: in the available data, there is no significant association with the RNA expression of the other genes located in the same LD block (Supplementary Table 3).
expression of CDK2AP1 is observed in the presence of the rs1790100 G susceptibility allele. No rs1790100 GG homozygotes were observed in this sample of LCLs. A black line denotes the mean value for each category.

Given this result, we analyzed CDK2AP1 gene expression in another publicly available data set, the ‘mRNA-by-SNP-browser’ (see URL). In this data set generated from 400 LCLs, two of the SNPs associated with the level of CDK2AP1 RNA expression are in strong LD with rs1790100, our SNP of interest (rs1727324, \(r^2 = 0.87\) in HapMap CEU). The minor alleles of these two SNPs correlate with a lower expression of CDK2AP1 (rs1727324, \(P = 1.70 \times 10^{-5}\); rs1060105, \(P = 1.60 \times 10^{-4}\)), confirming the results we obtained using the rs1790100 marker and the HapMap LCLs.

Finally, we further explore these in vitro observations derived from LCLs using ex vivo RNA data obtained from peripheral blood mononuclear cells (PBMCs) in 255 relapsing–remitting MS and clinically isolated demyelinating syndrome (CIS) subjects (Supplementary Table 4); as in LCLs, an additive model is used for this quantitative trait analysis. As some of the subjects were treated with an immunomodulatory drug at the time of sampling, we first used the Welch 2 sample t-test to compare the level of CDK2AP1 RNA expression across the treatment regimen categories (Glatiramer acetate/interferon-β/untreated). As seen in Supplementary Figure 6, there is no significant difference in the expression of CDK2AP1 RNA among individuals who were untreated \((n = 83)\), treated with Glatiramer acetate \((n = 67)\) or treated with an interferon-β formulation \((n = 105)\) at the time of sampling. Thus, we merged the RNA data from the three subject categories and correlated the CDK2AP1 expression levels with the rs1790100 genotype. The results of this ex vivo PBMC data analysis are consistent with the results of the LCL analyses: reduced CDK2AP1 expression is observed in the presence of the rs1790100 G susceptibility allele (additive model \(P = 0.01\)) (Figure 5). Adding treatment category as a covariate did not change the results \((P = 0.01)\), giving further evidence to the fact that immunomodulatory drugs do not affect the expression profile of CDK2AP1. Finally, we note that, in our ex vivo data, MS subjects who are homozygous for the risk (‘G’) allele have a substantial reduction of the CDK2AP1 expression profile compared with the other two genotype classes. We therefore tested a recessive model of inheritance, which seems to fit the data better \((P = 0.005)\). On the other hand, the LCL data are more consistent with an additive model, and additional in vitro and ex vivo data will be useful to dissect the exact nature of the effect of rs1790100 under these different conditions.

Discussion

We tested seven putative susceptibility loci for association with MS using an expanded repertoire of independent collections of case–control samples from Finland, Germany, Italy, Sweden, the United Kingdom and the United States of America. The effect observed for each polymorphism in the replication phase is in the same direction as the effect observed in the discovery phase. A joint analysis of replication and discovery data sets also provides robust evidence of association between MS susceptibility and the RGS1 (combined \(P = 3.55 \times 10^{-4}\)), IL12A (combined \(P = 3.08 \times 10^{-4}\)) and MPHOSPH9/CDK2AP1 (combined \(P = 3.96 \times 10^{-4}\)) loci. When comparing the effects of an allele across studies, the IL12A, OLIG3/TNFAIP3, PTGER4 and RGS1 loci were the most consistent: we detected no evidence of between-strata heterogeneity.
or infectious agents. Specifically, it has been observed the immune system and during responses to exogenous proteins are involved in lymphocyte migration and can meta-analysis.

population stratification in the different strata of cannot empirically assess and correct for possible susceptibility. However, we only genotyped seven SNPs across the various sample collections, and therefore we cannot empirically assess and correct for possible population stratification in the different strata of the meta-analysis.

The RGS1 gene encodes a molecule that is a member of the regulators of G-protein signaling (RGS) family. RGS proteins are involved in lymphocyte migration and can influence cell trafficking both during the development of the immune system and during responses to exogenous or infectious agents. Specifically, it has been observed that, in RGS1 knockout mice, B cells have a better adhesion to high endothelial venules in lymph nodes; they home better to lymph nodes and move more rapidly within lymph node follicles than do wild-type B cells, supporting an involvement of RGS1 in the B-cell mobility into and out of lymph nodes. We can therefore speculate that alterations in RGS1 function mediated by allelic variants could impact the migratory capability of B cells and possibly alter their recruitment to the central nervous system.

