Published in final edited form as:

Hum Genet. 2010 March ; 127(5): 525–535. doi:10.1007/s00439-010-0789-4.

Genetic Variation in the IL7RA/IL7 Pathway Increases Multiple Sclerosis Susceptibility

Rebecca L. Zuvich^{1,*}, Jacob L. McCauley^{2,*}, Jorge R. Oksenberg^{3,4}, Stephen J. Sawcer⁵, Philip L. De Jager^{6,7}, Consortium International Multiple Sclerosis Genetics, Cristin Aubin⁷, Anne H. Cross⁸, Laura Piccio⁸, Neelum T. Aggarwal^{9,10}, Denis Evans^{9,10}, David A. Hafler^{7,11}, Alastair Compston⁵, Stephen L. Hauser³, Margaret A. Pericak-Vance², and Jonathan L. Haines¹

¹ Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, TN 37232-0700, USA

² John P. Hussman Institute for Human Genomics, University of Miami, Miller School of Medicine, Miami, FL 33136, USA

³ Department of Neurology, School of Medicine, University of California, San Francisco, CA 94143-0435, USA

⁴ Institute for Human Genetics, School of Medicine, University of California, San Francisco, CA 94143-0435, USA

⁵ University of Cambridge, Department of Clinical Neurosciences, Addenbrooke's Hospital, BOX 165, Hills Road, Cambridge, CB2 2QQ, UK

⁶ Program in Translational NeuroPsychiatric Genomics, Department of Neurology, Brigham & Women's Hospital and Harvard Medical School, Boston, MA 02115

⁷ Program in Medical & Population Genetics, Broad Institute of Harvard University and Massachusetts Institute of Technology, Cambridge, MA 02139, USA

⁸ Department of Neurology, Washington University, St. Louis, Missouri, USA

⁹ Rush Alzheimer Disease Center & Department of Neurological Sciences, Rush University, Chicago, Illinois, USA

¹⁰ Rush Institute for Healthy Aging, Rush University, Chicago, Illinois, USA

¹¹ Department of Neurology, Yale School of Medicine, New Haven, CT, USA

Abstract

Multiple sclerosis (MS) is characterized as an autoimmune demyelinating disease. Numerous family studies have confirmed a strong genetic component underlying its etiology. After several decades of frustrating research, the advent and application of affordable genotyping of dense SNP maps in large datasets has ushered in a new era in which rapid progress is being made in our understanding of the genetics underlying many complex traits. For MS, one of the first discoveries to emerge in this new

Corresponding author: Rebecca L. Zuvich, Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, TN 37232-0700, PH: 615-936-2719, Fax: 615-343-8619, rebecca.l.zuvich@vanderbilt.edu.

*These authors contributed equally to this work.

Author contributions: RLZ, JLM, and JLH designed the study and wrote the manuscript. RLZ, JLM, and JLH performed analytical work. RLZ performed the quality control analysis. IK, PLW, and CA generated and processed genotype data for analysis. PLD, MB, SP, RB, SR, NTA, LP, WLM, DP, DE, AHC, BC, JDR, LFB, AJI, AC, DAH, SLH, JRO, SJS and MPV contributed to DNA sample collection, genetic data, discussion, and edits to the manuscript. All authors have read and contributed to the manuscript.

era was the association with rs6897932[T244I] in the interleukin-7 receptor alpha chain (*IL7RA*) gene (Gregory et al. 2007; International Multiple Sclerosis Genetics Consortium 2007; Lundmark 2007), a discovery that was accompanied by functional data that suggest this variant is likely to be causative rather than a surrogate proxy (Gregory et al. 2007). We hypothesized that variations in other genes functionally related to *IL7RA* might also influence MS. We investigated this hypothesis by examining genes in the extended biological pathway related to *IL7RA* to identify novel associations. We identified 73 genes with putative functional relationships to *IL7RA* and subsequently genotyped 7,865 SNPs in and around these genes using an Illumina Infinium BeadChip assay. Using 2,961 case-control dataset, two of the gene regions examined, *IL7* and *SOCS1*, had significantly associated single-nucleotide polymorphisms (SNPs) that further replicated in an independent case-control dataset (4,831 samples) with joint p-values as high as 8.29×10^{-6} and 3.48×10^{-7} , respectively, exceeding the threshold for experiment-wise significance. Our results also implicate two additional novel gene regions that are likely to be associated with MS: *PRKCE* with p-values reaching 3.47×10^{-4} and *BCL2* with p-values reaching 4.32×10^{-4} . The *TYK2* gene, which also emerged in our analysis, has recently been associated with MS (Ban et al. 2009). These results help to further delineate the genetic architecture of MS and validate our pathway approach as an effective method to identify novel associations in a complex disease.

Introduction

Multiple sclerosis (MS) is an idiopathic autoimmune demyelinating disease that strikes in early to mid-adult life and affects women more often than men. MS, like most common diseases, has several complicating characteristics including clinical and genetic heterogeneity, incomplete penetrance, fluctuations over time, polygenic inheritance, and substantial environmental risk factors (Compston and Coles 2002). Numerous family and twin studies have shown that there is a strong genetic component underlying the etiology of MS. The prevalence of the disease among first-degree relatives of affected individuals is substantially increased over the prevalence in the general population, and twin studies suggest an overall heritability of ~50% (Ebers, Sadovnick, and Risch 1995; Hauser and Goodkin 1998; Mumford et al. 1994; Robertson et al. 1996; Sadovnick et al. 1993; Sadovnick and Ebers 1995; Sadovnick et al. 1996).

In the 1970s the Major Histocompatibility Complex (MHC) association with MS was discovered (Bertrams and Kuwert 1972; Naito et al. 1972). However, linkage screens failed to consistently identify any regions of linkage outside the MHC (Ebers et al. 1996; Haines et al. 1996; Kuokkanen et al. 1997; Sawcer et al. 1996; Transatlantic Multiple Sclerosis Genetics Cooperative 2001; International Multiple Sclerosis Genetics Consortium 2005). Furthermore, although hundreds of candidate genes have been tested, none had gained consensus as carrying risk for MS (Hirschhorn et al. 2002). This slow progress made it clear that more powerful approaches were needed in the genetic analysis of MS. Genomic convergence was applied to explicitly identify candidate genes with the best genetic and biological relevance (Hauser et al. 2003). In 2007, more than three decades after HLA was first identified, the interleukin-7 receptor alpha chain (*IL7RA* or CD127) was identified as the first non-MHC MS susceptibility locus using this approach (Gregory et al. 2007) and simultaneously confirmed in an accompanying report (Lundmark 2007).

