Supporting Information

De Jager *et al.* 10.1073/pnas.0813310106

SI Methods

Genotyping. In the initial screen of the Brigham and Women's Hospital samples, SNPs were genotyped using the iPLEX™ Sequenom MassARRAY platform. We analyzed only those 24 SNPs with high quality data (>95% genotype call rate, HWE *P* value > 0.001 in healthy control subjects), and these data are reported for MS association in Table S2. In the fine mapping analysis, genotypes were generated using 1 of 2 platforms: (*i*) iPLEX™ Sequenom MassARRAY platform for the Australian trios, the Finnish cases and healthy control subjects, the U.S. and United Kingdom trios, the United Kingdom cases, and the U.S. cases and healthy control subjects or (*ii*) Taqman SNP Genotyping Assay C_31433800_10 (rs12044852) and C_15755405_10 (rs2300747) (Applied Biosystems) using a 7900 Sequence Detection System or 7300 Real-Time PCR System for the United Kingdom healthy control subjects and Belgian cases, healthy control subjects, and trios.

Resequencing for Polymorphism Discovery. To supplement the database from which to select polymorphisms for genotyping, we carried out bi-directional, PCR-based resequencing in the second block of linkage disequilibrium found within the *CD58* locus. We targeted the chromosome 1 segment stretching from 116,834,306 bp to 116,910,353 bp (human genome build hg18). A total of 16 individuals were used in the resequencing effort, including 8 subjects with MS and 2 healthy control subjects that are homozygous for the rs12044852^{A} allele that is associated with decreased risk of MS. The remaining 6 subjects are healthy control subjects that are homozygous for the major rs12044852^{G} allele. Twenty-seven putative novel SNPs were discovered in at least one resequenced individual, and 10 of these SNPs were validated to be polymorphic in the HapMap DNA samples (see below and Table S3). These new validated polymorphisms have been submitted to the dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP).

SNP Validation. We examined putative SNPs from 2 different sources: (*i*) our resequencing effort (see *Resequencing for Polymorphism Discovery*) and (*ii*) the dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP). Within our chromosome 1 segment of interest (116,834,306 bp to 116,910,353 bp), we considered 27 putative SNPs from the resequencing effort and 136 putative SNPs and indels from dbSNP that had no data in the HapMap resource (http://www.hapmap.org/) (1). We attempted to genotype both sets of SNPs and indels in the CEU and YRI populations of the HapMap using the Sequenom iPLEX platform as described above. The CHB and JPT samples were only genotyped with the putative SNPs from the resequencing effort. Table S3 presents the results of this SNP validation effort. We considered a SNP to be validated if at least one heterozygous individual was found in the panel of HapMap subjects.

Lymphoblastic Cell Line RNA Data and Quantitative Trait Analysis. The *CD58* RNA measurement for each LCL analyzed in this study was produced by the Wellcome Trust Sanger Institute using the Illumina platform and is publicly available (2). We downloaded this dataset which has data on the *CD58* RNA and all other RNA species included on the array; measurements for each RNA species were generated by each of the 4 technical replicates performed by the Sanger Institute. For each individual LCL, the values from the 4 technical replicates were log-transformed and mean intensity centered array-wide. The per gene median was taken as that LCL's value for each gene. The dataset was then quantile-normalized across cell lines. The resulting values for CD58 were used in the primary analysis correlating *CD58* RNA expression to rs2300747 genotype. rs2300747 genotype data for each LCL were downloaded from the HapMap resource (1). Only unrelated individuals with both genotype and expression data were used in the analysis: 88 CHB and JPT LCLs and 60 CEU LCLs. The primary analysis of the expression data were performed using the quantitative trait analysis option found in the PLINK toolkit v0.99r by S. Purcell (3). Secondarily, a separate pairwise comparison of genotype classes was performed using a Welch 2 sample *t* test to better illustrate the allele's dose-dependent effect on RNA expression in Fig. 2*A*.

Mononuclear Cell RNA Data from Subjects with MS. Between July 2002 and October 2007, peripheral blood mononuclear cell (PBMC) samples were collected from relapsing-remitting MS subjects and CIS subjects as part of the Comprehensive Longitudinal Investigation of MS at the Brigham and Women's Hospital (4)). The female: male ratio is 3:1. The mean disease duration at the time of RNA sampling is 7.5 years, ranging from 0 to 43 years. The mean age of symptom onset is 33 years old, with a range of 14 to 55years old. A total of 213 subjects had RRMS, and 26 had CIS at the time of RNA sampling. CIS subjects differ from MS subjects by having had only one clinical episode of demyelination; MS subjects, by definition, must have at least 2 such events or 1 event and evidence of disease activity in a paraclinical measure such as MRI.[29] Nonetheless, the pathophyisology is shared between these 2 sets of subjects, and they are treated in the same manner in a clinical environment. PBMCs were isolated from heparinized blood by centrifugation on a Ficoll-Hypaque (Amersham Biosciences) gradient and immediately frozen in 90% FBS and 10% DMSO. All blood samples were processed within 3 h of phlebotomy. Total RNA from frozen samples was isolated using a homogenization shredding system in a microcentrifuge spin-column format (QIAshredder, Qiagen), followed by total RNA purification using selective binding columns (RNeasy Mini Kit, Qiagen) according to the manufacturer's protocol.