The IL12A gene encodes the IL12p35 subunit, which, together with IL12p40, forms the heterodimeric IL-12 cytokine that has a broad range of biological activities. IL-12 is an immunomodulatory cytokine secreted predominantly by monocytes and dendritic cells, and its signaling is thought to be crucial for T-helper 1 (Th1) lymphocyte differentiation; moreover, IL-12 has also been shown to stimulate interferon gamma (IFN-γ) production by T cells and NK cells and to suppress the expansion of T-helper 2 (Th2) cell clones that are thought to have anti-inflammatory properties in MS. IL-12, together with IL-23, has been strongly implicated in the pathogenesis of both experimental autoimmune encephalomyelitis (EAE) and MS, and antibodies neutralizing IL12p40 prevent clinical EAE in nonhuman primates. However, the efficacy of this category of drugs in subjects with the risk allele. In the context of MS pathogenesis, we can therefore hypothesize an involvement of this gene in the proliferation and clonal expansion of inflammatory cells. Further functional investigations of this chromosomal region are necessary to refine our understanding of its role in MS and to identify the functional consequence of the causal variant at this locus.

For the remaining loci, we find no evidence of a correlation between the genetic variants and RNA data obtained using the probes contained in the Illumina WG-6 version 1 array of the HapMap phase II experiment. Clearly, more detailed investigations of different cell types and RNA isoforms are indicated to further explore this question.

The list of accepted MS susceptibility loci at this point consists largely of loci with primarily immunological functions: CD6, CD58, CLEC16A, CYP27B1, HLA B, HLA DRB1, IL2RA, IL7R, IRF8, STAT3, TNFRSF1A and TYK2. We can now add the IL12A, MPHOSPH9/CDK2AP1 and RGS1 loci to this list. Our results are in agreement with recently published genetic studies on MS: the newly confirmed susceptibility alleles are common in the general population, but confer only a modest risk to MS. Thus, individually, these alleles have no predictive power, but aggregate measures of genetic risk seem to be robust in MS and represent a potential strategy to use this information in a clinical setting.

Several other genes have been put forth as being associated in this and other reports, but they still need to be independently validated. The best available evidence for association is for the 13q31.3, CD40, CD226, CXCR4, GPC5, PTGER4 and TNFAIP3 loci. KIF1B is an interesting locus that had strong evidence of association in the initial report, but this effect has not been observed in additional sample collections tested so far (SJ Sawcer, personal communication). Overall, the list of involved loci is evolving rapidly and will be enriched further by ongoing genome-wide scans in MS.

Recent studies have also highlighted the observation that, in addition to common variants, less common variants can have a role in susceptibility to MS. This is the case in the TNFRSF1A locus, which harbors a less common variant (frequency 0.02) with a strong
effect (OR = 1.6). It is therefore clear that we need to supplement the current efforts aimed to comprehensively characterize common variants in MS with studies exploring the role of less common and rare variants. The current generation of sequencing platforms is making such studies practical and will allow us to enrich our understanding of the genetic architecture of MS.

Materials and methods

Subjects
This study was conducted within the framework of the International Multiple Sclerosis Genetics Consortium (IMSGC). Samples were collected across different countries, namely, Finland, Germany, Italy, Sweden, United Kingdom and the United States of America. For these last two countries, the subjects were recruited in two different stages: 2214 cases and 2116 controls were initially used as a first replication attempt in our published meta-analysis paper (they correspond to the UK-1 and USA-1 data sets in Table 1). Additional individuals were then assessed to confirm the results in the current study (UK-2 and USA-2 data sets in Table 1). Specifically, the latter patient samples were ascertained at two sites within the United States of America (Brigham and Women’s Hospital in Boston and the University of California in San Francisco) and from across the United Kingdom. Unrelated controls were obtained from the same US sites and from the British 1958 Birth Cohort Study. In Germany, MS subjects and controls were recruited from the whole country by primary care physicians and neurologists. In Italy, MS subjects and controls were recruited from the whole country by primary care physicians and neurologists. In the Italian collection, most of the cases and controls were resident of Northern Italy (mainly Northwest Italy). For the majority of them, the place of birth of their grandparents was excluded from our analysis, given the unique ancestry of the Sardinian population. The Finnish and Swedish sample collections have previously been described. Summary information on the number of cases and controls recruited in each sample collection is listed in Table 1. The demographic and clinical profiles of the samples are presented in Supplementary Table 1. All subjects were of self-reported European ancestry and samples are presented in Supplementary Table 1. All controls recruited in each sample collection (Supplementary Table 2). We performed a meta-analysis across all of the replication collections assuming a per-allelic model. The minor allele was selected as the reference allele for each polymorphism, and the analyses generated the OR per copy of the reference allele. In our primary analysis, we synthesized data using a fixed-effect meta-analysis approach. Secondarily, we also evaluated a random-effects approach for our meta-analysis of the replication data. Between-study heterogeneity was assessed with Cochran’s Q-test and quantified using I² (and respective 95% confidence intervals). We also performed a joint analysis, combining the results from the extended replication effort with the results of the discovery phase that was conducted through a genome-wide scan. To combine statistics across different strata, we calculated the respective Z score in each sample collection; in this manner, we summarized both the P value in its magnitude and the direction of the effect in its sign (Z > 0 for OR > 1). Thereafter, we obtained an overall Z statistic as a weighted average of each individual Z score, and we calculated the corresponding P value. The joint Z score was calculated on the basis of an estimated effective sample size of 2000 cases and 2000 controls for the discovery phase study in subjects genotyped genome wide and the number of informative individuals typed across the different SNPs for the replication phase results. All of these statistical analyses were performed in R statistical packages (see URLs).

