ORIGINAL ARTICLE

Identification of a functional variant in the *KIF5A-CYP27B1-METTL1-FAM119B* locus associated with multiple sclerosis

Antonio Alcina,[1] Maria Fedetz,[1] Óscar Fernández,[2] Albert Saiz,[3] Guillermo Izquierdo,[4] Miguel Lucas,[5] Laura Leyva,[6] Juan-Antonio García-León,[6] María del Mar Abad-Grau,[7] Iraide Alloza,[8] Alfredo Antigüedad,[9] María J Garcia-Barcina,[10] Koen Vandenbroeck,[8,11] Jezabel Varadé,[12] Belén de la Hera,[12] Rafael Arroyo,[13] Manuel Comabella,[14] Xavier Montalban,[14] Natalia Petit-Marty,[15] Arcadi Navarro,[15,16,17] David Otaegui,[18] Javier Olascoaga,[19] Yolanda Blanco,[3] Elena Urcelay,[12] Fuencisla Matesanz[1,4]

▶ Additional supplementary files are published online only. To view these files please visit the journal online (http://dx.doi.org/10.1136/jmedgenet-2012-101085).

For numbered affiliations see end of article

Correspondence to
Dr Fuencisla Matesanzor or Dr Antonio Alcina, Instituto de Parasitologia y Biomedicina 'Lopez-Neyra'-CSIC, Avda. Conocimiento S/N. Parque Tecnologico Ciencias de la Salud, 18100 Armilla, Granada, Spain; lindo@ipb.csic.es (FM), pulgoso@ipb.csic.es (AA)

AA, MF, OF, AS, EU and FM contributed equally.

Received 29 May 2012
Revised 24 September 2012
Accepted 10 October 2012
Published Online First
17 November 2012

To cite: Alcina A, Fedetz M, Fernández Ó, *et al*. *J Med Genet* 2013, **50**, 25–33.

ABSTRACT

Background and aim Several studies have highlighted the association of the 12q13.3–12q14.1 region with coeliac disease, type 1 diabetes, rheumatoid arthritis and multiple sclerosis (MS); however, the causal variants underlying diseases are still unclear. The authors sought to identify the functional variant of this region associated with MS.

Methods Tag-single nucleotide polymorphism (SNP) analysis of the associated region encoding 15 genes was performed in 2876 MS patients and 2910 healthy Caucasian controls together with expression regulation analyses.

Results rs6581155, which tagged 18 variants within a region where 9 genes map, was sufficient to model the association. This SNP was in total linkage disequilibrium (LD) with other polymorphisms that associated with the expression levels of *FAM119B, AVIL, TSFM, TSPAN31* and *CYP27B1* genes in different expression quantitative trait loci studies. Functional annotations from Encyclopedia of DNA Elements (ENCODE) showed that six out of these rs6581155-tagged-SNPs were located in regions with regulatory potential and only one of them, rs10877013, exhibited allele-dependent (ratio A/G=9.5-fold) and orientation-dependent (forward/reverse=2.7-fold) enhancer activity as determined by luciferase reporter assays. This enhancer is located in a region where a long-range chromatin interaction among the promoters and promoter-enhancer of several genes has been described, possibly affecting their expression simultaneously.

Conclusions This study determines a functional variant which alters the enhancer activity of a regulatory element in the locus affecting the expression of several genes and explains the association of the 12q13.3–12q14.1 region with MS.

INTRODUCTION

Multiple sclerosis (MS) is a chronic autoimmune disease with a complex pathogenesis in which demyelination and neurodegeneration are the main contributors to disability.[1] Susceptibility to MS is thought to be conferred by a combination of genetic and environmental factors.[2] The best characterised region implicated in predisposition to MS is the Major Histocompatibility Complex on chromosome 6p21, specifically the *HLA-DRB1* 15:01* class II allele, but this accounts for less than 50% of MS heritability.[3] New powerful methods as gene expression profiling and genome-wide association study (GWAS) have provided evidence for new susceptibility loci implicated in MS.[4–6] These comprise over 40 loci including IL2RA, CBLB, *IL7R, CLEC16A, TNFRSF1A, PTGER4* and *IRF8*.[7–9]

For many risk loci, the association signals do not directly implicate a single gene and the causative role for candidate genes in the region can only be speculated. A striking number of loci have demonstrated genome-wide evidence for association in multiple, distinct autoimmune disorders.[10 11] One of these loci is 12q13–14 which has been associated in GWAS for rheumatoid arthritis (RA), coeliac disease (CD) and MS.[8 12–14] However, different genes have been suggested in each study based on the main associated signal. A meta-analysis of two published GWAS totalling 3393 RA cases and 12 462 healthy controls identified an association at rs1678542 localised in the *KIF5A* intronic region.[13] Another meta-analysis of two published GWAS on CD (4533 cases and 10 750 controls) and RA (5539 cases and 17 231 controls) described the association of both diseases at rs10876993 localised in the intergenic region between *B4GALNT1* and *OS9* genes.[14] Association at this locus was also described in a MS GWAS performed by the Australian and New Zealand Multiple Sclerosis Genetics Consortium in 1618 MS-cases. In this case, an associated single nucleotide polymorphism (SNP) (rs703842) was located at the 3' untranslated region (3' UTR) of the *METTL1* gene.[12] The last GWAS performed by the International Multiple Sclerosis Genetics Consortium with 10 000 MS patients also reported the association with this region at rs12368653 in the*AGAP2* gene.[8] In candidate gene studies, the *KIF5A* variant was demonstrated to be associated to MS[15] and type 1 diabetes.[16] Also rare variants in the *CYP27B1* gene have been associated with MS.[17] Other candidate-gene studies had

demonstrated association of variants located at the *CYP27B1* gene with type 1 diabetes[18] and MS.[19]

The *FAM119B* gene has been pinpointed as the causal gene of the locus since its expression has been shown to be much lower in leukocytes of MS patients carrying the MS risk allele.[20] On the other hand, since, in several studies, MS patients have lower levels of precursors of vitamin D in serum compared with controls, it has been suggested that *CYP27B1*, an enzyme implicated in Vitamin D activation, may be the causal gene.[19]

In this work we performed a fine mapping of the 12q13.3–12q14.1 region by a Tag-SNP approach to identify the polymorphism leading the association with MS. We also performed functional studies to determine the involvement of the associated MS risk variants in gene expression.

MATERIAL AND METHODS

Study subjects and SNP genotyping

The entire data set studied consisted of 2876 MS patients diagnosed following Poser's criteria[21] and 2910 age and gender matched healthy controls, mostly blood donors and staff, matched by age, gender, ethnicity and place of recruitment. Patients and control subjects were included in the study based upon written informed consent. Demographic and clinical characteristics of the sample collections have been published before[15 22] (a summary is in the online supplementary table S1). SNPs were genotyped using the iPLEX Sequenom MassARRAY platform in the Spanish National Genotyping Center's facilities at Santiago de Compostela University (http://www.cegen.org). All DNAs ware received in the Centro Nacional de Genotipado (CEGEN) centre for genotyping in 96-well plates; the concentration of double-stranded DNA was assessed using PicoGreen (Molecular Probes). As quality control, two trios of Center for the Study of Human Polymorphisms (CEPH) samples were genotyped (CEPH Na10830, Na10831, Na12147, Na10860, Na10861, Na11984). The genotypes of these samples corresponded with the ones deposited in HapMap with no detection of Mendelian inconsistencies. Additionally, a 10% of random samples were subjected to regenotyping. The resulting data were concordant with an average accuracy over 99.9%. We performed quality filtering of both samples and SNPs to ensure robust association testing.

Ethics statement

All participants provided a written informed consent to participate in this study, which was approved by the Institutional Review Board of Hospital Regional Universitario Carlos Haya of Malaga, Hospital Clinic of Barcelona, Hospital Virgen Macarena of Sevilla, Hospital de Basurto of Bilbao, Hospital Clínico San Carlos of Madrid, Hospital Universitari Vall d´ Hebron of Barcelona and Hospital Donostia of San Sebastian.