RNA concentration was determined using the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Inc.). RNA quality was assessed on Agilent Bioanalyzer 2100 using the Agilent RNA 6000 Nano Assay kit (Agilent Technologies). The overall total RNA quality was evaluated by A260/A280 ratio (ratio >1.8) and electropherogram (score >7). Two micrograms of extracted RNA have been reversed transcribed in vitro (Two-Cycle cDNA Synthesis Kit, Affymetrix), labeled (IVT Labeling Kit, Affymetrix), and hybridized on Affymetrix gene chip - U133 2.0 plus. The GeneChip arrays were scanned on a GeneChip Scanner 3000.

Once generated, the RNA data underwent a rigorous quality control (QC) analysis using the recommended pipeline available in the R package simpleaffy and affyPLM (Bioconductor). The quality parameters that we monitored included: (*i*) background noise, (*ii*) percentage of present called probe sets, (*iii*) scaling factor, (*iv*) information about exogenous control transcripts from the Affymetrix PolyA control kit, and (*v*) the ratio of intensities of 3′/5′ probes for the housekeeping genes *GAPDH* and *ACTB*. We then normalized the data using GCRMA.

239 RNA profiles from our collection met our QC criteria, producing a value for CD58 RNA expression, and had genotypes

for rs2300747. Welch 2-sample t tests were applied to perform our analyses comparing treatment regimen (untreated/GA/IFNβ) and genotype categories for rs2300747. Since a single subject was homozygous for rs2300747^{G}, the rs2300747^{AA} subjects were compared to a combined set of subjects having at least one rs2300747^{G} allele (rs2300747^{AG} and rs2300747^{GG} subjects).

Whole Blood RNA Data from Subjects with MS. We extracted these *CD58* RNA expression values from an existing dataset (5). This dataset was generated from whole blood of healthy control subjects ($n = 20$), untreated subjects with MS at the time of a neurologist-confirmed clinical relapse ($n = 10$), and untreated subjects with MS during a neurologist-confirmed clinical remission ($n = 10$). The subjects in remission had had 1–2 relapses in the 2 years before blood sampling. Additional details regarding these subjects, RNA extraction, and data generation methods are provided elsewhere (6). RNA expression for 9,381 genes was measured using the 10.5K cDNA Human array at the Peter MacCallum Cancer Institute. For each subject with MS, the individual curated expression value of CD58 in the database was normalized to the mean CD58 expression of the healthy control subject population. We then compared the distribution of relative CD58 expression in remission to the distribution of relative CD58 expression observed in relapse using t test.

1. The International HapMap Consortium (2005) A haplotype map of the human genome. *Nature* 437:1299–1320.
2. Stranger BE, *et al.* (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* 315:848–853.
3. Purcell S, *et al.* (2007) PLINK: A toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet* 81:559–575.
4. Gauthier SA, *et al.* (2007) Predicting short-term disability in multiple sclerosis. *Neurology* 68: 2059–2065.
5. Dixon AL, *et al.* (2007) A genome-wide association study of global gene expression. *Nat Genet* 39:1202–1207.
6. Arthur AT, *et al.* (2008) Genes implicated in multiple sclerosis pathogenesis from expression profiles and genotyping. *BMC Med Genet* 9:17.
7. Lundmark F, *et al.* (2007) Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. *Nat Genet* 39:1108–1113.

