ONLINE METHODS

Study design. Briefly, we conducted a fixed-effects meta-analysis of three whole-genome association scans for MS susceptibility (**Table 1**) based on the observed (imputed) and expected allele dosage of each SNP, taking into account the empirically observed variance of the allele dosage. As each dataset was generated on a different SNP array, we used the MACH algorithm[10] to impute genotypes on a single panel of 2.56 million SNPs; the meta-analysis was executed over this entire collection of SNPs. They were ranked based on the meta-analysis P values, and 186 SNPs outside of the MHC as well as two known MHC SNPs were selected for replication in an independent set of subjects (**Table 1**). We used the Cochran-Mantel-Haenszel method implemented in PLINK[44] to analyze the replication data, and we conducted a joint analysis by combining the results of the meta-analysis and replication studies using a sample size–weighted Z score (**Table 2**). Then, RNA data from PBMCs of subjects with MS and CIS (**Supplementary Table 1c**) were explored for evidence that the *IRF8* risk allele (rs17445836[G]) influences expression of genes in the interferon response pathway. We therefore explored broader effects on transcription in our PBMC RNA dataset (Affymetrix U133 2.0 plus) by analyzing the results of a comprehensive quantitative trait analysis across all RNA probes in relationship to rs17445836[G] using a gene set enrichment methodology[33].

Human subjects and genome-wide data in the meta-analysis. Summary information on the MS subjects and healthy control subjects are shown in **Table 1**. Below, we offer additional details on each sample set. All subjects with MS meet the McDonald criteria for a diagnosis of MS[45]. The demographic profiles of the sample collections are presented in **Supplementary Table 1a**.

In the initial MS genome-wide association scan by the IMSGC[4], subjects with MS were genotyped on the Affymetrix platform using the GeneChip Human Mapping 500K Array set. All healthy control subjects also have data generated on the Affymetrix platform using the same array. The MS subjects from the US were matched to healthy control subject data generated by the National Institute of Mental Health, and the MS subjects from the UK were matched to healthy control subject data generated by the Wellcome Trust. All details related to the quality parameters of these data (SNPs and subjects) are presented elsewhere[4]. Here, we use the same dataset that was analyzed in our earlier study, after removing subjects that were also genotyped in one of the other two studies: 54 subjects were genotyped in the original IMSGC study and the BWH study, and 82 subjects were genotyped in the IMSGC study and in the Gene MSA study. All of these duplicates were US subjects, and they were removed from the US component of the IMSGC dataset.

The BWH dataset is new. Its subjects with MS ($n = 860$) and healthy control ($n = 270$) subjects were genotyped on the Affymetrix Genome-wide Human SNP Array 6.0 (Genechip 6.0) at the Broad Institute's Center for Genotyping and processed for quality control (QC) using the PLINK software suite. We applied its standard quality control pipeline for subjects (genotype success rate $>95\%$, sex concordant, excess inter/intra-heterozygosity) and for SNPs (Hardy-Weinberg equilibrium $P > 1 \times 10^{-6}$; MAF > 0.01, genotype call rate >0.95; misshap test $>1 \times 10^{-9}$) to these data. In a second step, EIGENSTRAT[46] was used to identify population stratification outliers. Given the limited number of healthy control subjects that we had genotyped, we selected additional control subjects from an existing dataset of 2,681 subjects with genotypes generated by the MIGen consortium using the Affymetrix Genechip 6.0 platform at the Broad Institute's Center for Genotyping in a scan for loci associated with susceptibility to early myocardial infarction (MI). As there is no known association between MS and early MI, we selected control subjects for our analysis from a combined pool of (i) healthy control subjects from BWH that were recruited for MS studies (spouses and friends of MS subjects), (ii) healthy control subjects from the MIGen study, and (iii) early MI cases from MIGen. The SNP content was reduced to the 709,690 SNPs that had passed quality control in both studies. Each subject with MS in the BWH dataset was then matched to two subjects drawn from the combined control subject pool (MIGen and BWH subjects) using the first principal component distance calculated by EIGENSTRAT as described below. As recommended by the authors of EIGENSTRAT, we excluded X-chromosome SNPs from the calculation of eigenvectors.