Luciferase assays

Six fragments in both orientations with potential regulatory activity, as indicated by ENCODE, containing the tagged variants, 24 inserts in total, plus a plasmid without insert, were assayed for luciferase activity. Thus, fragments I–VI were PCR amplified (the coordinates and the primer used are indicated in online supplementary table S2) using DNA from homozygous individual for each allele of rs6581155. The fragments were previously cloned in Vaccinia DNA Topoisomerase I (TOPO) T/A vector (Invitrogen) and sequenced. Then, each clone was introduced in forward and reverse orientation in pGL4.25 (luc2CP/minP) vector which contains a minimal promoter upstream of the luciferase reporter gene luc2CP. Triplicates were transfected into human Raji cells (ATCC CCL86). The Raji cell line was maintained in Roswell Park Memorial Institute (RPMI) 1640 medium, supplemented with 10% fetal calf serum, 2.0 mM glutamine, 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 1 mM streptomycin, 100 U/ml penicillin, 0.05 mM 2-mercaptoethanol. Exponentially growing cells were transfected by nucleofection Amaxa Cell Line Nucleofector Kit V (Lonza) according to manufacture protocols. Briefly, Raji cells (2×10^6) were mixed with I–VI plasmid constructs or pGL4.25 control plasmid 2 µg of DNA and 0.2 µg of DNA of Renilla luciferase reporter gene plasmid pRL-TK(Promega) for normalising transfection efficiency. The transfected cells were cultured in 12 well-plates for 24 h. Cells were harvested and washed twice in Phosphate Buffered Saline (PBS) at 4°C. Luciferase activity was evaluated using Dual-Luciferase Assay System (Promega) according to manufacturers. For analysis, lysates were thawed on ice and the protein concentration was measured by the Bradford method. Then, 40 µg of protein from all supernatants were analysed for both firefly and Renilla luciferase activities in a luminometer F12 (Berthold Detection Systems). The following equation was used to determine the normalised fold changes in activity between tested groups:

$$\text{FoldActivity} = (\text{Firefly}/\text{Renilla})\text{fromconstruct}/(\text{Firefly}/\text{Renilla})\text{fromemptyplasmid}$$

The normalised fold changes in activity from each of the three experiments are then averaged together.

Statistical analysis

Departure from Hardy-Weinberg equilibrium for all the biallelic SNP markers was tested using an exact test.[23] Raw data were analysed for comparison of allele and genotype counts using PLINK V.1.05. To avoid false-positive results due to multiple testing and considering that the SNPs analysed are not in complete disequilibrium, we applied the Benjamini-Hochberg method[24] that is robust against positive dependence and controls the false discovery rate.[25] In order to determine the effect and independence of variant association with MS, multiple logistic regression models were computed. Linkage disequilibrium (LD) patterns between SNPs were analysed with Haploview 4.2.[26] Gender was analysed for potential differential risk contribution. We used the Breslow-Day Test to evaluate the homogeneity between gender cohorts. The age of onset with respect to the genotypes was analysed by the Kruskal-Wallis test.

RESULTS

A tag-SNP association study of the region

We performed the fine mapping of the 12q13.3–12q14.1 region between coordinates 56227468 to 56469005 (B36) (figure 1). To analyse this region, 16 SNPs were chosen by pairwise tagging from the HapMap B36 Utah residents with Northern and Western European ancestry from the CEPH collection (CEU) population which captured 89 markers with $r^2\geq0.8$ (mean $r^2=0.93$) and a minor-allele frequency ≥0.1. These tag-SNPs captured the variants associated with RA, MS and CD previously reported in different GWAS (see online supplementary table S3).

Tag-SNPs were genotyped in 2895 MS patients and 2942 healthy controls, mostly blood donors, matched by age, gender, ethnicity and place of recruitment. The genotypic distribution between cases and controls is shown in table 1. The genotype frequencies were fitted to an additive model. The top associated SNP was rs6581155 (p=4.57E−7; OR=0.79; 95% CI 0.73 to 0.87). We performed a logistic regression analysis to determine

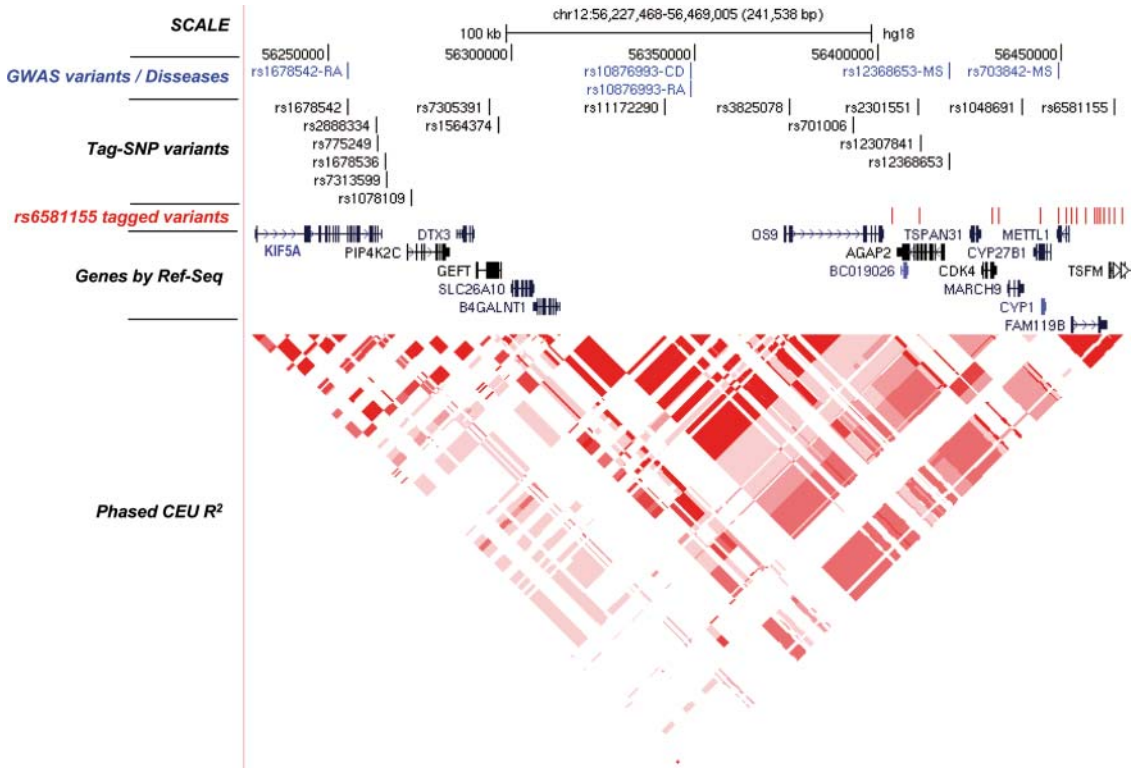

Figure 1 Scheme of the 12q13.3–12q14.1 region. The illustration shows from top to bottom the following parts as indicated : Scale and chromosomal position (B36/gh18); (in blue) diseases-associated variants as reported in different genome-wide association study (RA, rheumatoid arthritis; CD, coeliac disease; MS, multiple sclerosis); Tag-SNPs analysed in the present work; (in red) MS-associated rs6581155-tagged variants with $r^2 \geq 0.8$ and a minor-allele frequency <0.1; Ref-Seq genes present in the region; the LD structure represented by r-square for the CEU population.

the effect of further independent variants, by testing the addition of each SNP to rs6581155 and by adding this polymorphism to each one of the other Tag-SNPs (table 2). The results showed that, in addition to generate the strongest signal, rs6581155 alone was sufficient to model the association in the locus.

Stratification by gender and age at onset and genotypes of rs6581155 SNP showed absence of statistical significance.