The associated SNP was also identified in the first genome-wide association study (GWAS) in MS (International Multiple Sclerosis Genetics Consortium 2007) and in the non-synonymous SNP (nsSNP) screen performed by the Wellcome Trust Case Control Consortium (WTCCC) (Wellcome Trust Case Control Consortium 2007). The original GWAS study also identified interleukin-2 receptor α gene (*IL2RA* or CD25) as an MS risk gene, and highlighted the likely role of additional loci including *CLEC16A* and *CD58*. The effects of these additional

genes were replicated in multiple independent datasets confirming their status as MS associated loci. In addition, other genes were identified and replicated as MS risk genes, including *TYK2*, *CD226*, *TNFRSF1A* (or CD120a), *IRF8*, and *CD6* (Australia and New Zealand Multiple Sclerosis Genetics Consortium 2009; Ban et al. 2009; Baranzini, Wang, and Gibson 2009; De Jager et al. 2009; International Multiple Sclerosis Genetics Consortium 2009a; Kumpfel et al. 2008).

Rs6897932, a non-synonymous coding SNP in *IL7RA*, plays a functional role in MS, as the major allele (C) increases skipping of exon 6 and thus increases the soluble form of *IL7RA*. *IL7RA* is a likely functional candidate for this complex disease because it is involved in the development, maturation, and homeostasis of T and B cells. It is involved in two signaling pathways: *IL7* and thymic stromal lymphopoietin (*TSLP*) (Gregory et al. 2007).

IL7 is an essential cytokine for lymphocyte development. Its receptor consists of two components: *IL7RA* and *IL2RG* (gamma c or CD132); the latter also heterodimerizes with receptors for other interleukins, including *IL2*, *IL4*, *IL9*, *IL15*, and *IL21*. Both components are essential for the biological effects of *IL7* and are also required for the downstream signaling cascade, which involves the activation of Janus kinases (Jaks) and signal transducer activators of transcription (STATs). The downstream signaling cascades are involved in various cell processes including development of B and T cells, regulation of cell proliferation, regulation of survival for progenitor and mature lymphoid cells, and gene transcription following an immune response (Hofmeister et al. 1999; Jiang et al. 2005). We hypothesized that genes in the *IL7RA* pathway may have independent or interactive effects with the risk SNP in *IL7RA* (Gregory et al. 2007).

Other complex diseases, including schizophrenia and Crohn's disease, have employed the approach of examining the extended pathways identified through known susceptibility genes to identify new disease loci (Hennah et al. 2007; McGovern et al. 2009). With the prior knowledge that *IL7RA* is associated, it is plausible that this pathway (*IL7/IL7RA*) is involved in the autoimmune aspects of the disease. We designed our study to investigate the role of genes in the *IL7RA* pathway in MS susceptibility both for independent and interactive gene effects. We selected genes as part of this pathway if they were directly involved in the downstream signaling cascade or were known to be influenced by varying *IL7* concentrations.

Results

After all quality control was completed, the discovery dataset consisted of 7,637 SNPs genotyped in 1,343 cases and 1,379 controls. The results using the Armitage trend test found 578/7,637 nominally significant SNP results with estimated genotype relative risks (GRRs) of 1.11-1.64 (over 75% have GRRs<1.20). None of these SNPs met a Bonferroni experiment-wise correction, although one SNP in *IL7* (rs2587156) came close (p-value= 2.46×10^{-5}). There are seventeen SNPs of particular interest (p-values< 1×10^{-3} , Table I) (the full set of significant SNPs is given in Suppl. Table 1), representing five gene regions. The genotypic and haplotypic analyses did not yield any significant novel regions that were not detected using the Armitage trend test. Likewise, the Multi-factor dimensionality reduction (MDR) results did not display any significant 2-way or 3-way interactions. We selected 31 SNPs including the seventeen top SNPs and fourteen additional SNPs with suggestive p-values (p-values< 1×10^{-2} , Supplementary Table 1) from these five gene regions. We chose to focus on those most significant results (p-values < 1×10^{-3}) based on what we could reasonably follow up on in a replication study. To get better coverage of those five regions of particular interest, we included additional SNPs giving us a total of 31 SNPs.

Following QC analyses, 24/31 SNPs were analyzed in the independent validation dataset of 2,111 cases and 2,037 controls. We observed 11/24 significant SNPs (p -values <0.05) with estimated GRRs of 1.11-1.25 (Table II). In addition, we merged the two independent datasets together to obtain the overall Armitage p -values for all 3,454 cases and 3,416 controls. The SNPs on chromosome 16 flanking *SOCS1* were the most significant having an estimated GRR of 1.25; the most significant p -value (3.48×10^{-7}) was with SNP rs441349. *IL7* also had significant SNPs as well with estimated GRRs of 1.14 and 1.35 (Table III) with p -values of 6.92×10^{-4} and 8.29×10^{-6} for SNPs rs1054283 and rs2587156, respectively. Because of gender ratios in the replication dataset were different from the discovery set, we also examined the stratified dataset to confirm that gender is not a confounding variable. We saw no differences in the gender specific results (Table IV).

The high risk allele frequency of the associated *IL7RA* variant meant that our analysis based on stratifying according to this locus had only limited power. It was thus not surprising that no novel finding emerged. The HLA stratified data was more interesting. There were 2,584/6,870 individuals with at least one risk allele ("A" allele for rs3135388) (IMSGC 2007). The Armitage trend test for those individuals without a risk allele showed *IL7* had more significant p -values than in the overall dataset. For those individuals with the risk allele (at least one copy), SNPs in the flanking region of *SOCS1* yielded the most significant p -values for the Armitage trend test (Table V). However, a test of heterogeneity was only significant for the SNPs in *IL7* (p -values=0.03-0.04).

Discussion

Through our analysis of the genes implicated as part of the *IL7R* pathway, we identified two novel gene regions, *IL7* and *SOCS1*, replicated these in an independent dataset, and have highly significant overall results. Finding association with the biologically relevant ligand (*IL7*) of *IL7R* indicates that the known biological activity of these genes is relevant to MS, as opposed to some perhaps as yet unknown pleiotropic effect of these genes. The *IL7/IL7R* complex along with *IL2RG* is vital for the development, maturation, and survival of T cells, which is theorized to be where the breakdown in the body's immune response occurs. This breakdown could, therefore, play a role in the adverse autoimmunity found in MS patients (Goverman 2009).

Interestingly, *SOCS1* is located on chromosome 16 slightly downstream from a recently confirmed MS susceptibility locus, *CLEC16A*. After examining the linkage disequilibrium (LD) architecture of this region, *SOCS1* and *CLEC16A* appear to be two independent MS disease loci (the r^2 values between the *CLEC16A* SNP (rs7184083) and the most significant SNPs in the *SOCS1* region are 0.057 (rs1640923) and 0.049 (rs12922733)). The LD block containing *CLEC16A* is separated from the more proximal LD block that contains the *SOCS1* gene region that contains our significant SNPs. Having an associated SNP in the flanking region of a gene involved in suppression of cytokine signaling (i.e. *SOCS1* functions as a suppressor of cytokine signaling) suggests that this flanking region could be a regulatory element. Having multiple associated genes in close proximity may suggest that these genes could be co-regulated. These observations warrant further investigation of this region via functional studies.