The Gene MSA study consists of three sets of samples: (i) GeneMSA-NL: 253 subjects with MS and 208 healthy control subjects were collected at the Vrije Universiteit Medical Centre in Amsterdam, Netherlands, (ii) GeneMSA-CH: 230 subjects with MS and 232 healthy control subjects were collected at the University Hospital Basel, Basel, Switzerland, and (iii) GeneMSA-US: 486 subjects with MS and 431 healthy control subjects were collected at the University of California, San Francisco. All subjects were genotyped genome-wide in a single batch using the HumanHap550 Beadchip produced by Illumina. Details of the quality control pipeline for these data are described in detail in a prior publication[8].

Matching case and control subjects in the BWH discovery sample. As described above, we matched BWH MS cases to a pool of control samples consisting of subjects from the MIGen study as well as healthy control subjects from BWH. All of these subjects had Affymetrix 6.0 genome-wide SNP data, genotyped at the Broad Institute. We selected a subset of ancestry-informative markers (709,690 SNPs that passed stringent quality controls in both the MIGen study and our own study). We excluded X-chromosome SNPs and then used EIGENSTRAT[46] to define genetic eigenvectors. We then matched cases and controls using the following strategy. First, we randomly selected a subject with MS (case). Second, for each case we selected the most genetically similar control from the pool of available unmatched control subjects. We defined similarity with a Euclidean distance metric based on the top eigenvector:

$$d_{i,j} = \sqrt{(a_i - a_j)^2}$$

where $d_{i,j}$ is the distance between case i and control j, and a is the value of the case or control in the first eigenvector. Steps 1 and 2 are repeated until a total of two controls are selected for each case.

This resulted in a matched collection of 860 cases and 1,720 controls. Of note, this method was intended to match cases and controls on the basis of the top two eigenvectors, but an error in the code prevented the use of the second eigenvector in the calculation. The error was discovered in the final stage of manuscript preparation. The gain in matching from using a second eigenvector was tested and is marginal in this dataset.

Genome-wide genotype imputation and meta-analysis method. A meta-analysis was conducted in 9,844 unique individuals (2,624 cases and 7,220 controls) to identify genetic loci associated with multiple sclerosis. These samples came from three separate genome-wide association studies (six separate cohorts or strata) that are described in **Table 1**. To maintain the existing case-control relationships, our analysis approach involved combining the results of independent analyses performed in the six strata that are outlined in **Table 1**. Specifically, to control for population stratification, all individuals were stratified into 268 clusters on the basis of pairwise identity by state (IBS) between individuals within each of the six strata using PLINK: this process yielded 83 clusters in BWH, 73 clusters in IMSGC/UK, 46 clusters in IMSGC/US, 33 clusters in GeneMSA-US, 15 clusters in GeneMSA-NL, and 18 clusters in GeneMSA-CH. These clusters were defined using only those genotyped SNP data that were available in each cohort. In parallel, we used MACH version 1.0.5 (ref. 10) to impute the genotypes of 2,557,248 SNPs across the genome based on phased chromosomes (haplotypes) of the CEU population in HapMap release 21 (NCBI build 35). Imputation was conducted on all samples, ignoring case-control status, to avoid introducing artifacts between cases and controls. We elected to use probabilistic dosages in our statistical analysis rather than hard genotype calls to account for the uncertainty of imputation at each locus. Standard quality metrics were applied to the imputed data: we considered only those SNPs with $<5\%$ genotype missing rate, minor allele frequency >0.01, and Hardy-Weinberg $P > 10^{-6}$.