Moreover, we genotyped the rare variant rs118204009 located in the *CYP27B* gene that has been recently associated with MS in a subgroup of 1006 patients and 1044 healthy

Table 1 Association study of the different Tag-SNPs with multiple sclerosis

	Alleles		Multiple sclerosis			Healthy controls			Allelic p value		
SNP	1	2	11	12	22	11	12	22	Uncorrected	Benjamini-Hochberg Adjusted	Per allele OR (95% CI)
rs1678542	C	G	248 (9)	1149 (41.7)	1361 (49.3)	318 (11.2)	1229 (43.3)	1291 (45.5)	0.0005	0.0017	0.87 (0.8 to 0.94)
rs2888334	C	G	134 (4.7)	946 (33.2)	1770 (62.1)	166 (5.7)	1030 (36.9)	1676 (58.4)	0.002	0.004	0.87 (0.8 to 0.95)
rs775249	T	C	181 (6.4)	1047 (36.8)	1619 (56.9)	181 (6.3)	1100 (38.2)	1601 (55.5)	0.44	0.51	0.97 (0.89 to 1.05)
rs1678536	G	C	129 (4.5)	922 (32.4)	1796 (63.1)	163 (5.7)	1013 (35.2)	1701 (59.1)	0.001	0.0023	0.86 (0.79 to 0.94)
rs7313599	G	A	752 (26.4)	1399 (49.1)	696 (24.5)	660 (23)	1430 (49.7)	784 (27.3)	0.0008	0.0021	1.13 (1.05 to 1.22)
rs1078109	A	G	6 (0.2)	366 (12.9)	2472 (86.9)	8 (0.28)	306 (10.6)	2568 (89.1)	0.018	0.026	1.2 (1.03 to 1.4)
rs7305391	A	G	128 (4.5)	908 (31.9)	1813 (63.6)	119 (4.1)	967 (33.6)	1792 (62.3)	0.5	0.537	0.97 (0.89 to 1.06)
rs1564374	G	A	477 (17.2)	1383 (49.9)	909 (32.8)	570 (20.9)	1349 (49.4)	812 (29.7)	0.00037	0.0015	0.87 (0.81 to 0.94)
rs11172290	G	A	58 (2)	735 (25.8)	2052 (72.1)	73 (2.5)	726 (25.2)	2082 (72.2)	0.79	0.79	0.99 (0.89 to 1.09)
rs3825078	A	T	176 (6.2)	1115 (39.1)	1558 (54.7)	238 (8.2)	1149 (40)	1497 (51.9)	0.003	0.006	0.88 (0.81 to 0.96)
rs701006	A	G	222 (7.8)	1225 (43)	1403 (49.2)	318 (11)	1266 (44)	1295 (45)	0.000014	0.00008	0.84 (0.78 to 0.91)
rs2301551	G	C	76 (2.7)	867 (30.5)	1897 (66.8)	130 (4.5)	967 (33.7)	1769 (61.7)	3.22E−6	0.00003	0.8 (0.73 to 0.88)
rs12307841	C	T	26 (0.9)	518 (18.1)	2311 (80.9)	28 (0.97)	469 (16.3)	2386 (82.7)	0.1097	0.15	1.11 (0.98 to 1.25)
rs12368653	G	A	430 (15.1)	1403 (49.2)	1016 (35.7)	502 (17.5)	1413 (49.2)	955 (33.3)	0.009	0.015	0.91 (0.84 to 0.98)
rs1048691	T	C	135 (4.8)	958 (33.8)	1738 (61.4)	127 (4.4)	940 (32.7)	1805 (62.8)	0.24	0.29	1.06 (0.96 to 1.15)
rs6581155	G	A	110 (4)	909 (32.7)	1757 (63.3)	173 (6.2)	1002 (36.2)	1596 (57.6)	4.57E−7	7.3E−6	0.79 (0.73 to 0.87)

Counts (percentages in brackets) are represented for each genotype.

Table 2 Regression analyses of the 16 Tag SNPs

		*Add locus to rs6581155		†Add rs6581155 to locus	
SNP	A1	p	OR (95% CI)	p	OR (95% CI)
rs1678542	C	0.2223	0.94 (0.86 to 1.04)	0.0001393	0.82 (0.74 to 0.91)
rs2888334	C	0.8487	0.99 (0.89 to 1.1)	0.0001001	0.8 (0.72 to 0.9)
rs775249	T	0.1719	0.94 (0.86 to 1.03)	3.99E−07	0.79 (0.72 to 0.87)
rs1678536	G	0.7964	0.99 (0.88 to 1.1)	9.86E−05	0.8 (0.72 to 0.9)
rs7313599	G	0.1479	1.06 (0.98 to 1.15)	6.20E−05	0.82 (0.74 to 0.9)
rs1078109	A	0.09391	1.15 (0.98 to 1.35)	1.42E−06	0.8 (0.73 to 0.88)
rs7305391	A	0.2615	0.95 (0.87 to 1.04)	6.29E−07	0.8 (0.73 to 0.87)
rs1564374	G	0.2487	0.95 (0.86 to 1.04)	0.000759	0.83 (0.74 to 0.93)
rs11172290	G	0.634	0.98 (0.88 to 1.08)	4.10E−07	0.79 (0.73 to 0.87)
rs3825078	A	0.589	1.03 (0.93 to 1.14)	1.36E−05	0.78 (0.7 to 0.87)
rs701006	A	0.4727	0.96 (0.85 to 1.08)	0.003056	0.82 (0.72 to 0.94)
rs2301551	G	0.1587	0.88 (0.75 to 1.05)	0.1069	0.88 (0.75 to 1.03)
rs12307841	C	0.6206	1.03 (0.91 to 1.17)	1.23E−06	0.8 (0.73 to 0.87)
rs12368653	G	0.5573	1.03 (0.93 to 1.14)	1.74E−05	0.78 (0.69 to 0.87)
rs1048691	T	0.795	1.01 (0.92 to 1.11)	1.80E−06	0.8 (0.72 to 0.87)
rs6581155	G	NA	NA	NA	NA

Results for models assuming additive effects.
*Denotes results adjusted for rs6581155.
†Denotes results adjusted for locus tested.
NA, not applicable.

controls.[17] We did not detect this polymorphism in any individual of the MS or control groups (data not shown).

Association of rs6581155-tagged SNPs with gene transcription levels

To determine if selected Tag-SNPs could have a functional effect on the transcription levels of the genes located in the 12q13.3–12q141 region, we first used expression quantitative trait loci (eQTL) data from Zeller *et al*[27] obtained from monocyte samples of 1490 German individuals (http://genecanvas.ecgene.net/news.php?readmore=45). There were 23 eQTLs for *FAM119B*, 6 for *AVIL*, 18 for *XRCC6BP1*, 18 for *TSFM* and 20 for *TSPAN31* genes in the whole region. The variants best correlated with expression of *FAM119B*, *AVIL*, *TSFM* and *TSPAN31*were also SNPs tagged by rs6581155 and, therefore, in high LD with it ($r^2>0.8$) (table 3). The eQTLs for the *XRCC6BP* gene were in low LD with rs6581155 ($r^2=0.18$) indicating that they were not involved in MS risk. Interestingly, the minor alleles for the rs6581155-tagged SNPs, which were protective for MS, correlated with high expression of *FAM119B* and *AVIL* but low expression of *TSFM* and *TSPAN31* (see online supplementary figure S1) although there were large differences between the correlation coefficients obtained for each gene. *FAM119B* was the gene most strongly correlated with the rs6581155-tagged SNPs in monocytes.

The eQTL data from Dixon *et al*[28] obtained in lymphoblastic cell lines (LCL) from 206 families of British descent confirmed the eQTLs previously observed for *FAM119B* and *TSFM*, although the correlation coefficients were similar for both genes in contrast to the data from Zeller *et al*.[27] The eQTLs for *FAM119B* were also reported in three additional data sets carried out in different cell types, as shown in table 3.

The data from Montgomery *et al*,[29] obtained by RNA-Seq from LCL of 60 HapMap CEU individuals, have reported a rs6581155-tagged SNP to be eQTLs for *FAM119B* and *CYP27B1*, in which the minor allele inversely correlated with high expression of these genes (correlation coefficient (r) for *CYP27B1*=−0.472, and for *FAM119B*=−0.474).