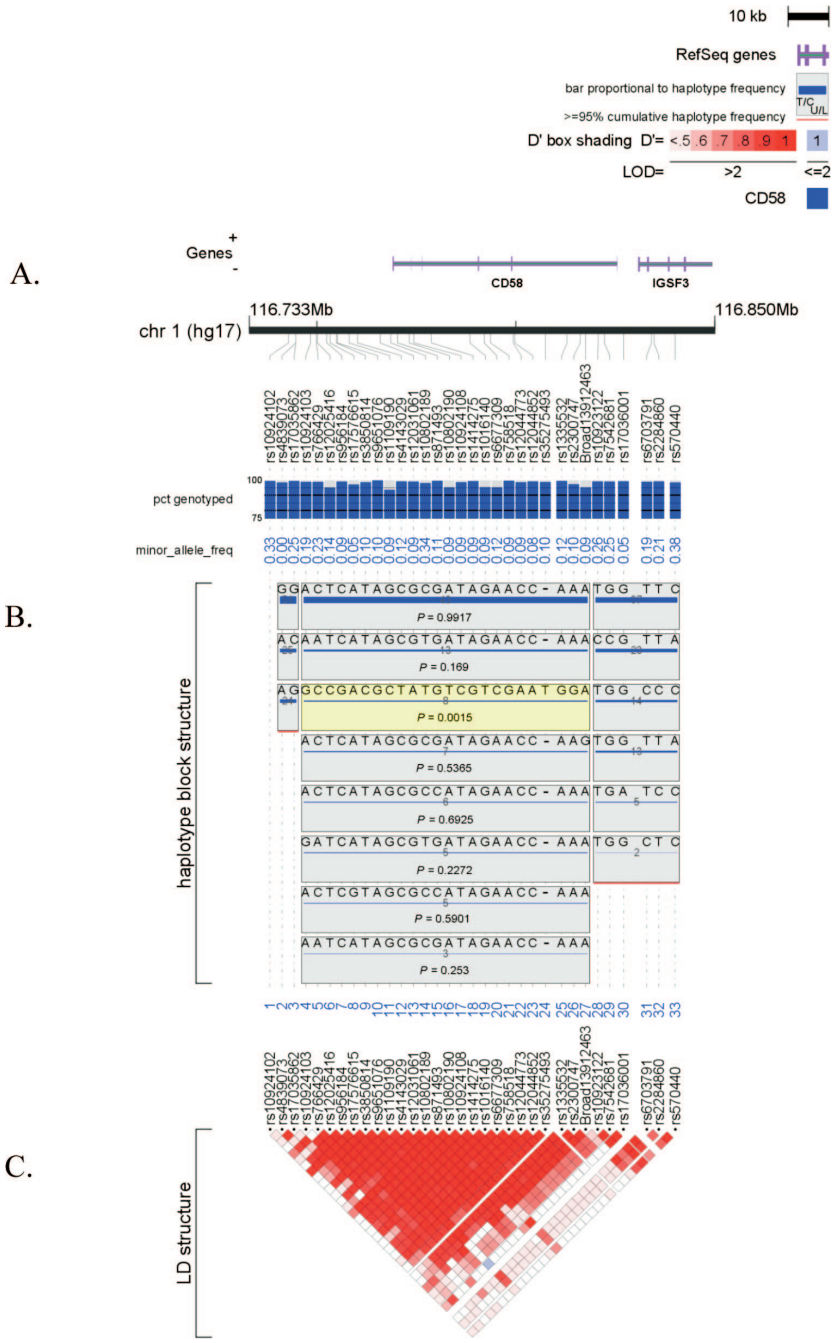

Fig. S1. A haplotype containing only the *CD58* gene is under-represented in subjects with MS. This figure was derived from the genotype data available in the BWH samples used in the analysis and highlights the association of a single haplotype (in yellow) with susceptibility to MS. The detailed association results of each SNP are reported in Table S2. (*A*) The genomic segment of interest is outlined in black at the top of the figure, and its physical boundaries are noted in megabases (Mb). Human genome assembly 18 (hg18) is the reference selected for these physical positions. Above the genomic segment, the position of the *CD58* and *IGSF3* genes are noted in blue and and gray; *ATP1A1* is found telomeric to *CD58* but is outside the segment of DNA shown here. Below the genomic segment, 33 SNPs used in the study are listed in the genomic order, along with their minor allele frequencies. One SNP, rs17036591, is not shown because it lies out of the genomic segment under consideration. (*B*) Each haplotype of >0.01 frequency found in the BWH sample collection is shown here with its constituent alleles; data from both cases and healthy control subjects were used to produce this figure. We interpret the linkage disequilibrium data as being consistent with the presence of 3 blocks of linkage disequilibrium within the chromosomal segment that is interrogated as part of this work. The haplotype associated with MS is highlighted in yellow. The rs2300747^{G} allele is a surrogate marker for this haplotype and is used throughout the replication and expression studies. Regardless of which of the 3 haplotype block definition found in Haploview is used, a haplotype containing rs12044852^{A} and rs2300747^{G} is associated with MS with its boundaries varying depending on the definition used. None of these haplotypes extend to include *IGSF3*. The P value association of each haplotype with susceptibility to MS is listed in the figure. (*C*) We illustrate the linkage disequilibrium structure of the region by reporting pairwise SNP correlations using squares whose intensity of red is directly proportional to the r^2 value between 2 SNPs.