We also identified three additional strong candidate genes that have biological relevance to the etiology of MS, including Protein kinase C epsilon (*PRKCE*), B-cell lymphoma 2 (*BCL2*), and tyrosine kinase 2 (*Tyk2*). *PRKCE* is involved in neuron channel activation and apoptosis. *BCL2* is involved in activation of apoptosis. *Tyk2* transmits cytokine signals by phosphorylating receptor subunits (Jiang et al. 2005; Shirai, Adachi, and Saito 2008). Through independent efforts, we have already identified *Tyk2* as a susceptibility gene in MS. Its re-

emergence here thus provides a valuable positive control and further support for the importance of *Tyk2* (Ban et al. 2009).

The results indicate that the multiple genetic variations in the *IL7/IL7RA* pathway are important in MS pathogenesis, and by implication, that this pathway plays an important biological role in the etiology of MS. Furthermore our data support the approach of using candidate pathways as an effective way of identifying additional MS susceptibility genes. Examining the extended pathways of previously identified associated genes has recently proven useful in the identification of new disease loci (Hennah et al. 2007; McGovern et al. 2009). Therefore, this method of screening the biologically interacting genes of confirmed disease genes may help to elucidate additional aspects of the underlying genetic architecture of numerous complex diseases.

These data have also identified several genes that have modest significance and should be examined further. Some of the genes mentioned previously (i.e. *PRKCE* and *BCL2*) should be investigated in independent datasets. They were not strongly replicated, but had several SNPs with nominally significant p-values in the discovery dataset and trending favorably in the validation dataset. Thus, additional examination of these genes is necessary. Additionally, other genes (i.e. *PRKCA*, *Stat5a/b*, *IL12A/B*, *T-BET*, *BAX*, and *BIM*) also had several SNPs with interesting trends that were not followed up in this study. All of these genes have biological plausibility and need further exploration.

The HLA-DR*1501 allele identified decades ago confers a much larger effect on MS susceptibility (odds ratio (OR) ~3) than the effects of recently identified associations. Both the *IL7RA* and the *IL2RA* alleles have much smaller effects (OR~1.2 and OR~1.25, respectively) (Gregory et al. 2007; International Multiple Sclerosis Genetics Consortium 2007). It is possible that common alleles with smaller effects could explain a significant proportion of the genetic component of MS (Baranzini et al. 2009; Sawcer 2008; Wang et al. 2005). In this study of the *IL7/IL7RA* pathway, the nominally significant SNPs all have estimated GRRs of 1.11-1.64, and over 75% have GRRs<1.20. Using the pathway approach in conjunction with these large datasets and dense SNP genotyping, the etiology of MS is finally starting to be dissected. By building on the knowledge of these smaller gene effects, functional studies can help us understand the underlying mechanism of these genes and their relationships to the pathogenesis of MS.

Methods

Samples

Samples for the discovery screen came from both the United States and the United Kingdom. Through the International Multiple Sclerosis Genetics Consortium (IMSGC) we compiled a total of 2,961 samples (1,479 cases and 1,482 controls). The McDonald criteria were used to provide a positive diagnosis for MS (McDonald et al. 2001). Healthy unrelated controls were selected from the same US sites and from the 1958 Birth Cohort in the UK. Controls were selected to match the appropriate ratio of female to male cases. Controls were also age-matched to the cases when possible. Subjects were chosen with all four main types of MS and varying disability. In the UK subset, those with clinically isolated syndrome (CIS) were excluded from this study; however, about 7% of the US subset had CIS at their examination dates, as the US experience indicates that these individuals virtually always develop MS. All participants self-reported as non-Hispanic whites. The table of the demographics of these samples can be found elsewhere (Table 4 of International Multiple Sclerosis Genetics Consortium 2009b).

Samples for the replication dataset also were obtained from the United States and the United Kingdom through the IMSGC. The same criteria were applied for diagnosis (McDonald et al.

2001). The controls were collected from the same ascertainment sites as the cases with the UK controls coming from the 1958 Birth Cohort. Like the discovery dataset, the UK cases with CIS were removed; whereas, CIS cases only made up 4% of the US cases. All samples were self-reported non-Hispanic whites. This dataset consisted of 4,831 samples (2397 cases and 2434 controls) (Table VI). Genotype concordance tests were performed to exclude any possible duplicate samples, thereby ensuring that the two datasets (discovery dataset and the replication dataset) are independent.

Genotyping

We identified 73 genes demonstrating putative biological involvement within the *IL7RA* pathway (Table VII) by performing a PubMed search (Ansel et al. 2006; Fry and Mackall 2005; Gregory et al. 2007; Hofmeister et al. 1999; Jiang et al. 2004; Jiang et al. 2005; Kisseleva et al. 2002; Motsinger et al. 2007; O'Shea 2004). All known SNPs within the genomic sequence of each gene and 50 kb flanking upstream and downstream of the gene were identified using dbSNP version 127 (for complete coordinates of these genes see Suppl. Table 2). The flanking regions of each gene were included to thoroughly examine the gene and any possible regulatory regions surrounding the gene.

We used the Illumina Infinium 60K BeadChip assay as our genotyping platform (Stemers et al. 2006) for the initial screening experiment. Both case and control samples were included on each BeadChip in a randomized order. With additional collaborators from the International Multiple Sclerosis Genetics Consortium (IMSGC), we aggregated a total of 48,767 SNPs for multiple projects into a single 60K BeadChip design. This study selected and used 7,865 SNPs for the 73 gene regions. These SNPs represented 16% of the total 60K BeadChip. The criteria for picking SNPs were: minor allele frequencies greater than 5%, validated SNPs (from dbSNP), and Illumina Infinium score greater than 0.60 (proprietary score used to determine the overall success of the assay). The Illumina protocol involves amplification of genomic DNA, fragmentation of this DNA, hybridization of the DNA on the BeadChip, extension on the BeadChip, and imaging to read the chip (Stemers et al. 2006).

For the validation experiment, the Sequenom MassARRAY iPLEX platform was used. The Sequenom protocol involves a multiplex PCR reaction preceding a single base primer extension reaction. The individual SNPs are identified by using Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI TOF MS). The protocol for this platform is described in detail elsewhere (Gabriel, Ziaugra, and Tabbaa 2009). In total, 90 SNPs were genotyped across 3 Sequenom pools for the aggregated multiple IMSGC projects. This study selected 31 SNPs across five gene regions. The SNPs chosen for replication were those with the highest Armitage p-values and were derived from five biologically relevant gene regions.