Enhancer activity

Since all the rs6581155-tagged polymorphisms were located in intronic and intergenic non-coding regions, another strategy to search for functionality was to use annotations of gene regulatory regions from ENCODE (http://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=173506627&c=chr8&g=wgEncodeReg). Six out of 18 captured SNPs were located in remarkable regions associated with histone methylation, binding sites of transcription factors or in DNAse I hypersensitive site as shown in figure 2A.

To test their potential activities as enhancers or repressors of gene expression, we cloned the fragments containing these six variants upstream of the basic promoter of a luciferase reporter system and transfected into Raji lymphoblastic B cells.

We cloned four inserts for each of the six regions carrying both alleles in forward and reverse orientation. Only region III containing variant SNP rs10877013 had allele-dependent and orientation-dependent enhancer activity. Thus, allele A versus allele G in forward orientation showed 9.5-fold higher ratio of luciferase activity (figure 2B). We observed a 2.7 times higher enhancer activity of the T allele in the reverse fragment.

TFSEARCH, Match and TESS software were used to predict potential transcription factor binding sites that could be responsible of the different allele-depending activity of fragment III. The SNP rs10877013 falls within the putative binding sites for the CCAAT/enhancer binding proteins (C/EBP) α and β. Although both alleles appear to be consistent with the consensus sequence for C/EBP TF it is possible that a base change could affect the affinity or isotype binding of the C/EBP transcription factor to the sequence.

DISCUSSION

The 12q13.3–12q14.1 region has been associated to several autoimmune diseases[8 12–16] but no causal variant has been found so far for any of them exception made for a rare variant

Table 3 rs6581155 tagged SNPs that have been reported to be eQTLs in the region

Study	Zeller *et al*[27]					Dixon *et al*[28]		Montgomery *et al*[29]		Stranger *et al*[30]	Dimas *et al*[31]		
Population	1490 German individuals					206 families of British descent		60 CEU individuals of the HapMap collection		60 individuals of the HapMap collection CEU	80 individuals of Western European origin		
Gene	AVIL	XRCC6BP1	TSFM	FAM119B	TSPAN31	FAM119B	TSFM	CYP27B1	FAM119B	FAM119B	FAM119B	FAM119B	FAM119B
Cell type			Monocytes			LCL		LCL		LCL	T-cells	LCL	Fibroblasts
Probe/Transcript	ILMN_20148	ILMN_16915	ILMN_27946	ILMN_12426	ILMN_11441	213861_s_at	212656_at	ENST00000228606	ENST00000300209	GI_24308058-S	ILMN_1723846	ILMN_1723846	ILMN_1723846
	p Value/effect					p Value/effect		p Value/effect		p Value/effect	p Value		
SNP													
rs6581155										2.1E−10/−0.57			
rs10747783						5.00E−21/−0.731	9.00E−13/0.564			7.5E−11/−0.58			
rs923829						5.00E−21/−0.731	9.00E−13/0.564			6E−11/−0.58			
rs10431552						2.80E−21/−0.745	1.00E−12/0.569			2.1E−10/−0.57			
rs2072052						7.70E−19/−689	1.00E−13/0.588			5E−10/−0.56			
rs2291617						4.10E−11/−0.62	7.60E−11/0.626			6E−11/−0.58			
rs4646536	1.7E−07/−0.021	5E−38/0.11	7.71E−19/0.055	3.34E−236/−0.52	1.4E−11/0.033					2.1E−10/−0.57			
rs724834										6E−11/−0.58			
rs2069502	1.84E−08/−0.024	4.4E−38/0.11	6.19E−17/0.049	7.78E−221/−0.5	2.29E−12/0.036	7.7E−19/−0.689	1.40E−13/0.583			5.3E−10/−0.56			
rs10877013						8.80E−20/−0.711	1.80E−12/0.56			4.1E−10/−0.56	0.000012/−0.481	0.000292/−0.407	0.0000001/−0.571
rs10877011						8.90E−14/−0.612	9.00E−13/0.6	0.00014157/−0.472	0.00013264/−0.474	9.5E−11/−0.57			
rs701008						2.30E−12/−0.563	2.30E−11/0.55			1.6E−8/−0.51			
rs8181644	1.2E−07/−0.022	3.8E−38/0.11	1.4E−19/0.057	3.9E−234/−0.52	5.5E−12/0.034					2.1E−10/−0.57			
rs10877019	4.2E−07/−0.023	6.31E−38/0.11	1.024E−190.058	8.32E−233/−0.52	1.52E−11/0.034					2.1E−10/−0.57			
rs11172335										2.1E−10/−0.57			
rs703842	1.67E−07/−0.021	2.42E−37/0.11	1.42E−18/0.054	7.5E−235/−0.52	1.41E−11/0.037	7.80E−17/−0.694	2.50E−11/0.568			2.1E−10/−0.57			
rs10877015										2.1E−10/−0.57			
rs11172333	1.11E−07/−0.023	7.31E−39/0.11	1.74E−19/0.057	2.07E−233/−0.51	5.63E−12/0.035					2.1E−10/−0.57			

Zeller *et al* Genome-Wide Expression analysis performed with Illumina HT-12 v3 BeadChip. Association significance was obtained by ANalysis Of VAriance (ANOVA).
Dixon *et al* Genome-Wide Expression analysis performed by U133 Plus 2.0 GeneChips (Affymetrix). Association analysis was obtained by fitting a simple regression model to each trait and used a variance component approach to account for correlation between different observed phenotypes within each family.
Montgomery *et al* Genome-Wide Expression analysis performed by RNA sequencing in an Illumina GAII. Association analysis and multiple testing corrections were conducted by spearman rank correlation and permuting the expression phenotype 10 000 times and summarising the extreme p value distribution for each particular exon, transcript and probe.
Stranger *et al* Genome-Wide Expression analysis performed with Sentrix Human-6 Expression BeadChip V.1, Illumina. Association significance was obtained by spearman rank correlation with 10 000 permutations of the expression phenotypes.
Dimas *et al* Gene expression was performed by commercial whole-genome expression array (WG-6 V.3 Expression BeadChip, Illumina). Association analysis and multiple testing corrections were conducted by spearman rank correlation. Thresholds for each gene were assigned after 10 000 permutations of expression values relative to genotypes where the haplotypic (LD) structure was maintained in each interaction.
eQTL, expression quantitative trait loci; LCL, lymphoblastic cell lines.