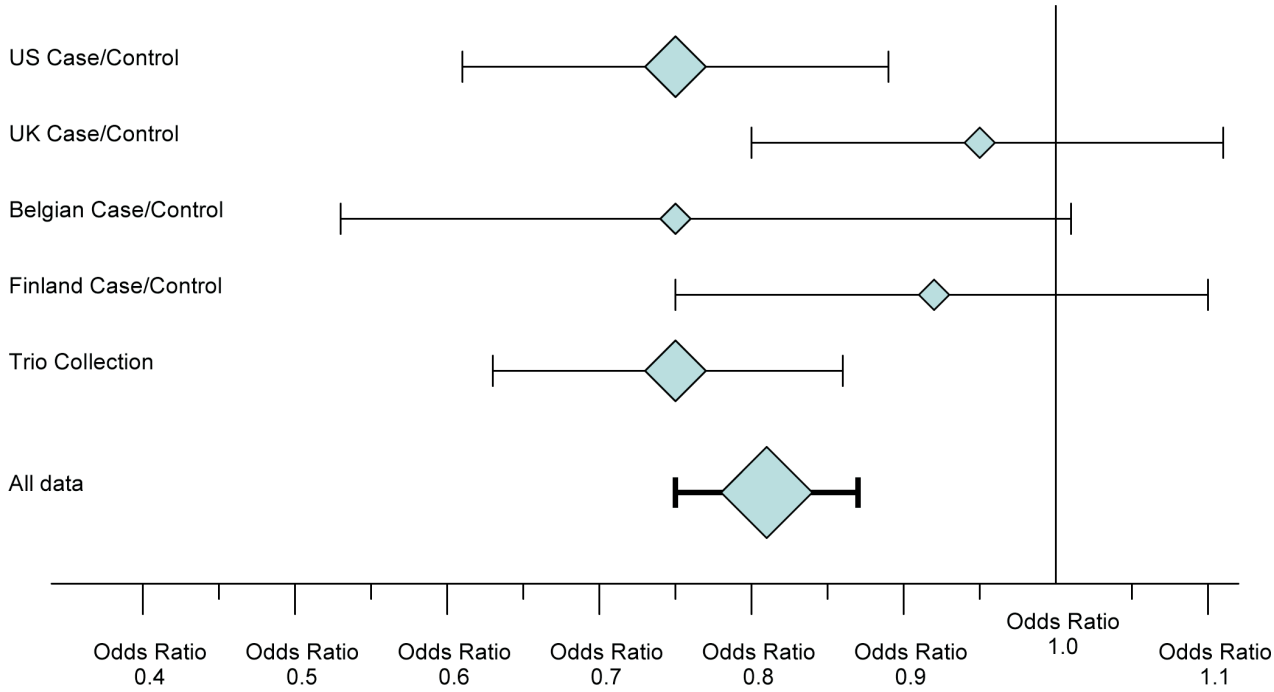

Fig. S2. Summary of the association analysis of rs2300747^{G} with decreased susceptibility with MS (All data: odds ratio 0.82, 95% CI 0.75–0.89; $P = 1.1 \times 10^{-6}$). Our Mantel-Haenszel analysis of rs2300747^{G} pools data from 5 strata that are represented here. We plot the odds ratio of each stratum and its 95% confidence interval; the diamond marks the estimated odds ratio. The size of the diamond is proportional to the number of cases in each sample set: U.S.: 1,557 cases/855 controls, United Kingdom: 961 cases/2,466 controls, Belgium: 348 cases/372 controls, Finland 692 cases/727 controls, and 1,768 trios from Australia, Belgium, the United Kingdom and the U.S. Table S5*b* reports these results in detail. N.B. the trio data were generated in 3 different batches that are combined into one analysis for this SNP. None of the affected subjects used in the trio analysis is used in the case/control analyses; each of the 5 strata in this analysis contains a set of independent, unique subjects.

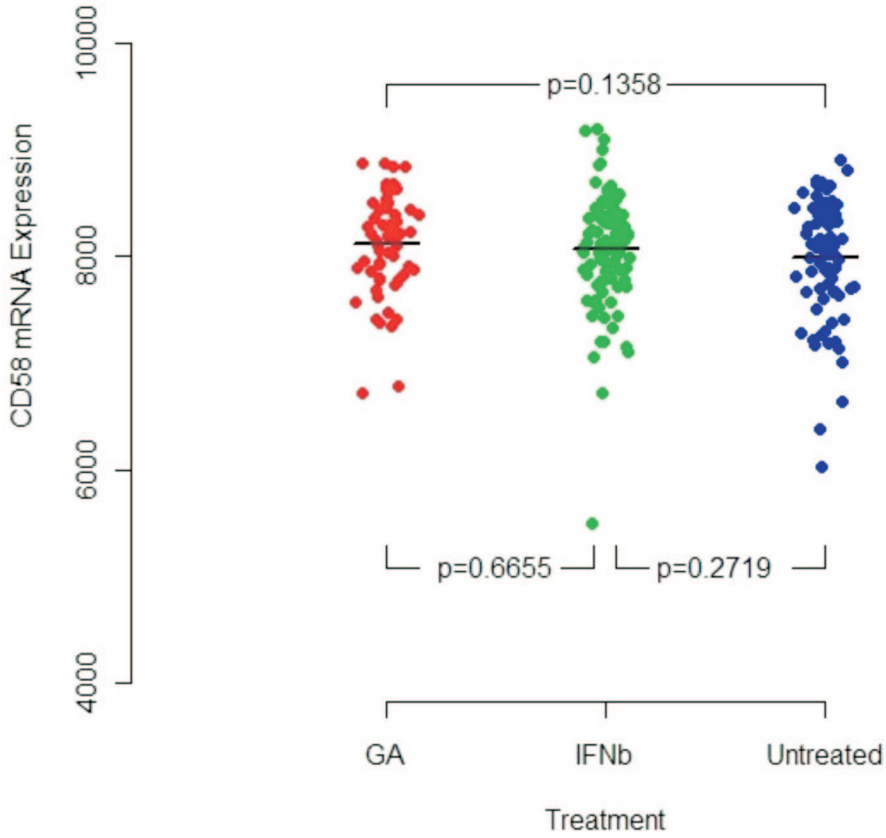

Fig. S3. Treatment with either glatiramer acetate (GA) or IFN beta (IFNβ) does not have a strong effect on the expression of CD58 RNA in PBMCs of subjects with MS or CIS. The 239 subjects with MS or CIS are assessed in this analysis: 81 who are untreated, 64 treated with GA, and 94 treated with IFNb. Given the limited number of subjects, we pooled all subjects receiving one of the IFNβ formulations (Avonex, Betaseron, or Rebif) into a single IFNβ category. None of these subjects were on combination therapy for MS at the time of sampling.