We conducted a fixed-effects meta-analysis across all clusters based on the observed and expected allele dosage, taking into account the empirically observed variance of the allele dosage:

$$z_{meta} = \frac{\sum (p_o - p_e)}{\sqrt{\sum \mathrm{var}(p)}}$$

where p_o and p_e are the cumulative observed and expected allele dosage in cases per cluster, respectively. We take into account imputation uncertainty by taking the empirically observed variance of the allele dosage (computed per cohort) for var(p) if the average maximal posterior probability of an imputed SNP

© 2009 Nature America, Inc. All rights reserved.
npg

 doi:10.1038/ng.401

<0.99 (that is, for poorly imputed SNPs). Otherwise, if the average maximal posterior probability of an imputed SNP >0.99 (that is, for well imputed SNPs), we take for var(p) the binomial variance of the allele dosage (which is equal to $p(1-p)$).

For the 2,557,248 SNPs examined in 9,813 individuals, the genomic inflation factor was $\lambda = 1.054$. Given the unique role of the major histocompatibility complex (located on chromosome 6) in MS, we also computed the genomic inflation factor after excluding SNPs found on chromosome 6: $\lambda = 1.048$. There were 2,142 SNPs that reached genome-wide significance ($P < 5 \times 10^{-8}$) in the meta-analysis. Of these, 2,141 SNPs were on chromosome 6, and one SNP (rs12025416, $P = 4.7 \times 10^{-8}$) was found within the *CD58* locus on chromosome 1. The most significant SNP (rs6931337, $P = 2.8 \times 10^{-167}$) was on chromosome 6. A SNP within HLA-DRA (rs3135388, $P = 7.4 \times 10^{-164}$) previously identified to be associated with MS is in high LD with rs6931337. As rs3135388 has been used previously as a surrogate for the HLA-DRB1*1501 susceptibility allele, we used it again for this purpose in the replication analysis.

Identifying independent blocks of association for replication. To organize our top results of our meta-analysis and select loci for the replication study, we used the "clump" routine from PLINK[44]. We applied an iterative approach where the marker with the most extreme evidence of association was used as the starting point and all other SNPs with an $r^2 > 0.5$ with best marker were grouped into one locus. The process was then repeated until 100 independent loci were defined. Known susceptibility alleles were included, based on previous work (ref. 4 and IMAGEN consortium, unpublished data). For HLA-B*4402, which emerged from a parallel set of analyses by the IMAGEN consortium (J. Oksenberg, unpublished data), the best surrogate marker (rs2743951) could not be designed as a SNP assay, and therefore rs2523393 was selected based on its strong LD ($r^2 = 0.92$) with rs2743951 in HapMap CEU samples (J. Oksenberg, unpublished data).

Subjects used in the replication analysis. There are two strata in our panel of replication samples (**Table 1**). As we do not have genome-wide data on these individuals, we matched them by country of origin and limited the analysis to subjects of self-reported European ancestry. **Supplementary Table 1b** contains the pertinent demographic details of these subjects. The UK component consists of an additional 831 subjects with MS collected at the University of Cambridge. These cases are matched to 1,030 subjects from the 1958 birth cohort who had not been genotyped genome-wide as part of the Wellcome Trust project and therefore represent an independent set of UK control samples from those used in the meta-analysis. The US stratum of the replication panel consists of subjects with MS from four different collections (demographic details are provided in **Supplementary Table 1b**): (i) BWH, 228 cases and 14 healthy controls; (ii) Accelerated Cure Project, 597 cases and 35 healthy controls; (iii) Washington University. 152 cases and 13 healthy controls; and (iv) UCSF, 407 cases and 142 healthy controls. To supplement the pool of healthy control subjects, we also included subjects of self-declared European ancestry from (i) the PhenoGenetic project, 292 healthy control subjects from a tissue bank of samples from subjects recruited in the Greater Boston metropolitan area to provide fresh blood for immunogenetic and other analyses; (ii) the healthy control collection at the Harvard/Partners Center for Genetics & Genomics, 101 healthy control subjects recruited from the Greater Boston metropolitan area (see URLS section below), and (iii) 489 healthy control subjects from the Chicago Health and Aging Project, a population-based study of healthy, nondemented, aging subjects centered in a suburb of Chicago[47].

Genotyping platforms. The platforms used to generate genome-wide data in each component of our meta-analysis are listed above in the description of these components. The Sequenom MASS Array platform in its iPLEX format was used to genotype the panel of SNPs selected for replication in the replication samples. Of the 188 SNPs selected for replication, 180 SNPs had data that met our quality control parameters: HWE $P > 1 \times 10^{-6}$; MAF > 0.01, genotype call rate >0.95.