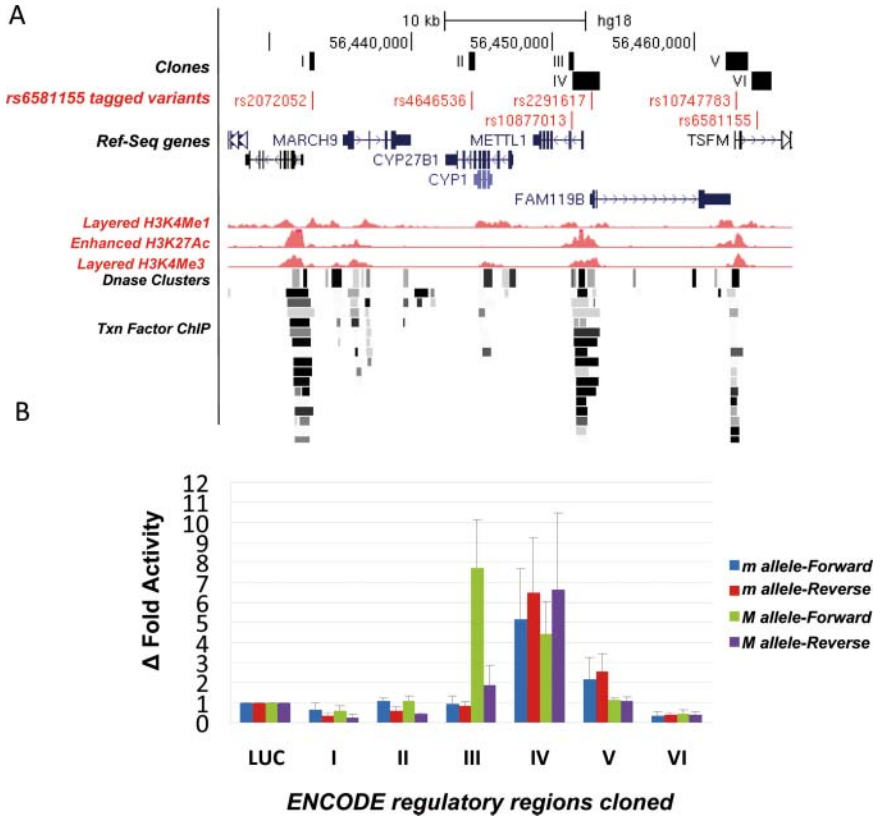

Figure 2 Enhancer-suppressor activity of the rs6581155-tagged variants. (A) Schematic illustration to show the localisation of the six tagged SNPs (in red) in potential regulatory regions (in black, I–VI) as indicated by ENCODE for the GM12878 lymphoblastoid cell line and the genes present in the region: *Enhancer H3K4Me1* track shows where modification of histone proteins is suggestive of enhancer; *Promoter H3K4Me3* track shows a histone mark associated with promoters; *Layered H3K4Me3* track shows histone mark associated with promoters that are active or poised to be activated; *DNase Clusters* track shows regions where chromatin is hypersensitive to DNase I enzyme; *Txn Factor ChIP* track shows DNA regions where transcription factors bind to DNA as assayed by chromatin immunoprecipitation (ChIP) with antibodies specific to the transcription factor followed by sequencing of the precipitated DNA (ChIP-Seq). (B) Δ Fold luciferase activity respect to the empty vector of the different constructs corresponding to the six regions (I–VI) with potential regulatory activity, containing the rs6581155 tagged polymorphisms transfected into Raji B cells. Four clones for each region bearing the different alleles (allele m, minor and M, major) and in both orientations (Forward and Reverse) from three independent transfection experiments are represented. Luciferase activity levels are referred to the level of the control plasmid containing only the basic promoter and the Renilla activity.

in the coding region of *CYP27B1* gene associated with MS.[17] This variant was not detected in our sample set.

In the present work, by integrating several strategies, we have uncovered the variant most strongly associated with MS in the region. This variant also affects the activity of an enhancer in an allele-dependent and orientation-dependent manner and correlates with the expression of five genes of the locus. Interestingly, the major allele of this SNP (rs10877013) correlates with high luciferase activity and at the same time with high expression of *TSFM* and *TSPAN31* and low expression of *FAM119B, CYP27B1* and *AVIL*. Although it could seem contradictory that one enhancer may differentially affect the expression of different genes, parallel situations have been described extensively in two scenarios: first, the same transcription factor is capable of activating or repressing depending on the context;[32 33] second, a single DNA sequence can be bound by different factors with opposite activities.[34] On the other hand, we cannot forget the important role of H3K4me preventing aberrant gene expression or modulating transcriptional response. H3K4me turns to mark for both positive and negative regulation. The role of H3K4me2 in histones deacetylation could serve as a fine-tuning regulatory mechanism.[35]

It has been described that interacting promotes are coexpressed in a tissue-specific manner. Based on the Chromatin Interaction Analysis with Paired-End-Tag sequencing results mapping long-range chromatin interactions associated with RNAPII and H3K4me2,[36 37] we observed that the genes in the analysed region are engaged through promoter–promoter, promoter–enhancer interactions forming a multigene complex that could provide a structural framework for cotranscription (figure 3). In this sense, the effect of a polymorphism on one enhancer embedded in this multigene complex could have diverging effects depending on the interaction of this enhancer with the different promoters. Additionally, the interactions could be tissue-type or cell-type specific, as has been shown by Li *et al*.[36] Therefore, it is not evident which of the genes in the region may be responsible for the genetic association with MS, since they seem to be coordinately regulated, and the enhancer containing the MS-associated variant may be affecting either of them; in fact, it acts as an eQTL for all of them. We observed that the Spearman's correlation coefficients (r) between the expression levels of the different genes in this cluster and the rs6581155-tagged SNPs were different in the different assays performed. For instance, in monocytes the best correlation is observed with *FAM119B* (r=−0.52), however in data obtained from LCL by RNA-seq a medium-high correlation coefficient is found for both *CYP27B1* (r=−0.472) and *FAM119B* (r=−0.474), and with microarrays the correlation coefficient is

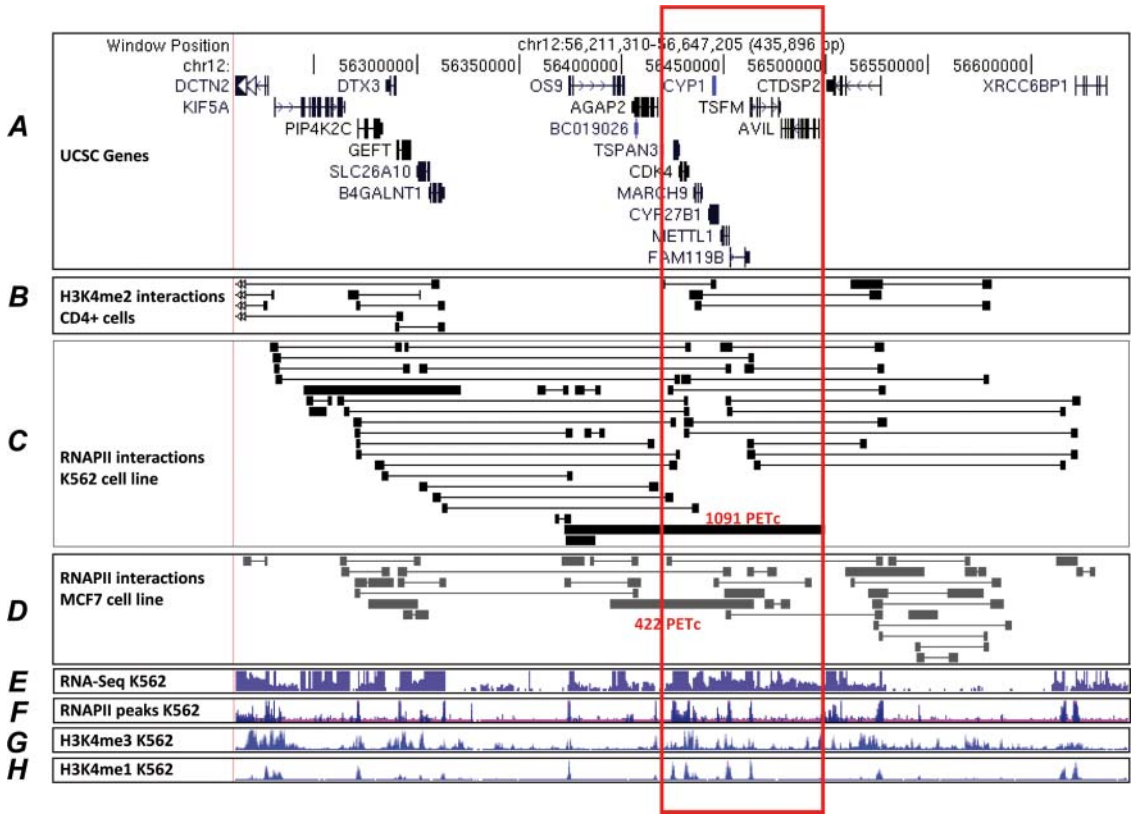

Figure 3 Datasets on Long-Range Chromatin Interaction with Paired-End-Tag sequencing (ChIA-PET) in the studied region. (A) Coordinates of chr12 in B37 and known genes by University of California Santa Cruz (UCSC) browser. (B) *H3K4me2* ChIA-PET data obtained from CD4 T cells, data from Chepelev *et al.*[37] (C) RNAPII ChIA-PET data obtained in K562 cells (data from Li *et al*[36]) In red, the number of the PET-counts (PETc) of the most frequent interactions is indicated. (D) RNAPII ChIA-PET data obtained in MCF7 cells.[36] The most frequent interactions are indicated in red. (E) ENCODE information on transcription levels assayed by sequencing of polyadenylated RNA from K562 cells. (F) ENCODE information on the RNAPII binding sites by ChIP-seq. (G) *H3K4Me3* histone mark in K562 cells associated with promoters. (H) *H3K4Me1* modification of histone proteins in K562 cells, suggestive of enhancer and, to a lesser extent, other regulatory activity.