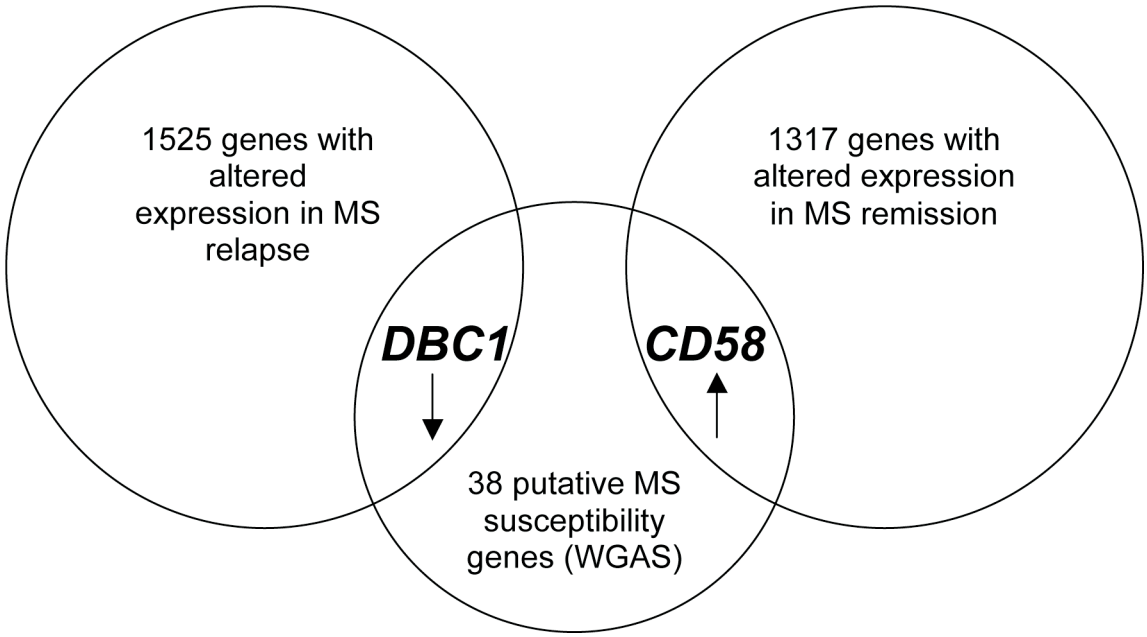

Fig. S4. Two putative MS susceptibility loci have RNA species that are significantly differentially regulated in blood sampled from the remission ($n = 10$) and relapse phases ($n = 10$) of MS. We searched an existing dataset of genes whose expression has been reported to be significantly increased or decreased in 1 of 2 clinical states of MS: remission and relapse (6). We limited our search to those 38 genes from our IMSGC whole genome association scan that displayed enhanced evidence of association to MS susceptibility after the replication effort (7); we consider these 38 genes to be putative susceptibility loci that await further validation. Of these 38 genes, only 2 are differentially expressed in these data: *DBC1* is down-regulated during relapses, and *CD58* is up-regulated during remissions. The change in *CD58* expression during relapses is not significantly different from the baseline expression of *CD58* in 20 healthy control individuals (7). For this analysis, a significant difference in RNA expression was defined as a $P < 0.05$ after correction for the testing of multiple hypotheses (7).

Table S1. Clinical features of the subject collections used in this analysis

	BWH collection	Additional collections							
Subject source	BWH	Australia	Belgium		Finland	United Kingdom		U.S.	
Analysis	C/C	Trios	C/C	Trios	C/C	C/C	Trios	C/C	Trios*
Number of trio families	—	372	—	134	—	—	833	—	707
Number of cases	707	—	368	—	715	921	—	679	—
Female:Male ratio	3.0:1.0	4.0:1.0	1.7:1.0	1.6:1.0	2.5:1.0	2.7:1.0	3.0:1.0	2.4:1.0	3.1:1.0
Mean disease duration (range)	12 (0–46)	9 (6–41)	16 (0–47)	10 (0–39)	NA	13 (0–45)	11 (0–40)	12 (0–47)	9 (0–36)
Mean age at onset, years (range)	33 (10–69)	33 (7–55)	35 (10–65)	28 (9–45)	NA	33 (13–60)	27 (10–53)	34 (3–63)	30 (9–51)
Disease course, n, %	—	—	—	—	—	—	—	—	—
″Bout onset″	659 (93)	315 (85)	324 (88)	132 (99)	NA	761 (83)	778 (93)	408 (60)	533 (75)
Relapse remitting	518 (73)	249 (67)	NA	NA	NA	482 (52)	534 (64)	159 (24)	112 (16)
Secondary progressive	141 (20)	66 (18)	NA	NA	NA	279 (30)	244 (29)	47 (7)	32 (5)
Primary progressive	28 (4)	18 (5)	42 (11)	2 (1)	NA	132 (14)	55 (7)	56 (8)	23 (3)
Clinically isolated syndrome	7 (1)	0	0	0	NA	0	0	9 (1)	7 (1)
Unknown	12 (2)	37 (5)	2 (1)	0	NA	28 (3)	0	0	0

Glossary: BWH = Brigham and Women's Hospital; C/C = case/control collection; Trios = trio families (affected child and biological parents); and NA = not available. Information on whether a subject was in the relapsing remitting or secondary progressive phase of the disease was not available for the Belgian collection. ″Bout onset MS″ is used here to describe subjects whose disease course started with discrete attacks of inflammatory demyelination. EDSS = Expanded Disability Status Scale (EDSS). The EDSS used is the EDSS value recorded on or within one year of enrollment into any of the studies. These EDSS data are not available for the Belgian collection. *, The U.S. cases are all collected through University of California, San Francisco, CA; 153 of the ″U.S. trios″ were collected at BWH and are pooled with the University of California, San Francisco, CA trios. The subjects with MS in the trios are distinct from those used in the case/control analyses. Each subject with MS is used only once in our analyses.