Replication analysis and joint analysis. The replication analysis was conducted using a Cochran-Mantel Haenszel (CMH) approach as implemented in PLINK[44]. As genome-wide data were not available, we divided subjects (cases and controls) into two strata based on the country in which the sample was collected (US or UK); these two strata were used in the CMH analysis that is reported in detail in **Supplementary Table 3**. To perform a joint analysis of the meta-analysis and replication data, we calculated Z scores for each of the two components of the analysis that were then added to calculate a joint Z score (based on an estimated effective sample size (cases + controls) of 4,000 subjects for the meta-analysis and 4,500 subjects for the replication stage), from which the final P values are determined. This approach is described in detail in a prior publication[11].

Logistic regression for assessing independence of loci. To assess whether SNPs in the same locus may have distinct effects on susceptibility to MS, we implemented a logistic regression analysis using stepwise selection, with the rank 1 SNP (SNP with most extreme P value) being forced into the model first. We then calculated the residual effect of each of the other SNPs after accounting for the effect of the SNP with the most extreme evidence of association. The covariate in the model is the country of origin (US/UK) to account for possible population heterogeneity between the US and UK samples.

Secondary analyses with a covariate for gender. In the secondary analysis of the data included in our meta-analysis, we implemented a logistic regression analysis, with case-control as the outcome variable, and cohort of origin (six cohorts outlined in **Table 1**) and gender as two covariates in the model to account for possible heterogeneity between the cohorts and different sex ratio between case and control groups of the six cohorts. We selected 41 of the top SNPs that were not redundant with the results of the meta-analysis for inclusion in the replication effort. We used the same method to perform a secondary analysis of the replication data (see **Supplementary Table 7**); in that case, a covariate for the country of origin (UK or US) was included to minimize the possible effect of population stratification.

RNA data and analysis. Between July 2002 and October 2007, PBMC samples were collected from relapsing-remitting MS subjects and CIS subjects as part of the Comprehensive Longitudinal Investigation of MS at the Brigham & Women's Hospital (**Supplementary Table 1c**)[48]. CIS subjects differ from MS subjects by having had only one clinical episode of demyelination; MS subjects, by definition, must have at least two such events or one event and evidence of disease activity in a paraclinical measure such as MRI[45]. Nonetheless, the pathophyisology is shared between these two sets of subjects, and they are treated in the same manner in a clinical environment[49]. 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 micro-centrifuge 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). 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 were 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 Poly-A control kit, and (v) the ratio of intensities of 3′/5′ probes for the housekeeping genes *GAPDH* and *β-actin*. We then normalized the data using GCRMA.

From our collection, 240 RNA profiles met our QC criteria and had genotypes for rs1800693 (*TNFRSF1A*), rs17445836 (*IRF8*) and rs17824933 (*CD6*). Next, we analyzed the correlation of rs17445836 (*ICSBP1*) with probes from its three target genes: *AICD*, *BCL6* and *TLR4*. For these analyses, all

© 2009 Nature America, Inc. All rights reserved.
npg

subjects were considered together. We also performed secondary analyses in each treatment group (untreated, $n = 82$; IFNβ-treated, $n = 94$; GA-treated, $n = 64$), and these individual results mirrored the results obtained in the pooled analysis in each case (data not shown).

Gene set enrichment analysis of quantitative trait analysis results We implemented a quantitative trait analysis, using the Wald test as implemented in PLINK, of the rs17445836 polymorphism with all 54,676 Affymetrix U133 2.0 Plus probes; in all, 22,757 genes are sampled. Each treatment category (subjects are either untreated, IFNβ-treated and GA-treated) was analyzed separately using the normalized expression data described in the previous section. Only probes with an expression value >3 in each of the 240 subjects were considered for the next phase of the analysis. We use the Wald test as implemented in PLINK[44] for our quantitative trait analysis. The β value from the Wald test is used as the input variable of each probe in downstream analyses. If two or more probes mapped to the same gene, they were collapsed into one mean β value for that gene.