Quality Control (QC)

Upon completion of genotyping, QC processes were run for the entire BeadChip to guarantee the accuracy of the genotyped dataset. The discovery dataset contained 2,961 individuals (1,479 cases and 1,482 controls). Using the Whole-genome Association Study Pipeline (WASP) (Sexton et al. 2007), we tested for SNP and sample genotyping efficiency, allele frequencies, gender errors, and Hardy-Weinberg equilibrium (HWE). To ensure thorough cleaning of the data, this process was performed recursively.

Overall, individuals were dropped based on low genotyping efficiency (<95%), on gender discrepancies (X-chromosome heterozygosity scores), on apparent relatedness (using IBS statistics in Graphical Relationship Representation), and on population stratification (using Structure). In total, 239 individuals were eliminated: 92 based on low genotyping efficiency, 68 based on gender discrepancies and/or apparent relatedness, and 79 based on population

stratification. Over 800 X-chromosome SNPs were on the BeadChip and could be used to determine possible gender discrepancies. Population stratification was investigated, and samples were dropped if the probability was <0.90 of being from European descent.

SNPs were eliminated based on low genotyping efficiency ($<95\%$), on the absence of a minor allele in our dataset (monomorphic), and on the deviation from HWE (<0.0001). Seventy-three SNPs were out of HWE in the overall dataset but were in HWE (>0.05) in either the cases alone or the controls alone and were subsequently retained for analysis.

After application of stringent QC, we eliminated $\sim 8\%$ of the samples and $\sim 4\%$ of the SNPs. The estimate of residual genotypic error (based on X-chromosome heterozygosity in males) was 1.87×10^{-3} . Overall, QC was completed yielding 2,722 individuals (1,343 cases and 1,379 controls) and 46,874 SNPs. This QC process was the same for this experiment and for an extensive follow-up of GWAS data from the IMSGC (International Multiple Sclerosis Genetics Consortium 2009b). This project included 7,637 SNPs (16.3% of the total BeadChip SNPs passing QC).

Similar QC processes were applied to the 31 SNPs in the replication phase. The replication dataset contained 4,831 individuals (2,398 cases and 2,433 controls). Using the Whole-genome Association Study Pipeline (WASP), we tested for marker and sample genotyping efficiency, allele frequencies, and Hardy-Weinberg equilibrium (HWE). Overall, 474 individuals were eliminated based on low genotyping efficiency ($<85\%$) or the individuals self-reporting as other than non-Hispanic Caucasians. In addition, 219 individuals were eliminated due to overlap with the discovery dataset. No X-chromosome SNPs were run so gender discrepancies could not be addressed, and there were not enough SNPs genotyped to look at sample relatedness. Markers were eliminated based on low genotyping efficiency ($<95\%$) and deviation from HWE (<0.0001). Four SNPs had genotyping efficiencies lower than 95%, and three SNPs were out of HWE. The QC process eliminated $\sim 14.3\%$ of the samples and $\sim 22.5\%$ of the SNPs. Overall, QC was completed yielding 4,148 individuals (2,111 cases and 2,037 controls) and 24 SNPs.

Analysis

Armitage trend tests and genotypic chi square analyses were performed using WASP software for the minor allele. In addition, haplotype blocks were identified and examined for haplotypic associations using HaploView (Barrett et al. 2005) and haplo.stats from the R project (<http://www.r-project.org/>). The Armitage trend test and genotypic chi square analyses were performed for all 7,637 SNPs on all individuals using an additive scoring model for the minor allele. The haplotypic analysis was performed on the “tagging SNPs” that tag (using r^2 value=0.80) only the 578 nominally significant (p-values <0.05) SNPs from the Armitage trend test on all individuals. Rs3135388, a SNP that is highly predictive of the HLA-DRB1*1501 allele, was included on the chip for purposes of secondary analyses such as stratification and interactive effects (International Multiple Sclerosis Genetics Consortium 2007). We also included the rs6897932 SNP in *IL7RA*. Stratification by this SNP was used to explore possible interactions as previously suggested (Gregory et al. 2007). The replication dataset had an unequally distributed sex ratio (Table VI); thus, to ensure that there was not a possible confounder, we stratified our dataset by gender to see if there were differences in the results.

We report nominal p-values (p-values <0.05) (unless otherwise indicated). A stringent Bonferroni corrected p-value is 2.0×10^{-5} calculated using $\sim 2,500$ independent SNPs (given the linkage disequilibrium in each gene region using an r^2 value=0.80). MDR was used to look for gene-gene interactions within the pathway (Ritchie et al. 2001). We have adequate power ($>80\%$) in the discovery dataset to see odds ratios of 1.3 and larger.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a grant from the NIH R01NS32830. JLM was partially supported by a National MS Society (NMSS) Research Award (RG 4201-A-1). The International MS Genetics Consortium is supported by R01NS049477. PLD is a Harry Weaver Neuroscience Scholar Award of the National MS Society (NMSS); he is also a William C. Fowler Scholar in Multiple Sclerosis Research and is supported by an NINDS K08 grant, NS46341. LP was a fellow of the National MS Society, USA (FG-1665-A-1) during these studies. DAH is a Jacob Javits Scholar of the NIH. This work was also supported by the Wellcome Trust (084702/Z/08/Z), the Medical Research Council (G0700061) and the Cambridge NIHR Biomedical Research Centre. We also acknowledge use of genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We acknowledge use of genotype data from the British 1958 Birth Cohort DNA collection, and thank the Accelerated Cure Project for its work in collecting samples from subjects with MS and for making these samples available to IMSGC investigators. We also thank the following clinicians for contributing to sample collection efforts: Accelerated Cure project – Drs. Elliot Frohman, Benjamin Greenberg, Peter Riskind, Saud Sadiq, Ben Thrower, and Tim Vollmer; Washington University – Drs. B.J. Parks and R.T. Naismith. Finally, we thank the Brigham & Women's Hospital PhenoGenetic Project for providing DNA samples from healthy subjects that were used in the Stage 2 follow-up effort of this study. We acknowledge the work done by the Center for Genome Technology Genotyping Core within the Miami Institute for Human Genomics (University of Miami); specifically Ashley Anderson and Luis Espinosa for aid in sample processing and genotyping. We also thank the Computational Genomics Core within the Center for Human Genetics Research (Vanderbilt University); specifically Justin Giles, Yuki Bradford, and David Sexton for their support in data processing.