Table S2. Results of the association analyses for the SNPs genotyped in the 2 phases of the project: Phase 1 (screening) and Phase 2 (map refinement)

SNP (source)	Chr 1 position hg18 (bp)	Genotyping phase	Location with respect to *CD58*	BWH minor allele MS, HC	BWH MS, HC MAF	BWH χ^2 *P* value
rs10924102 (HM)	116832505	1	3′region	T, T	0.339, 0.300	0.1288
rs4839073 (HM)	116834306	1	3′ region	A, G	0.481, 0.487	0.2479
rs17035862 (HM)	116834408	2	3′ region	C, C	0.240, 0.303	0.0075
rs10924103 (HM)	116838074	1	3′ region	G, G	0.183, 0.229	0.037
rs766429 (HM)	116839563	1	3′ region	A, A	0.342, 0.195	0.0444
rs12025416 (HM)	116839810	2	3′ region	C, C	0.128, 0.170	0.0289
rs956184 (HM)	116842145	1	3′ region	G, G	0.082, 0.126	0.0047
rs17576615 (HM)	116842994	1	3′ region	G, G	0.052, 0.063	0.4085
rs3850814 (HM)	116844567	1	3′ region	C, C	0.097, 0.134	0.0311
rs9651076 (HM)	116844825	1	3′ region	G, G	0.097, 0.136	0.0211
rs1109190 (HM)	116847786	2	3′ region	C, C	0.085, 0.130	0.0076
rs4143029 (HM)	116849853	1	3′ region	T, T	0.110, 0.149	0.0286
rs12031061 (HM)	116855534	1	3′ region	A, A	0.081, 0.127	0.0032
rs10802189 (HM)	116858253	2	3′ region	T, T	0.336, 0.348	0.6542
rs871493 (HM)	116859912	1	Intron 5	C, C	0.112, 0.114	0.8934
rs10802190 (HM)	116862907	2	Intron 5	T, T	0.081, 0.125	0.0066
rs10924108 (HM)	116863997	2	Intron 4	C, C	0.081, 0.126	0.0036
rs1414275 (HM)	116868430	1	Intron 3	G, G	0.086, 0.129	0.0076
rs1016140 (HM)	116878070	2	Intron 3	T, T	0.082, 0.126	0.0081
rs6677309 (HM)	116881689	2	Intron 2	C, C	0.110, 0.152	0.0215
rs758518 (HM)	116883333	1	Intron 2	G, G	0.084, 0.126	0.0085
rs12044773 (HM)	116889025	1	Intron 1	A, A	0.077, 0.128	0.0011
rs12044852 (HM)	116889302	1	Intron 1	A, A	0.074, 0.129	0.0003
rs35275493** (dbSNP)	116897025	2	Intron 1	T, T	0.094, 0.138	0.0093
rs1335532 (HM)	116902480	1	Intron 1	G, G	0.113, 0.153	0.0228
rs2300747 (HM)	116905738	1	Intron 1	G, G	0.090, 0.149	0.0004
rs55842195 ** (reseq)	116907113	2	Intron1	G, G	0.088, 0.078	0.4987
rs10923122 (HM)	116910353	1	Intron 1	C, C	0.268, 0.223	0.0597
rs7542681 (HM)	116911865	1	Intron 1	C, C	0.256, 0.206	0.037
rs17036001 (HM)	116916652	1	5′ region	A, A	0.053, 0.057	0.7811
rs6703791 (HM)	116923653	1	*IGSF3* cSNP	C, C	0.184, 0.215	0.1569
rs2284860 (HM)	116924410	1	*IGSF3* Intron	C, C	0.210, 0.235	0.2708
rs570440 (HM)	116929695	1	*IGSF3* Intron	A, A	0.391, 0.349	0.1227
rs17036591 (HM)	117274942	1	*PTGFRN Intron*	A, A	0.221, 0.231	0.6559

Glossary: BWH = Brigham and Women's Hospital; Chr = chromosome; HC = healthy control; hg18 = human genome assembly 18; MAF = minor allele frequency, and MS = multiple sclerosis. The selected Phase 1 SNPs consist of a panel of tag SNPs that captured all polymorphisms with an $r^2 > 0.8$ using the Tagger algorithm implemented in Haploview[31]. The CEU genotype data from HapMap[11] were used to derive the tag SNPs. Only SNPs with a frequency > 0.05 were selected for our analysis given our sample size, which precludes the robust assessment of less common SNPs. The complete *CD58* gene as well as 20 kilobases (kb) of upstream and 20 kb of downstream chromosomal DNA were targeted for tag selection.

Phase 2 SNPs were selected to give further information regarding the structure of this locus. The Tagger SNP selection analysis was performed again after adding new genotype data from HapMap samples that were generated during the validation of putative SNPs from dbSNP or the resequencing effort to our original dataset of HapMap genotypes (Table S3). The Phase 2 tag SNPs originating from dbSNP or the resequencing effort were selected to capture, in combination with the Phase 1 SNPs, all SNPs in the enhanced HapMap dataset with an $r^2 > 0.8$. These Phase 2 tag SNPs are marked with "**". The other Phase 2 SNPs were selected to refine the association with rs12044852: these SNPs had an $0.8 > r^2 > 1.0$ with rs12044852.