Gene set enrichment analysis (GSEA) version 2.0.1 (ref. 33) and the manually curated section C2 of MSigDB database were used in our subsequent analyses. Gene sets were preprocessed to exclude gene sets which contained <15 or >200 genes from our collection of 828 probes from untreated subjects that exceeded our threshold of $P < 0.05$ in the quantitative trait analysis. 93 out of a possible 1,892 gene sets met these criteria and were tested in our analysis. We performed 1,000 permutations of the analysis using the weighted enrichment statistic to estimate the statistical significance of our results. The process was repeated for the IFNβ-treated subjects (1,413 probes and 147 gene sets met our criteria and were tested) and the GA-treated subjects (3,191 probes and 223 gene sets). We consider replicated those results that were associated in both the untreated subjects and at least one of the treatment categories at an FDR q value <0.05 (with the same direction of effect). Lack of replication between the two treated categories is difficult to interpret given the different mechanisms of action for GA and IFNβ. Sixteen genes tested met this criterion of replication (**Table 3**). Detailed results are presented in **Supplementary Table 6a**. The GA-treated group (the smallest subject group) does not have a significant overlap with either of the other two groups of subjects.

The GSEA report for each gene set includes (i) the number of genes used to evaluate a particular gene set, (ii) a normalized enrichment score (NES) which accounts for differences in gene set size and number of permutation performed, and (iii) a false discovery rate (FDR) q value, a measure of statistical significance that accounts for the number of hypotheses tested.

Ingenuity analysis. The Ingenuity Pathway Analysis (IPA) tool (IPA Tool; Ingenuity Systems) was used to test whether its pre-defined "canonical interferon response pathway" was enriched in genes whose expression is correlated with the rs17445836[G] allele in our dataset generated from the PBMCs of subjects with MS. The data file we uploaded into this analysis tool is the same that was explored using the Gene Set Enrichment method described above. In short, it consists of all Affymetrix probesets who meet a $P < 0.05$ threshold in our quantitative trait analysis testing an additive model of association with rs17445836[G] (see previous section). For each probeset, its Affymetrix probeset ID and its β value from the quantitative trait analysis are loaded into the analysis tool. Each probeset ID is mapped to its corresponding gene object in the Ingenuity Pathways Knowledge Base.

The interferon signaling pathway that we have tested is the one found in the Ingenuity Pathways Analysis library of canonical pathways. This pathway contains 29 genes. Since we are testing a single hypothesis (association of higher interferon pathway gene expression with to rs17445836[G]) in this analysis, we use Fisher's exact test to derive a P value that estimates the significance of the enrichment of interferon pathway genes in the list of genes that we have found to be associated with rs17445836[G].

URLs. MACH algorithm, http://www.sph.umich.edu/csg/abecasis/MACH/download/; Infevers, http://fmf.igh.cnrs.fr/ISSAID/infevers; Ingenuity Systems, http://www.ingenuity.com; http://www.hpcgg.org/BiosampleServices/overview.jsp.

44. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
45. McDonald, W.I. *et al.* Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann. Neurol.* **50**, 121–127 (2001).
46. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
47. Bienias, J.L., Beckett, L.A., Bennett, D.A., Wilson, R.S. & Evans, D.A. Design of the Chicago Health and Aging Project (CHAP). *J. Alzheimers Dis.* **5**, 349–355 (2003).
48. Gauthier, S.A., Glanz, B.I., Mandel, M. & Weiner, H.L. A model for the comprehensive investigation of a chronic autoimmune disease: the multiple sclerosis CLIMB study. *Autoimmun. Rev.* **5**, 532–536 (2006).
49. Miller, D., Barkhof, F., Montalban, X., Thompson, A. & Filippi, M. Clinically isolated syndromes suggestive of multiple sclerosis, part 2: non-conventional MRI, recovery processes, and management. *Lancet Neurol.* **4**, 341–348 (2005).

© 2009 Nature America, Inc. All rights reserved.