References

1. Ansel KM, et al. Regulation of Th2 differentiation and Il4 locus accessibility. *Annu Rev Immunol* 2006;24:607–656. [PubMed: 16551261]
2. Australia and New Zealand Multiple Sclerosis Genetics Consortium. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosome 12 and 20. *Nat Genet* 41(7):824–828. [PubMed: 19525955]
3. Ban M, et al. Replication analysis identifies TYK2 as a multiple sclerosis susceptibility factor. *Eur J Hum Genet* 2009;17(10):1309–1313. [PubMed: 19293837]
4. Baranzini SE, et al. Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. *Hum Mol Genet* 2009;18(11):2078–2090. [PubMed: 19286671]
5. Baranzini SE, Wang J, Gibson RA. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet* 2009;18(4):767–778. [PubMed: 19010793]
6. Barrett JC, et al. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21(2):263–265. [PubMed: 15297300]
7. Bertrams J, Kuwert E. HL-A antigen frequencies in multiple sclerosis. Significant increase of HL-A3, HL-A10 and W5, and decrease of HL-A12. *European Journal of Neurology* 1972;7(74):78.
8. Compston A, Coles A. Multiple sclerosis. *Lancet* 2002;359(9313):1221–1231. [PubMed: 11955556]
9. De Jager PL, et al. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat Genet* 2009;41(7):776–782. [PubMed: 19525953]
10. Ebers GC, et al. A full genome search in multiple sclerosis. *Nat Genet* 1996;13(4):472–476. [PubMed: 8696345]
11. Ebers GC, Sadovnick AD, Risch NJ. A genetic basis for familial aggregation in multiple sclerosis. Canadian Collaborative Study Group. *Nature* 1995;377(6545):150–151. see comments. [PubMed: 7675080]
12. Fry TJ, Mackall CL. The many faces of IL-7: from lymphopoiesis to peripheral T cell maintenance. *J Immunol* 2005;174(11):6571–6576. [PubMed: 15905493]
13. Gabriel S, Ziaugra L, Tabbaa D. SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Curr Protoc Hum Genet* 2009;Chapter 2(Unit)

14. Goverman J. Autoimmune T cell responses in the central nervous system. *Nat Rev Immunol* 2009;9(6):393–407. [PubMed: 19444307]
15. Gregory SG, et al. Interleukin 7 receptor alpha chain (IL7R) shows allelic and functional association with multiple sclerosis. *Nat Genet* 2007;39(9):1083–1091. [PubMed: 17660817]
16. Haines JL, et al. A complete genomic screen for multiple sclerosis underscores a role for the major histocompatibility complex. The Multiple Sclerosis Genetics Group. *Nat Genet* 1996;13(4):469–471. [PubMed: 8696344]
17. Hauser MA, et al. Genomic convergence: identifying candidate genes for Parkinson's disease by combining serial analysis of gene expression and genetic linkage. *Hum Mol Genet* 2003;12(6):671–677. [PubMed: 12620972]
18. Hauser, SL.; Goodkin, DE. Multiple Sclerosis and Other Demyelinating Diseases. In: Fauci, AD., et al., editors. *Harrison's Principle of Internal Medicine*. McGraw Hill; New York: 1998. p. 2409-2419.
19. Hennah W, et al. Families with the risk allele of DISC1 reveal a link between schizophrenia and another component of the same molecular pathway, NDE1. *Hum Mol Genet* 2007;16(5):453–462. [PubMed: 17185386]
20. Hirschhorn JN, et al. A comprehensive review of genetic association studies. *Genet Med* 2002;4(2):45–61. [PubMed: 11882781]
21. Hofmeister R, et al. Interleukin-7: physiological roles and mechanisms of action. *Cytokine Growth Factor Rev* 1999;10(1):41–60. [PubMed: 10379911]
22. International Multiple Sclerosis Genetics Consortium. A high-density screen for linkage in multiple sclerosis. *Am J Hum Genet* 2005;77:454–467. [PubMed: 16080120]
23. International Multiple Sclerosis Genetics Consortium. Risk Alleles for Multiple Sclerosis Identified by a Genomewide Study. *N Engl J Med* 2007;357(9):851–862. [PubMed: 17660530]
24. International Multiple Sclerosis Genetics Consortium. The expanding genetic overlap between multiple sclerosis and type I diabetes. *Genes Immun* 2009a;10(1):11–14.
25. International Multiple Sclerosis Genetics Consortium. Comprehensive Follow-up of the First Genome-wide Association Study of Multiple Sclerosis identifies KIF21B and TMEM39A as Susceptibility Loci. *Hum Mol Genet*. 2009b Epub ahead of print.
26. Jiang Q, et al. Cell biology of IL-7, a key lymphotrophin. *Cytokine Growth Factor Rev* 2005;16(4-5):513–533. [PubMed: 15996891]
27. Jiang Q, et al. Distinct regions of the interleukin-7 receptor regulate different Bcl2 family members. *Mol Cell Biol* 2004;24(14):6501–6513. [PubMed: 15226449]
28. Kisseleva T, et al. Signaling through the JAK/STAT pathway, recent advances and future challenges. *Gene* 2002;285(1-2):1–24. [PubMed: 12039028]
29. Kumpfel T, et al. Multiple sclerosis and the TNFRSF1A R92Q mutation: clinical characteristics of 21 cases. *Neurology* 2008;71(22):1812–1820. [PubMed: 19029521]
30. Kuokkanen S, et al. Genomewide scan of multiple sclerosis Finnish multiplex families. *Am J Hum Genet* 1997;61:1379–1387. [PubMed: 9399895]
31. Lundmark F. Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. *Nat Genet* 2007;39(9):1108–1113. [PubMed: 17660816]
32. McDonald WI, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* 2001;50(1):121–127. [PubMed: 11456302]
33. McGovern DP, et al. Genetic epistasis of IL23/IL17 pathway genes in Crohn's disease. *Inflamm Bowel Dis* 2009;15(6):883–889. [PubMed: 19235914]
34. Motsinger AA, et al. Complex gene-gene interactions in multiple sclerosis: a multifactorial approach reveals associations with inflammatory genes. *Neurogenetics* 2007;8(1):11–20. [PubMed: 17024427]
35. Mumford GL, et al. The British Isles survey of multiple sclerosis in twins. *Neurology* 1994;44(11):15.
36. Naito S, et al. Multiple sclerosis: association with HL-A3. *Tissue Antigens* 1972;2(1):1–4. [PubMed: 5077731]