very similar for *TSFM* and *FAM119B*. Based on data from LCL cells, *TSFM, FAM119B* or *CYP27B1* would be candidates, while *FAM119B* appears to emerge from monocytes. In fact, Gandhi *et al*[20] reported that *FAM119B* expression level was much lower in MS patients carrying the susceptibility haplotype. We observed that this correlation occurs also in samples from healthy donors as they are the CEU HapMap samples. However, we do not know the precise conditions (cell type, stage of disease, etc) under which expression of these genes could be determinant for the development of MS. For the same reason the variant-specific enhancer activity of fragment III, reported in this work, could vary in different cell-types. Here we have used in the activity assays the same cell-type used in the eQTL analysis to reproduce the effect observed in the expression of these genes. The Raji cells are lymphoblast-like cells derived from B-lymphocyte as the LCL from HapMap. The effect of the enhancer on the gene expression of proximal tubule cells of the kidney and the disease-activated macrophages, which are the major source of CYP27B1,[38] could shed light on the involvement of this variant in MS.

Relying on published data from other groups on the concentration of precursors of vitamin D in serum, two independent groups reported a correlation between the major alleles of rs703842, tagged by rs6581155 in the present work, with high 25-hydroxyvitamin D (25(OH)D) concentrations in MS patients.[39 40] On the other hand, Lange *et al*[41] reported a correlation of the major allele of rs10877012, in total LD with rs6581155, and low $1,25(OH)_2D$ serum levels. The major allele of the MS-associated SNPs, in total LD with the previously mentioned SNPs, correlated with low expression of the *CYP27B1* gene. This enzyme catalyses the conversion of 25(OH)D to $1,25(OH)_2D$, thus a low expression of this enzyme, could reconcile the above seemingly conflicting data, via accumulation of 25(OH)D and a low $1,25(OH)_2D$ product. These data support the *CYP27B1* gene as a very plausible candidate gene for association of the region with MS.

Vitamin D, besides having well-known control functions of calcium and phosphorus metabolism, bone formation and mineralisation, also has a complex role in the maintenance of immune homeostasis. The administration of vitamin D in animal models leads to improvement of immune-mediated symptoms. The correlation between the MS- associated variants and circulating levels of vitamin D supports the important role of vitamin D in susceptibility to MS. Thus, if this is the case in human MS, both the pharmacogenomics and the expression analysis of all the genes affecting the causing factor can reveal the kind of intervention necessary for neutralising a pathogenic situation.[42]

In this work we localise a group of variants in almost total LD that explain the association of the 12q13.3–12q14.1 region with MS. One of these variants (rs10877013) strongly affects the activity of one enhancer in the region, in an allele-dependent and orientation-dependent way, that could be the cause of the alteration of *TSFM*, *TSPAN31*, *FAM119B*,

CYP27B1 and *AVIL* gene expression, due to the promoter multigene interactions observed in the region. Some of the tagged variants have been described to be associated with the vitamin D serum concentration. Our results point to one of these as the causal gene for MS association in the locus.

Author affiliations
[1]Department of Cell Biology and Immunology Instituto de Parasitología y Biomedicina 'López Neyra', Consejo Superior de Investigaciones Científicas (IPBLN-CSIC), Granada, Spain
[2]Servicio de Neurología, Instituto de Neurociencias Clínicas, Hospital Regional Universitario Carlos Haya, Málaga, Spain
[3]Neurology Service, Hospital Clinic and I. d'Investigació Biomèdica Pi iSunyer (IDIBAPS), Barcelona, Spain
[4]Unidad de Esclerosis Múltiple, Hospital Virgen Macarena, Sevilla, Spain
[5]Servicio de Biología Molecular, Hospital Virgen Macarena, Sevilla, Spain
[6]Laboratorio de Investigación, Instituto de Neurociencias Clínicas, Hospital Regional Universitario Carlos Haya, Málaga, Spain
[7]Departamento de Lenguajes y Sistemas Informáticos, CITIC, Universidad de Granada, Granada, Spain
[8]Neurogenomiks Group, Universidad del País Vasco (UPV/EHU), Leioa, Spain
[9]Servicio de Neurología, Hospital de Basurto, Bilbao, Spain
[10]Servicio de Genética, Hospital de Basurto, Bilbao, Spain
[11]IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
[12]Immunology Department, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain
[13]Multiple Sclerosis Unit, Neurology Department, Hospital Clínico San Carlos, Instituto de Investigation Sanitaria del Hospital Clinico San Carlos (IdISSC), Madrid. Spain
[14]Centre d'Esclerosi Múltiple de Catalunya, CEM-Cat, Unitat de Neuroimmunologia Clínica, Hospital Universitari Vall d'Hebron (HUVH), Barcelona, Spain
[15]Institute of Evolutionary Biology (UPF-CSIC), PRBB, Barcelona, Spain
[16]National Institute for Bioinformatics, Universitat Pompeu Fabra, Barcelona, Spain
[17]Institució Catalana de Recerca i EstudisAvançats (ICREA), Catalonia, Spain
[18]Multiple Sclerosis Unit, Neuroscience Area, Biodonostia Health Research Institute, Donostia-San Sebastian, Spain
[19]Servicio de Neurología, Unidad de Esclerosis Múltiple, Hospital Donostia, San Sebastián, Spain

Acknowledgements We thank patients with multiple sclerosis and control subjects for making this study feasible. SNP genotyping services were provided by the Spanish Centro Nacional de Genotipado CEGEN-USC (http://www.cegen.org).

Contributors AA, EU, KV and FM designed the study, performed the main statistical analyses and co-wrote the manuscript. MF carried part of the experiments. OF, AS, GI, ML, LL, JAGL, IA, AA, MJGB, JV, BH, RA, MC, XM, DO, JO, YB coordinated sample collection. MAA, NPM, AN carried out some bioinformatics analysis.

Funding Financial support for the study was provided by Ministerio de Economía y Competitividad (MINECO)-Fondos Europeos de Desarrollo Regional (FEDER) (grant number SAF2009-11491); Fondo de Investigación Sanitaria (FIS) (grant numbers RETICS-REEM RD07/0060, CP 10/00526, PI10/1985, PS09/02105); Junta de Andalucía- Fondos Europeos de Desarrollo Regional (FEDER) (grant numbers P07-CVI-02551, P09-CTS-5218); Ikerbasque; the Basque Foundation for Science (Bilbao) and Fundación Ilundain.

Competing interests None.

Patient consent Obtained.

Ethics approval This study was approved by the Institutional Review Boards of Hospital Regional Universitario Carlos Haya of Malaga, Hospital Clinic of Barcelona, Hospital Virgen Macarena of Sevilla, Hospital de Basurto of Bilbao, Hospital Clínico San Carlos of Madrid, Hospital Universitari Vall d'Hebron of Barcelona and Hospital Donostia of San Sebastian.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The data presented in the manuscript are available on request.