Lymphoblastic cell line RNA data and quantitative trait analysis
To explore the possible functional consequences of the seven selected genetic variants, we initially used the gene expression data from LCLs of 60 individuals of European ancestry (CEU subjects) genotyped in the International HapMap Project phase II. Transcript levels were obtained from a public database (see URL) in which the RNA expression of each LCL was profiled with an Illumina whole-genome expression array (Sentrix Human-6 Expression BeadChip version 1). Out of the 47 294 transcripts that were interrogated, the normalized values for 14 925 transcripts (14 072 genes) were available for a cis expression QTL (eQTL) analysis. We tested for association between SNP and expression variation using a linear regression model. To discover cis effects, only those genes found within the same block of LD as the SNP of interest were tested. Possible regulatory effects of the seven SNPs were also assessed using the ‘mRNA-by-SNP-browser’ (see URL), which contains the results of SNP/RNA expression
Mononuclear cell RNA data from subjects with MS
We used ex vivo data to explore the association of rs1790100 and CDK2AP1 RNA expression. Between July 2002 and October 2007, PBMC samples were collected from relapsing-remitting MS and CIS subjects, as part of the comprehensive longitudinal investigation on MS at the Brigham and Women’s Hospital.43 CIS subjects have the clinical event and evidence of disease activity in a paracrine measure such as MRI. Nevertheless, these two clinical categories share common pathophysiological events and are also treated in the same manner in the clinical environment. This is why for this analysis we pooled the RNA data from relapsing-remitting MS and CIS individuals. Some of these subjects were treated with an immunomodulatory drug (Glatiramer acetate or interferon-β) at the time of recruitment. Given the limited number of subjects, we pooled all individuals receiving one of the interferon-β formulations (IFNb1a IM, IFNb1a SC or IFNb1b SC) into a single interferon-β category. None of these subjects were on combination therapy for MS at the time of sampling, PBMC isolation, RNA extraction and quantification and quality control steps are described in detail in a previous publication; 3 RNA expression levels.

From our collection, 255 subjects had a CDK2AP1 RNA expression profile and also had a genotype for rs1790100. The quantitative trait analysis module implemented in PLINK v1.06 was used for the analysis of expression data. A Welch 2 sample t-test was applied to compare the treatment regimen categories (Glatiramer acetate/interferon-β/untreated) across samples for the CDK2AP1 RNA expression levels.

Conflict of interest
The authors declare no conflict of interest.

Acknowledgements
We are grateful to the patients and the healthy controls for their participation in this study. PLD is a Harry Weaver Neuroscience Scholar Award Recipient of the National MS Society (NMSS). DAH is a Jacob Javits Scholar of NIH. SD is supported by a FISM grant (2008/R/11). BF and IR are authors on behalf of REFGENSEP, a national French clinical and research network funded by INSERM, ARSEP and AFM. The International MS Genetics Consortium is supported by R01NS049477 and by the National Multiple Sclerosis Society. This work was also supported by the Medical Research Council (G0700061) and by the Cambridge NIHR Biomedical Research Centre. The sampling and analysis of the Swedish cohorts have received grant support from The Swedish Research Council, the EU fp6 program neuromyelitis optica (LSHM-CT-2005-018637), as well as from the Bibbi and Niels Jensens Foundation, The Montel Williams Foundation and the Söderberg Foundation.

Web Resources

The URLs for data presented herein are as follows:
- GENe Expression VARIation, http://www.sanger.ac.uk/humgen/genevar/
- mRNA-by-SNP-browser’ v1.0.1, http://www.sph.umich.edu/csg/liang/asthma/

The Gene Expression Omnibus accession number for the RNA data obtained from PBMCs and reported in this paper is GSE16214.

References

Genes and Immunity

Appendix

Members of the International Multiple Sclerosis Genetics Consortium

F Esposito1,2, NA Patasopoulos3,4,5, S Cepok6, I Kockum7,8, Members of the International Multiple Sclerosis Genetics Appendix

IL12A, MPHOSPH9/CDK2AP1 and RGS1
IMSGC

and PL De Jager4,5,48, International Multiple Sclerosis Genetics Consortium members.