Table S3. Validation of putative SNPs from dbSNP and the resequencing effort using the HapMap DNA samples

SNP ID	Alleles	Chrom. 1 physical position (hg18)	Number of subjects with minor allele in resequencing panel	CEU minor allele	CEU MAF	YRI minor allele	YRI MAF	JPT minor allele	JPT MAF	CHB minor allele	CHB MAF
Putative SNPs from resequencing											
Broad13909142	A/C	116836303	1 control het	C	0.025	A	0	A	0	A	0
Broad13909631	A/T	116847619	1 MS het	A	0.017	T	0	T	0	T	0
Broad13910476	C/T	116863301	1 control het	T	0.025	C	0	C	0	C	0
Broad13911192	C/T	116877001	1 control het	T	0.034	C	0	C	0	C	0
Broad13912407	C/G	116905356	1 MS homo	G	0.051	C	0	C	0	C	0
Broad13912463	A/G	116907113	1 control het	G	0.085	A	0	A	0	A	0
Broad13911505	C/T	116885481	2 control hets	C	0.254	T	0	C	0.034	C	0.044
Broad13909666	C/G	116848500	2 control hets	G	0	C	0.083	G	0	G	0
Broad13911980	A/G	116896436	1 MS het	G	0	A	0.008	G	0	G	0
Broad13909880	C/T	116853796	1 control het	C	0	C	0	T	0.011	C	0
Broad13909541	A/G	116844852	1 control het	G	0	G	0	G	0	G	0
Broad13909569	A/C	116845681	1 control het	C	0	C	0	C	0	C	0
Broad13910591	C/T	116865367	1 MS het	C	0	C	0	C	0	C	0
Broad13910841	A/G	116873544	1 control het	A	0	A	0	A	0	A	0
Broad13911280	C/T	116879948	1 control het	T	0	T	0	T	0	T	0
Broad13911396	C/T	116882509	1 control het	C	0	C	0	C	0	C	0
Broad13912348	C/T	116904436	1 control het	T	0	T	0	T	0	T	0
Putative SNPs and indels from dbSNP											
rs10924104	A/G	116850426	N/A	A	0.154	A	0.154	N/D	N/D	N/D	N/D
rs12131793	A/G	116876285	N/A	A	0.022	G	0	N/D	N/D	N/D	N/D
rs12732875	C/G	116901937	N/A	C	0.015	G	0	N/D	N/D	N/D	N/D
rs35275493	-/T	116897025	N/A	T	0.154	T	0.075	N/D	N/D	N/D	N/D
rs11269409	-/G	116900506	N/A	-	0.015	-	0	N/D	N/D	N/D	N/D
rs1414273	C/T	116904172	N/A	T	0.142	C	0.142	N/D	N/D	N/D	N/D
rs2300746	C/G	116883706	N/A	G	0.149	G	0.149	N/D	N/D	N/D	N/D
rs10754444	C/T	116851199	N/A	C	0.154	C	0.154	N/D	N/D	N/D	N/D
rs1034919	C/G	116877028	N/A	C	0.157	C	0.157	N/D	N/D	N/D	N/D
rs10754445	A/T	116878982	N/A	T	0.154	T	0.154	N/D	N/D	N/D	N/D
rs1034920	C/T	116877922	N/A	C	0.152	C	0.152	N/D	N/D	N/D	N/D
rs4468196	A/G	116855744	N/A	A	0.152	A	0.152	N/D	N/D	N/D	N/D
rs10802191	A/T	116866606	N/A	A	0.152	A	0.152	N/D	N/D	N/D	N/D
rs12141411	A/G	116885933	N/A	A	0.145	A	0.145	N/D	N/D	N/D	N/D
rs10924106	C/T	116855268	N/A	C	0.154	C	0.154	N/D	N/D	N/D	N/D
rs10924105	C/T	116850738	N/A	C	0.396	T	0.396	N/D	N/D	N/D	N/D
rs7554038	A/G	116843211	N/A	A	0.205	A	0.205	N/D	N/D	N/D	N/D

Glossary: "-" denotes a deletion for one of the two indel (insertion/deletion) polymorphisms in the table; Chrom = chromosome; MAF = minor allele frequency; N/A = not applicable. N/D = not done (the SNPs from dbSNP were not genotyped in the East Asian HapMap samples).

Resequencing: Twenty-seven putative SNPs were identified from the generated sequence. Assays for 6 of these SNPs could not be designed as part of our multiplex assay, and 4 failed to genotype on the Sequenom platform. Of the 17 SNPs successfully genotyped in HapMap samples, all were seen at least once in a sequenced heterozygous individual. Seven of the 17 genotyped SNPs are polymorphic in the HapMap CEU samples. The 10 SNPs that are polymorphic in at least one HapMap population have been submitted to dbSNP and await assignment of rs numbers. Their temporary IDs are as follows: Broad13909142/rs56152595, Broad13909631/rs55704829, Broad13910476/rs55971473, Broad13911192/rs55982573, Broad13912407/rs55888098, Broad13912463/rs55842195, Broad13911505/rs56302466, Broad13909666/rs56383778, Broad13911980/rs56208406, Broad13909880/rs56258936.

dbSNP: A total of 136 putative SNPs were identified as being available in dbSNP but not in the HapMap resource. Thirty of these SNPs did not fit in our initial genotyping pool design for genotyping the HapMap samples. Fourteen of the remaining 106 SNPs or indels tested failed clustering analyses of the genotyping data generated by the Sequenom platform. Seventy-five of the remaining 92 SNPs were monomorphic in the CEU samples from HapMap. The remaining 17 SNPs are presented in the table above.

The SNP selection process for genotyping the BWH samples is described in detail in Table S2.