37. O'Shea JJ. Targeting the Jak/STAT pathway for immunosuppression. *Ann Rheum Dis* 2004;63:ii67–ii71. [PubMed: 15479876]
38. Ritchie MD, et al. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am J Hum Genet* 2001;69(1):138–147. [PubMed: 11404819]
39. Robertson NP, et al. Age-adjusted recurrence risks for relatives of patients with multiple sclerosis. *Brain* 1996;119(Pt 2):449–455. [PubMed: 8800940]
40. Sadovnick AD, et al. A population based study of multiple sclerosis in twins: update. *Annal of Neurology* 1993;33:281–285.
41. Sadovnick AD, Ebers GC. Genetics of multiple sclerosis. *Neurologic Clinics* 1995;13:99–118. [PubMed: 7739507]
42. Sadovnick AD, et al. Evidence for genetic basis of multiple sclerosis. *Lancet* 1996;347(1728):1730.
43. Sawcer S. The complex genetics of multiple sclerosis: pitfalls and prospects. *Brain* 2008;131(Pt 12):3118–3131. [PubMed: 18490360]
44. Sawcer S, et al. A genome screen in multiple sclerosis reveals susceptibility loci on chromosome 6p21 and 17q22. *Nature Genetics* 1996;13:464–468. [PubMed: 8696343]
45. Sexton DP, et al. The Whole-genome Association Study Pipeline (WASP): A Comprehensive Tool for Large-Scale Association Studies. *Am J Hum Genet* 2007;57:413S.
46. Shirai Y, Adachi N, Saito N. Protein kinase Cepsilon: function in neurons. *FEBS J* 2008;275(16):3988–3994. [PubMed: 18637121]
47. Steemers FJ, et al. Whole-genome genotyping with the single-base extension assay. *Nat Methods* 2006;3(1):31–33. [PubMed: 16369550]
48. Transatlantic Multiple Sclerosis Genetics Cooperative. A meta-analysis of genomic screens in multiple sclerosis. *Mult Scler* 2001;7(1):3–11. [PubMed: 11321191]
49. Wang WY, et al. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet* 2005;6(2):109–118. [PubMed: 15716907]
50. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447(7145):661–678. [PubMed: 17554300]

Table 1

Most Significant SNPs in the IL7R Pathway

Chr.	SNP	Location (bp)	Gene	Nearby Gene	Location	Armitage p-value	Odds Ratio *	Minor Allele	Minor Allele Frequency
8	rs2587156	79885663	IL7	IL7	Flanking 5'UTR	2.46E-05	0.63	A	0.07
2	rs2711285	46083353	PRKCE	PRKCE	intron	3.47E-04	1.23	G	0.37
19	rs368835	10286986	MGC19604	TYK2	intron	3.97E-04	0.77	A	0.17
18	rs8092360	58924818	BCL2	BCL2	Flanking 3'UTR	4.32E-04	1.26	A	0.22
16	rs1794019	11287276	PRM1	SOC1	Flanking 5'UTR	4.33E-04	0.78	C	0.18
16	rs181694	11292330	PRM1	SOC1	Flanking 5'UTR	4.35E-04	0.78	A	0.18
16	rs1646026	11290196	PRM1	SOC1	Flanking 5'UTR	4.86E-04	0.78	T	0.18
16	rs12922733	11288171	PRM1	SOC1	Flanking 5'UTR	5.38E-04	0.78	G	0.18
16	rs12599908	11303416	PRM1	SOC1	Flanking 5'UTR	5.47E-04	0.78	A	0.18
2	rs2595218	46079469	PRKCE	PRKCE	intron	5.65E-04	1.22	A	0.37
19	rs281423	10296493	RAVER1	TYK2	intron	5.72E-04	0.78	A	0.17
18	rs9947451	58894080	BCL2	BCL2	Flanking 3'UTR	6.12E-04	1.24	T	0.26
16	rs2867936	11306199	PRM1	SOC1	Flanking 5'UTR	6.42E-04	0.78	T	0.18
16	rs12596260	11305766	PRM1	SOC1	Flanking 5'UTR	7.31E-04	0.79	C	0.18
16	rs441349	11279260	PRM2	SOC1	Flanking 5'UTR	7.68E-04	0.79	T	0.18
16	rs497184	11297090	PRM1	SOC1	Flanking 5'UTR	9.40E-04	0.79	T	0.18
8	rs6993386	79816700	IL7	IL7	Flanking 5'UTR	9.58E-04	1.21	G	0.33

* OR calculated for minor allele

Table II

Armitage Trend Test Results for Replication Study (24 SNPs)

Chr.	SNP	Location (bp)	Gene	Nearby Gene	Location	Armitage pvalue	Odds Ratio	Minor Allele	Minor Allele Frequency
16	rs441349	11279260	PRM2	SOCS 1	Flanking 5UTR	6.90E-05	0.80	T	0.18
16	rs423674	11280906	PRM1	SOCS 1	Flanking 3UTR	2.35E-04	0.82	T	0.19
16	rs12599908	11303416	PRM1	SOCS 1	Flanking 5UTR	2.80E-04	0.82	A	0.19
16	rs2867936	11306199	PRM1	SOCS 1	Flanking 5UTR	2.86E-04	0.82	A	0.19
16	rs12922733	11288171	PRM1	SOCS 1	Flanking 5UTR	4.14E-04	0.82	G	0.19
16	rs1794019	11287276	PRM1	SOCS 1	Flanking 5UTR	4.81E-04	0.83	C	0.19
8	rs2587156	79885663	IL7	IL7	Flanking 5UTR	1.04E-02	0.81	T	0.07
16	rs2301365	11282787	PRM1	SOCS 1	Flanking 5UTR	2.20E-02	1.12	T	0.27
8	rs2953475	79800693	IL7	IL7	Flanking 3UTR	3.08E-02	1.11	A	0.28
8	rs1483573	79786172	C8orf70	IL7	intron	3.58E-02	1.11	A	0.28
8	rs1054283	79773265	C8orf70	IL7	coding	3.78E-02	1.10	A	0.28
8	rs894221	79811612	IL7	IL7	intron	4.75E-02	1.10	C	0.28
18	rs8088739	58918068	BCL2	BCL2	Flanking 3UTR	4.86E-02	1.10	C	0.26
18	rs8092360	58924818	BCL2	BCL2	Flanking 3UTR	6.70E-02	1.10	T	0.21
8	rs7816065	79925213	IL7	IL7	Flanking 5UTR	8.34E-02	1.08	T	0.46
8	rs6993386	79816700	IL7	IL7	Flanking 5UTR	8.73E-02	1.08	G	0.35
8	rs4739140	79915565	IL7	IL7	Flanking 5UTR	1.94E-01	0.94	C	0.38
8	rs1545298	79924391	IL7	IL7	Flanking 5UTR	2.22E-01	1.05	T	0.49
18	rs8083921	58917082	BCL2	BCL2	Flanking 3UTR	2.44E-01	0.94	T	0.27
8	rs2887502	79919146	IL7	IL7	Flanking 5UTR	2.48E-01	0.95	C	0.40
19	rs368835	10286986	MGC 19604	TYK2	intron	4.39E-01	0.95	A	0.14
18	rs9947451	58894080	BCL2	BCL2	Flanking 3UTR	5.56E-01	1.03	A	0.26
5	rs17624563	110484162	WDR 36	TSLP	coding	6.88E-01	1.03	G	0.09
19	rs2304256	10336652	TYK2	TYK2	coding	8.54E-01	1.01	A	0.28