REFERENCES

1 Hauser SL, Oksenberg JR. The neurobiology of multiple sclerosis: genes, inflammation, and neurodegeneration. *Neuron* 2009;52:61–76.
2 McElroy JP, Oksenberg JR. Multiple sclerosis genetics. *Curr Top Microbiol Immunol* 2008;318:45–72.
3 Ramagopalan SV, Knight JC, Ebers GC. Multiple sclerosis and the major histocompatibility complex. *Curr Opin Neurol* 2009;22:219–25.
4 Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. *Nature* 2009;461:747–53.
5 Bomprezzi R, Kovanen PE, Martin R. New approaches to investigating heterogeneity in complex traits. *J Med Genet* 2003;40:553–9.
6 Fierabracci A, Milillo A, Locatelli F, Fruci D. The putative role of endoplasmic reticulum aminopeptidases in autoimmunity: insights from genomic-wide association studies. *Autoimmun Rev*. Published Online First: May 2012. doi:10.1016/j.autrev.2012.04.007.
7 Corrado L, Bergamaschi L, Barizzone N, Fasano ME, Guerini FR, Salvetti M, Galimberti D, Benedetti MD, Leone M, D'Alfonso S. Association of the CBLB gene with multiple sclerosis: new evidence from a replication study in an Italian population. *J Med Genet* 2011;48:210–11.
8 International Multiple Sclerosis Genetics Consortium; Wellcome Trust Case Control Consortium 2, Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, Moutsianas L, Dilthey A, Su Z, Freeman C, Hunt SE, Edkins S, Gray E, Booth DR, Potter SC, Goris A, Band G, Oturai AB, Strange A, Saarela J, Bellenguez C, Fontaine B, Gillman M, Hemmer B, Gwilliam R, Zipp F, Jayakumar A, Martin R, Leslie S, Hawkins S, Giannoulatou E, D'alfonso S, Blackburn H, Martinelli Boneschi F, Liddle J, Harbo HF, Perez ML, Spurkland A, Waller MJ, Mycko MP, Ricketts M, Comabella M, Hammond N, Kockum I, McCann OT, Ban M, Whittaker P, Kemppinen A, Weston P, Hawkins C, Widaa S, Zajicek J, Dronov S, Robertson N, Bumpstead SJ, Barcellos LF, Ravindrarajah R, Abraham R, Alfredsson L, Ardlie K, Aubin C, Baker A, Baker K, Baranzini SE, Bergamaschi L, Bergamaschi R, Bernstein A, Berthele A, Boggild M, Bradfield JP, Brassat D, Broadley SA, Buck D, Butzkueven H, Capra R, Carroll WM, Cavalla P, Celius EG, Cepok S, Chiavacci R, Clerget-Darpoux F, Clysters K, Comi G, Cossburn M, Cournu-Rebeix I, Cox MB, Cozen W, Cree BA, Cross AH, Cusi D, Daly MJ, Davis E, de Bakker PI, Debouverie M, D'hooghe MB, Dixon K, Dobosi R, Dubois B, Ellinghaus D, Elovaara I, Esposito F, Fontenille C, Foote S, Franke A, Galimberti D, Ghezzi A, Glessner J, Gomez R, Gout O, Graham C, Grant SF, Guerini FR, Hakonarson H, Hall P, Hamsten A, Hartung HP, Heard RN, Heath S, Hobart J, Hoshi M, Infante-Duarte C, Ingram G, Ingram W, Islam T, Jagodic M, Kabesch M, Kermode AG, Kilpatrick TJ, Kim C, Klopp N, Koivisto K, Larsson M, Lathrop M, Lechner-Scott JS, Leone MA, Leppä V, Liljedahl U, Bomfim IL, Lincoln RR, Link J, Liu J, Lorentzen AR, Lupoli S, Macciardi F, Mack T, Marriott M, Martinelli V, Mason D, McCauley JL, Mentch F, Mero IL, Mihalova T, Montalban X, Mottershead J, Myhr KM, Naldi P, Ollier W, Page A, Palotie A, Pelletier J, Piccio L, Pickersgill T, Piehl F, Pobywajlo S, Quach HL, Ramsay PP, Reunanen M, Reynolds R, Rioux JD, Rodegher M, Roesner S, Rubio JP, Rückert IM, Salvetti M, Salvi E, Santaniello A, Schaefer CA, Schreiber S, Schulze C, Scott RJ, Sellebjerg F, Selmaj KW, Sexton D, Shen L, Simms-Acuna B, Skidmore S, Sleiman PM, Smestad C, Sørensen PS, Søndergaard HB, Stankovich J, Strange RC, Sulonen AM, Sundqvist E, Syvänen AC, Taddeo F, Taylor B, Blackwell JM, Tienari P, Bramon E, Tourbah A, Brown MA, Tronczynska E, Casas JP, Tubridy N, Corvin A, Vickery J, Jankowski J, Villoslada P, Markus HS, Wang K, Mathew CG, Wason J, Palmer CN, Wichmann HE, Plomin R, Willoughby E, Rautanen A, Winkelmann J, Wittig M, Trembath RC, Yaouanq J, Viswanathan AC, Zhang H, Wood NW, Zuvich R, Deloukas P, Langford C, Duncanson A, Oksenberg JR, Pericak-Vance MA, Haines JL, Olsson T, Hillert J, Ivinson AJ, De Jager PL, Peltonen L, Stewart GJ, Hafler DA, Hauser SL, McVean G, Donnelly P, Compston A. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* 2011;476:214–19.
9 Matesanz F, González-Pérez A, Lucas M, Sanna S, Gayán J, Urcelay E, Zara I, Pitzalis M, Cavanillas ML, Arroyo R, Zoledziewska M, Marrosu M, Fernández O, Leyva L, Alcina A, Fedetz M, Moreno-Rey C, Velasco J, Real LM, Ruiz-Peña JL, Cucca F, Ruiz A, Izquierdo G. Genome-wide association study of multiple sclerosis confirms a novel locus at 5p13.1. *PLoS One* 2012;7:e36140.
10 Kristjansdottir G, Sandling JK, Bonetti A, Roos IM, Milani L, Wang C, Gustafsdottir SM, Sigurdsson S, Lundmark A, Tienari PJ, Koivisto K, Elovaara I, Pirttilä T, Reunanen M, Peltonen L, Saarela J, Hillert J, Olsson T, Landegren U, Alcina A, Fernández O, Leyva L, Guerrero M, Lucas M, Izquierdo G, Matesanz F, Syvänen AC. Interferon regulatory factor 5 (IRF5) gene variants are associated with multiple sclerosis in three distinct populations. *J Med Genet* 2008;45:362–9.
11 Goris A, Boonen S, D'hooghe MB, Dubois B. Replication of KIF21B as a susceptibility locus for multiple sclerosis. *J Med Genet* 2010;47:775–6.