Table S4. The analysis conditional on rs12044852 and rs2300747 reveals no other evidence of association with MS within the sample of SNPs tested to date

SNP	BWH χ^2 *P* value	Condition on rs12044852, *P* value	Condition on rs2300747, *P* value
rs10924102	0.1288	0.1346	0.095
rs4839073	0.2479	0.6724	0.4573
rs17035862	0.0075	0.1553	0.06364
rs10924103	0.037	0.8512	0.6863
rs766429	0.0444	0.1504	0.186
rs12025416	0.0289	0.8693	0.5173
rs956184	0.0047	0.5438	0.5983
rs17576615	0.4085	0.3885	0.3833
rs3850814	0.0311	0.207	0.6234
rs9651076	0.0211	0.4065	0.9103
rs1109190	0.0076	0.3735	0.6551
rs4143029	0.0286	0.5058	0.5942
rs12031061	0.0032	0.9978	0.6231
rs10802189	0.6542	0.2346	0.2997
rs871493	0.8934	0.9997	0.8449
rs10802190	0.0066	0.9981	0.3626
rs10924108	0.0036	0.9979	0.5716
rs1414275	0.0076	0.2033	0.8186
rs1016140	0.0081	0.9978	0.6241
rs6677309	0.0215	0.4914	0.3897
rs758518	0.0085	0.9973	0.9142
rs12044773	0.0011	0.9982	0.4891
rs12044852	0.0003	NA	0.2119
rs35275493	0.0093	0.7616	0.9729
rs1335532	0.0228	0.5753	0.274
rs2300747	0.0004	0.6674	NA
rs55842195	0.4987	0.6727	0.6647

Glossary: BWH = Brigham and Women's Hospital; Aus = Australia; Bel = Belgium; C/C = case/control study; MAF = minor allele frequency; OR = odds ratio. For this analysis, the BWH case/control collection is pooled with the other collection of U.S. cases and controls from University of California, San Francisco, CA.

* MAF in parents of Australian trios = 0.12; MAF in parents of Belgian trios = 0.13; MAF in parents of United Kingdom trios = 0.11; and MAF in parents of U.S. trios = 0.10.

In CEU data from HapMap, the r^2 between rs2300747 and rs12044852 is 0.929.

The second column shows the *P* values derived from our Phase 1 analysis that are shown in greater detail in Table S2. NA, not applicable.

Table S5*a*. Fine mapping of the *CD58* locus in 1278 U.S. and United Kingdom trios

SNP	Minor allele	T:U	*P* value	OR (95% CI)
rs3850814	T	170:238	0.0008	0.71 (0.59–0.87)
rs9651076	A	175:242	0.001	0.72 (0.60–0.88)
rs4143029	C	186:267	0.0001	0.70 (0.58–0.84)
rs12031061	G	171:233	0.002	0.73 (0.60–0.89)
rs10924108	T	171:234	0.0017	0.73 (0.60–0.89)
rs1414275	A	175:237	0.0023	0.74 (0.61–0.90)
rs758518	A	174:246	0.0004	0.71 (0.58–0.86)
rs12044773	C	166:223	0.0039	0.74 (0.61–0.91)
rs12044852	C	166:232	0.0009	0.77 (0.65–0.91)
rs35275493	A	198:281	0.0001	0.70 (0.59–0.84)
rs1335532	A	208:294	0.0001	0.71 (0.59–0.84)
rs2300747	A	196:285	5×10^{-5}	0.73 (0.63–0.86)
rs10923122	C	503:524	0.5123	0.96 (0.85–1.08)
rs7542681	C	485:498	0.6784	0.97 (0.86–1.10)
rs570440	A	594:606	0.729	0.98 (0.88–1.10)

Glossary: T:U = ratio of transmitted to untransmitted minor alleles; OR = odds ratio, with a 95% confidence interval (CI). The listed allele is the minor allele of each SNP. We have highlighted in bold rs12044852, the SNP which was genotyped in the whole genome scan[5], as well as rs2300747, the SNP with the most extreme evidence of association in this TDT analysis. All SNPs were genotyped together in a single pool on the Sequenom iPLEX platform.

Table S5*b*. Extension of rs12044852 and rs2300747 genotyping to all available samples

Samples	All trios	U.S. C/C	United Kingdom C/C	Belgium C/C	Finland C/C	All samples	
Trios	1768					1768	
Cases		1557	961	348	692	3558	
Controls		855	2466	372	727	4420	
Samples			rs2300747 MAF				
Cases		0.093	0.11	0.10	0.17		
Controls		0.12	0.12	0.13	0.18		
SNPs			OR (95% CI)			OR (95% CI)	*P* value
rs12044852	0.77 (0.65–0.91)	0.77 (0.63–0.95)	0.89 (0.74–1.08)	0.90 (0.63–1.27)	0.93 (0.75–1.16)	0.84 (0.76–0.92)	$8.8 \times 10–5$
rs2300747	0.73 (0.63–0.86)	0.73 (0.61–0.89)	0.95 (0.80–1.12)	0.74 (0.53–1.02)	0.91 (0.75–1.11)	0.82 (0.75–0.89)	$1.1 \times 10–6$

Table S6. Regulatory T cells respond to increasing CD2 co-stimulation with increased FoxP3 expression

Stimulation:	IL-2 only	IL-2 αCD3$^{(1.0)}$	IL-2 αCD3$^{(1.0)}$ αCD2$^{(0.125)}$	IL-2 αCD3$^{(1.0)}$ αCD2$^{(0.5)}$	IL-2 αCD3$^{(1.0)}$ αCD2$^{(2.0)}$	IL-2 αCD3$^{(1.0)}$ αCD28$^{(0.5)}$
donor 1	72	303	435	586	678	246
donor 2	82	270	331	406	501	314

FACS sorted Treg were stimulated under specified conditions for 3 days before staining with anti-FoxP3 (mAb 206D, BioLegend). The data above represent mean fluorescence intensity for FoxP3.