Table III

Armitage Trend Test Results for the Combined Datasets (3454 cases and 3416 controls)

Chr.	SNP	Location (bp)	Gene	Nearby Gene	Location	Merged Armitagep-value	Odds Ratio
16	rs441349	11279260	PRM2	SOCS1	flanking 5UTR	3.48E-07	0.80
16	rs12599908	11303416	PRM1	SOCS1	flanking 5UTR	1.12E-06	0.81
16	rs2867936	11306199	PRM1	SOCS1	flanking 5UTR	1.26E-06	0.81
16	rs1794019	11287276	PRM1	SOCS1	flanking 5UTR	1.46E-06	0.81
16	rs12922733	11288171	PRM1	SOCS1	flanking 5UTR	1.49E-06	0.81
16	rs423674	11280906	PRM1	SOCS1	flanking 3UTR	2.34E-06	0.81
8	rs2587156	79885663	IL7	IL7	flanking 5UTR	8.29E-06	0.74
8	rs1483573	79786172	C8orf70	IL7	intron	3.63E-04	1.15
8	rs2953475	79800693	IL7	IL7	flanking 3UTR	3.75E-04	1.15
8	rs6993386	79816700	IL7	IL7	flanking 5UTR	5.07E-04	1.13
8	rs1054283	79773265	C8orf70	IL7	coding	6.92E-04	1.14
16	rs2301365	11282787	PRM1	SOCS1	flanking 5UTR	7.85E-04	1.14
8	rs894221	79811612	IL7	IL7	intron	8.06E-04	1.14
18	rs8092360	58924818	BCL2	BCL2	flanking 3UTR	1.17E-03	1.15
18	rs8088739	58918068	BCL2	BCL2	flanking 3UTR	2.84E-03	1.12
19	rs368835	10286986	MGC19604	TYK2	intron	2.90E-03	0.86
8	rs7816065	79925213	IL7	IL7	flanking 5UTR	4.02E-03	1.10
8	rs1545298	79924391	IL7	IL7	flanking 5UTR	5.44E-03	1.10
8	rs4739140	79915565	IL7	IL7	flanking 5UTR	5.65E-03	0.91
18	rs8083921	58917082	BCL2	BCL2	flanking 3UTR	8.81E-03	0.90
8	rs2887502	79919146	IL7	IL7	flanking 5UTR	1.32E-02	0.92
18	rs9947451	58894080	BCL2	BCL2	flanking 3UTR	3.08E-02	1.09
5	rs17624563	1.1E+08	WDR36	TSLP	coding	9.07E-02	0.91
19	rs2304256	10336652	TYK2	TYK2	coding	9.19E-02	0.94

Table IV

Replication dataset Armitage trend test results stratified by gender (24 SNPs)

Chr.	SNP	Location (bp)	Female		Overall		Male	
			Armitage p-value	Odds Ratio	Armitage p-value	Odds Ratio	Armitage p-value	Odds Ratio
16	rs441349	11279260	5.96E-03	0.82	6.90E-05	0.80	4.66E-03	0.76
16	rs12599908	11303416	2.37E-02	0.85	2.80E-04	0.82	5.48E-03	0.76
16	rs2867936	11306199	2.26E-02	0.85	2.86E-04	0.82	5.95E-03	0.77
16	rs1794019	11287276	2.53E-02	0.86	4.81E-04	0.83	6.52E-03	0.77
16	rs423674	11280906	1.54E-02	0.84	2.35E-04	0.82	5.80E-03	0.76
16	rs12922733	11288171	1.86E-02	0.85	4.14E-04	0.82	7.92E-03	0.77
8	rs2587156	79885663	4.64E-02	0.80	1.04E-02	0.81	3.20E-01	0.87
8	rs4739140	79915565	2.64E-01	0.94	1.94E-01	0.94	8.95E-01	0.99
8	rs2953475	79800693	1.19E-01	1.10	3.08E-02	1.11	1.37E-01	1.13
5	rs17624563	110484162	2.66E-01	1.12	6.88E-01	1.03	3.35E-01	0.88
8	rs1545298	79924391	7.87E-01	1.02	2.22E-01	1.05	1.19E-01	1.12
19	rs368835	10286986	2.19E-01	0.90	4.39E-01	0.95	5.86E-01	1.06
8	rs894221	79811612	1.77E-01	1.09	4.75E-02	1.10	1.48E-01	1.12
8	rs1483573	79786172	1.03E-01	1.11	3.58E-02	1.11	2.17E-01	1.10
8	rs6993386	79816700	1.27E-01	1.10	8.73E-02	1.08	4.20E-01	1.06
8	rs1054283	79773265	1.27E-01	1.10	3.78E-02	1.10	1.61E-01	1.12
8	rs2887502	79919146	4.59E-01	0.96	2.48E-01	0.95	5.25E-01	0.95
8	rs7816065	79925213	2.66E-01	1.07	8.34E-02	1.08	2.19E-01	1.09
16	rs2301365	11282787	1.33E-01	1.10	2.20E-02	1.12	4.31E-02	1.18
18	rs8088739	58918068	2.07E-01	1.09	4.86E-02	1.10	1.22E-01	1.14
18	rs8092360	58924818	2.22E-01	1.09	6.70E-02	1.10	1.89E-01	1.12
18	rs8083921	58917082	2.82E-01	0.93	2.44E-01	0.94	9.22E-01	1.01
19	rs2304256	10336652	4.64E-01	1.05	8.54E-01	1.01	6.09E-01	0.96
18	rs9947451	58894080	8.78E-01	1.01	5.56E-01	1.03	4.21E-01	1.07

Table V

Stratified HLA data for the combined datasets (HLA+ stratum are those individuals with at least 1 copy of the risk “A” allele at rs3135388 and HLA- stratum are those individuals without this risk allele)