12 Australia and New Zealand Multiple Sclerosis Genetics Consortium (ANZgene). Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nat Genet* 2009;41:824–8.
13 Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Guiducci C, Burtt NP, Gianniny L, Korman BD, Padyukov L, Kurreeman FA, Chang M, Catanese JJ, Ding B, Wong S, van der Helm-van Mil AH, Neale BM, Coblyn J, Cui J, Tak PP, Wolbink GJ, Crusius JB, van der Horst-Bruinsma IE, Criswell LA, Amos CI, Seldin MF, Kastner DL, Ardlie KG, Alfredsson L, Costenbader KH, Altshuler D, Huizinga TW, Shadick NA, Weinblatt ME, de Vries N, Worthington J, Seielstad M, Toes RE, Karlson EW, Begovich AB, Klareskog L, Gregersen PK, Daly MJ, Plenge RM. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. *Nat Genet* 2008;40:1216–23.
14 Zhernakova A, Stahl EA, Trynka G, Raychaudhuri S, Festen EA, Franke L, Westra HJ, Fehrmann RS, Kurreeman FA, Thomson B, Gupta N, Romanos J, McManus R, Ryan AW, Turner G, Brouwer E, Posthumus MD, Remmers EF, Tucci F, Toes R, Grandone E, Mazzilli MC, Rybak A, Cukrowska B, Coenen MJ, Radstake TR, van Riel PL, Li Y, de Bakker PI, Gregersen PK, Worthington J, Siminovitch KA, Klareskog L, Huizinga TW, Wijmenga C, Plenge RM. Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. *PLoS Genet* 2011;7:e1002.
15 Alcina A, Vandenbroeck K, Otaegui D, Saiz A, Gonzalez JR, Fernandez O, Cavanillas ML, Cénit MC, Arroyo R, Alloza I, García-Barcina M, Antigüedad A, Leyva L, Izquierdo G, Lucas M, Fedetz M, Pinto-Medel MJ, Olascoaga J, Blanco Y, Comabella M, Montalban X, Urcelay E, Matesanz F. The autoimmune disease-associated KIF5A, CD226 and SH2B3 gene variants confer susceptibility for multiple sclerosis. *Genes Immun* 2010;11:439–45.
16 Fung EY, Smyth DJ, Howson JM, Cooper JD, Walker NM, Stevens H, Wicker LS, Todd JA. Analysis of 17 autoimmune disease-associated variants in type 1 diabetes identifies 6q23/TNFAIP3 as a susceptibility locus. *Genes Immun* 2009;10:188–91.
17 Ramagopalan SV, Dyment DA, Cader MZ, Morrison KM, Disanto G, Morahan JM, Berlanga-Taylor AJ, Handel A, De Luca GC, Sadovnick AD, Lepage P, Montpetit A, Ebers GC. Rare variants in the CYP27B1 gene are associated with multiple sclerosis. *Ann Neurol* 2011;70:881–6.
18 Bailey R, Cooper JD, Zeitels L, Smyth DJ, Yang JH, Walker NM, Hyppönen E, Dunger DB, Ramos-Lopez E, Badenhoop K, Nejentsev S, Todd JA. Association of the vitamin D metabolism gene CYP27B1 with type 1 diabetes. *Diabetes* 2007;56:2616–21.
19 Sundqvist E, Bäärnhielm M, Alfredsson L, Hillert J, Olsson T, Kockum I. Confirmation of association between multiple sclerosis and CYP27B1. *Eur J Hum Genet* 2010;18:1349–52.
20 Gandhi KS, McKay FC, Cox M, Riveros C, Armstrong N, Heard RN, Vucic S, Williams DW, Stankovich J, Brown M, Danoy P, Stewart GJ, Broadley S, Moscato P, Lechner-Scott J, Scott RJ, Booth DR; ANZgene Multiple Sclerosis Genetics Consortium. The multiple sclerosis whole blood mRNA transcriptome and genetic associations indicate dysregulation of specific T cell pathways in pathogenesis. *Hum Mol Genet* 2010;19:2134–43.
21 Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, Johnson KP, Sibley WA, Silberberg DH, Tourtellotte WW. (1983) New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol* 1983;13:227–31.
22 Cavanillas ML, Fernández O, Comabella M, Alcina A, Fedetz M, Izquierdo G, Lucas M, Cénit MC, Arroyo R, Vandenbroeck K, Alloza I, García-Barcina M, Antigüedad A, Leyva L, Gómez CL, Olascoaga J, Otaegui D, Blanco Y, Saiz A, Montalbán X, Matesanz F, Urcelay E. Replication of top markers of a genome-wide association study in multiple sclerosis in Spain. *Genes Immun* 2010;12:110–15.
23 Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* 2005;76:887–93.
24 Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Statist Sic B* 1995;57:289–300.
25 Sarkar SK. False discovery and false nondiscovery rates in single-step multiple testing procedures. *Ann Stat* 2006;34:394–415.
26 Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005,21:263–5.
27 Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R, Maouche S, Germain M, Lackner K, Rossmann H, Eleftheriadis M, Sinning CR, Schnabel RB, Lubos E, Mennerich D, Rust W, Perret C, Proust C, Nicaud V, Loscalzo J, Hübner N, Tregouet D, Münzel T, Ziegler A, Tiret L, Blankenberg S, Cambien F. Genetics and beyond—the transcriptome of human monocytes and disease susceptibility. *PLoS One* 2010;5:e10693.
28 Dixon AL, Liang L, Moffatt MF, Chen W, Heath S, Wong KC, Taylor J, Burnett E, Gut I, Farrall M, Lathrop GM, Abecasis GR, Cookson WO. A genome-wide association study of global gene expression. *Nat Genet* 2007;39:1202–7.
29 Montgomery SB, Sammeth M, Gutierrez-Arcelus M, Lach RP, Ingle C, Nisbett J, Guigo R, Dermitzakis ET. Transcriptome genetics using second generation sequencing in a Caucasian population. *Nature* 2010;464:773–7.
30 Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, Ingle CE, Dunning M, Flicek P, Koller D, Montgomery S, Tavaré S, Deloukas P, Dermitzakis ET. Population genomics of human gene expression. *Nat Genet* 2007;39:1217–24.
31 Dimas AS, Deutsch S, Stranger BE, Montgomery SB, Borel C, Attar-Cohen H, Ingle C, Beazley C, Gutierrez Arcelus M, Sekowska M, Gagnebin M, Nisbett J, Deloukas P, Dermitzakis ET, Antonarakis SE. Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* 2009;325:1246–50.
32 Ceribelli M, Dolfini D, Merico D, Gatta R, Viganò AM, Pavesi G, Mantovani R. The histone-like NF-Y is a bifunctional transcription factor. *Mol Cell Biol* 2008;28:2047–58.
33 Sakabe NJ, Aneas I, Shen T, Shokri L, Park SY, Bulyk ML, Evans SM, Nobrega MA. Dual transcriptional activator and repressor roles of TBX20 regulate adult cardiac structure and function. *Hum Mol Genet* 2012;21:2194–204.
34 Zarnegar MA, Chen J, Rothenberg EV. Cell-type-specific activation and repression of PU.1 by a complex of discrete, functionally specialized cis-regulatory elements. *Mol Cell Biol* 2010;30:4922–39.
35 Pinskaya M, Morillon A. Histone H3 lysine 4 di-methylation: a novel mark for transcriptional fidelity? *Epigenetics* 2009;4:302–6.
36 Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, Poh HM, Goh Y, Lim J, Zhang J, Sim HS, Peh SQ, Mulawadi FH, Ong CT, Orlov YL, Hong S, Zhang Z, Landt S, Raha D, Euskirchen G, Wei CL, Ge W, Wang H, Davis C, Fisher-Aylor KI, Mortazavi A, Gerstein M, Gingeras T, Wold B, Sun Y, Fullwood MJ, Cheung E, Liu E, Sung WK, Snyder M, Ruan Y. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* 2012;148:84–98.
37 Chepelev I, Wei G, Wangsa D, Tang Q, Zhao K. Characterization of genome-wide enhancer-promoter interactions reveals co-expression of interacting genes and modes of higher order chromatin organization. *Cell Res* 2012;22:490–503.
38 Adams JS, Hewison M. Extrarenal expression of the 25-hydroxyvitamin D-1-hydroxylase. *Arch Biochem Biophys* 2012;523:95–102.
39 Simon KC, Munger KL, Kraft P, Hunter DJ, De Jager PL, Ascherio A. Genetic predictors of 25-hydroxyvitamin D levels and risk of multiple sclerosis. *J Neurol* 2011;258:1676–82.
40 Orton SM, Morris AP, Herrera BM, Ramagopalan SV, Lincoln MR, Chao MJ, Vieth R, Sadovnick AD, Ebers GC. Evidence for genetic regulation of vitamin D status in twins with multiple sclerosis. *Am J Clin Nutr* 2008;88:441–7.
41 Lange CM, Bojunga J, Ramos-Lopez E, von Wagner M, Hassler A, Vermehren J, Herrmann E, Badenhoop K, Zeuzem S, Sarrazin C. Vitamin D deficiency and a CYP27B1-1260 promoter polymorphism are associated with chronic hepatitis C and poor response to interferon-alfa based therapy. *J Hepatol* 2011;54:887–93.
42 Szodoray P, Nakken B, Gaal J, Jonsson R, Szegedi A, Zold E, Szegedi G, Brun JG, Gesztelyi R, Zeher M, Bodolay E. The complex role of vitamin D in autoimmune diseases. *Scand J Immunol* 2008;68:261–9.