1Department of Neurology, Scientific Institute San Raffaele, Milan, Italy; 2Institute of Experimental Neurology, Scientific Institute San Raffaele, Milan, Italy; 3Division of Genetics, Department of Medicine, Brigham & Women’s Hospital, Boston, MA, USA; 4Harvard Medical School, Boston, MA, USA; 5Program in Translational NeuroPsychiatric Genomics, Department of Neurology, Brigham & Women’s Hospital, Boston, MA, USA; 6Klinikum rechts der Isar, Technische Universität, München, Germany; 7Neuroimmunology Unit, Department of clinical neuroscience, Karolinska University Hospital, Solna, Sweden; 8Center for Molecular Medicine, Karolinska Institutet at Karolinska University Hospital, Solna, Sweden; 9Public Health Genomics Unit, National Institute for Health and Welfare, Helsinki, Finland; 10Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland; 11University of Sydney, Institute for Immunology and Allergy Research, Westmead Millennium Institute, Westmead Hospital, NSW, Australia; 12University of Newcastle, Callaghan, NSW Australia; 13Section for Experimental Neurology, Katholieke Universiteit Leuven, Leuven, Belgium; 14Université de Montréal, Montréal, Québec, Canada; 15Montreal Heart Institute, H1 T 1C8, Montréal, Québec, Canada; 16The Danish Multiple Sclerosis Research Center, Copenhagen University Hospital, Rigshospitalet, DK-2100, Copenhagen, Denmark; 17Department of Neurology, Oulu University Hospital, Oulu, Finland; 18Central Hospital of Seinäjoki, Seinäjoki, Finland; 19INSERM, UMR S975, Paris, France; 20UPMC Univ Paris 06, UMR S975, Centre de Recherche Institut du Cerveau et de la Moelle, CNRS 7225, Paris, France; 21Département de Neurologie, Pitié –Salpêtrière Hospital, AP-HP, 75651 Paris, France; 22Klinik für Neurologie, Technische Universität München, München, Deutsch; 23Institut für Humangenetik, Technische Universität München, München, Deutsch; 24Institut für Humangenetik, Helmholz Zentrum München, München, Deutsch; 25Institute of Epidemiology, Helmholtz Zentrum München-German Research Center for Environmental Health, Munich, Germany; 26Charité-Universitätsmedizin Berlin, Cecile Vogt Clinic for Neurology, Berlin, Germany; 27Department of Neurology, University Medicine Mainz, Johannes Gutenberg University, Mainz, Germany; 28Department of Medical Sciences and Interdisciplinary Research Center of Autoimmune Diseases, University of Eastern Piedmont, Novara, Italy; 29Clinica Neurologica, AOU Maggiore della Carità, Novara, Italy; 30Neurological Institute C. Mondino, IRCCS, Pavia, Italy; 31Department of Neurology, Ospedale San Giovanni Battista, Torino, Italy; 32Department of Neurology, Faculty Division Ullevål University Hospital, University of Oslo, Oslo, Norway; 33Institute of Immunology, Rikshospitalet, Oslo University Hospital, Oslo, Norway; 34Department of Neurology, Oslo University Hospital, Ullevål, Oslo, Norway; 35Institute of Basal Medical Sciences, University of Oslo, Blindern, Oslo, Norway; 36Unité de Neuroimmunologie Clinica, CEM-Cat, Hospital Universitari Vall d’Hebron, Barcelona, Spain; 37Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden; 38University of Pavia, Department of Applied Sciences, Pavia, Italy; 39University of Cambridge, Statistical Laboratory, Centre for Mathematical Sciences, Wilberforce Road, Cambridge, UK; 40Department of Neurology, University Hospital of Wales, Heath Park, Cardiff, UK; 41Human Genomics research group, Keele University, Stoke-on-Trent, UK; 42Department of Neurology, University Hospital North Staffordshire, Stoke-on-Trent, UK; 43Division of Epidemiology, School of Public Health, University of California at Berkeley, Berkeley, CA, USA; 44John P. Hussman Institute for Human Genomics, The University of Miami Miller School of Medicine, Miami, FL, USA; 45Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, TN, USA; 46Department of Neurology, University of California San Francisco, San Francisco, CA, USA; 47Massachusetts General Hospital, Boston, MA, USA; 48Program in Medical & Population Genetics, Broad Institute of Harvard University and Massachusetts Institute of Technology, Cambridge, MA, USA; 49Harvard NeuroDiscovery Center, Harvard Medical School, Boston, MA, USA; 50University of Cambridge, Department of Clinical Neuroscience, Addenbrooke’s Hospital, Cambridge, UK; 51The Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK and 52Department of Neurology, Yale University Medical School, New Haven, CT, USA.

Supplementary Information accompanies the paper on Genes and Immunity website (http://www.nature.com/gene)