Chr.	SNP	Location (bp)	HLA-Arrimtag p-value	HLA-Odds Ratio	Overall Arrimtag p-value	Overall Odds Ratio	HLA+Arrimtag p-value	HLA+ Odds Ratio
8	rs1483573	79786172	9.59E-07	1.27	3.63E-04	1.15	9.01E-01	1.01
8	rs1054283	79773265	1.73E-06	1.27	6.92E-04	1.14	9.56E-01	1.00
8	rs6993386	79816700	1.73E-06	1.25	5.07E-04	1.13	9.59E-01	1.00
8	rs2953475	79800693	1.74E-06	1.27	3.75E-04	1.15	7.55E-01	1.02
8	rs894221	79811612	1.86E-06	1.27	8.06E-04	1.14	8.58E-01	1.01
18	rs8092360	58924818	2.53E-04	1.22	1.17E-03	1.15	2.30E-01	1.09
16	rs441349	11279260	2.60E-04	0.80	3.48E-07	0.80	1.39E-05	0.73
16	rs2867936	11306199	3.32E-04	0.81	1.26E-06	0.81	5.58E-05	0.74
16	rs12922733	11288171	3.61E-04	0.81	1.49E-06	0.81	7.23E-05	0.75
16	rs1794019	11287276	3.70E-04	0.81	1.46E-06	0.81	6.43E-05	0.75
16	rs12599908	11303416	3.81E-04	0.81	1.12E-06	0.81	4.39E-05	0.74
16	rs423674	11280906	6.16E-04	0.82	2.34E-06	0.81	5.72E-05	0.75
18	rs8088739	58918068	7.71E-04	1.19	2.84E-03	1.12	1.47E-01	1.11
8	rs2587156	79885663	1.23E-03	0.75	8.29E-06	0.74	3.15E-03	0.72
18	rs8083921	58917082	4.05E-03	0.86	8.81E-03	0.90	2.25E-01	0.92
18	rs9947451	58894080	4.42E-03	1.15	0.031	1.09	9.01E-01	1.01
16	rs2301365	11282787	1.73E-02	1.13	7.85E-04	1.14	1.40E-02	1.19
8	rs4739140	79915565	3.14E-02	0.90	5.65E-03	0.91	1.21E-01	0.91
8	rs2887502	79919146	3.27E-02	0.91	0.013	0.92	2.98E-01	0.94
8	rs7816065	79925213	7.76E-02	1.08	4.02E-03	1.10	1.39E-02	1.16
8	rs1545298	79924391	8.38E-02	1.08	5.44E-03	1.10	1.93E-02	1.15
19	rs368835	10286986	2.02E-01	0.92	2.90E-03	0.86	9.94E-03	0.80
19	rs2304256	10336652	3.23E-01	0.95	0.092	0.94	1.05E-01	0.90
5	rs17624563	110484162	5.37E-01	0.95	0.091	0.91	4.92E-01	0.93

Table VI

Demographics of Case-control dataset from replication screen after Quality Control

	Case	Control
Sex (no)		
Female	1572	976
Male	539	1061
Ratio (female-male)	2.91	0.92
Age (yrs)		
Average	45.83	N/A
Range	17-79	N/A
Age at first symptoms (yrs)		
Average	19.05	N/A
Range	1-71	N/A
Disease course (% of known)		
Relapsing remitting	66.36	N/A
Secondary progressive	21.56	N/A
Primary progressive	7.02	N/A
Progressive relapsing	0.59	N/A
CIS	4.48	N/A

Genes Related to the IL7/IL7RA Pathway

Gene Symbol	Gene Name	Gene Symbol	Gene Name	Gene Symbol	Gene Name
PI3KCD	phosphoinositide-3-kinase delta	IL12B	interleukin-12 subunit beta	SOC57	suppressors of cytokine signaling 7
JAK1	Janus kinase 1	MAPK14	mitogen-activated protein kinase 14	STAT5B	signal transducer and activator of transcription 5B
IL-12RB2	interleukin-12 receptor, beta 2	SKAP2	src kinase associated phosphoprotein 2	STAT5A	signal transducer and activator of transcription 5A
MCL1	myeloid cell leukemia sequence 1	PI3KCG	phosphoinositide-3-kinase gamma	STAT3	signal transducer and activator of transcription 3
SHC1	Src homolog 2 domain containing transforming protein 1	PTK2B	protein tyrosine kinase 2 beta	TBET	Th-1-specific T box factor
PLA2G4A	phospholipase A2, group IVA	IL7	interleukin-7	PRKCA	protein kinase C subunit alpha
PRKCE	protein kinase C subunit epsilon	C-MYC	myelocytomatosis oncogene homolog	SOC53	suppressors of cytokine signaling 3
SOC55	suppressors of cytokine signaling 5	JAK2	Janus kinase 2	BCL2	B-cell CLL/lymphoma 2
IL1R2	interleukin-1-receptor type II	PTEN	phosphatase and tensin homolog	SOC56	suppressors of cytokine signaling 6
BIM	Bcl-2-like 11 (apoptosis facilitator) (Bcl2l11)	FRAT1	frequently rearranged in advanced T-cell lymphomas (GSK3)	NFATC1	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1
STAT1	signal transducer and activator of transcription 1	BAD	Bcl-2-associated agonist of cell death	TYK2	tyrosine kinase 2
IRS1	insulin receptor substrate 1	CPLA2	cytosolic phospholipase a2	IL27RA	interleukin-27-receptor subunit alpha
SHIP	inositol polyphosphate-5-phosphatase (INPP5D)	IL18	interleukin-18	JAK3	Janus kinase 3
CDC25A	cell division cycle 25 homolog A	LAG	lymphocyte-activation gene 3	IL12RB1	interleukin-12 receptor, beta 1
GCET2	germinal center expressed transcript 2	CDKN1B	Cyclin-dependent kinase inhibitor member 1B	BCL3	B-cell CLL/lymphoma 3
PI3KCB	phosphoinositide-3-kinase beta	STAT6	signal transducer and activator of transcription 6	PLA2G4C	phospholipase A2, group IVc
IL12A	interleukin-12 subunit alpha	IFNG	interferon gamma	BAX	Bcl-2-associated X protein
PI3KCA	phosphoinositide-3-kinase alpha	SOC52	suppressors of cytokine signaling 2	BCL2L1	Bcl-2-like 1
IL2	interleukin-2	IRS2	insulin receptor substrate 2	NFATC2	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2
IL15	interleukin-15	SOC54	suppressors of cytokine signaling 4	BID	BH3 interacting domain death agonist
IL7RA	interleukin-7 receptor subunit alpha	AKT	v-akt murine thymoma viral oncogene homolog 1	CRLF2	cytokine receptor-like factor 2
TSLP	thymic stromal lymphopoietin	PLA2G4B	phospholipase A2, group IVB	IL2RG	interleukin-2 receptor subunit gamma (gamma c)

Gene Symbol	Gene Name	Gene Symbol	Gene Name	Gene Symbol	Gene Name
IL5	interleukin-5	CSK	c-src tyrosine kinase	IL13RA1	interleukin-13-receptor alpha 1
IL13	interleukin-13	SOC1	suppressors of cytokine signaling 1		
IL4	interleukin-4	IL4R	interleukin-4 receptor		