Title

Deconstruction of HLA-DRB1*04:01:01 and HLA-DRB1*15:01:01 class II haplotypes using next-generation sequencing in European-Americans with multiple sclerosis

Journal

Multiple Sclerosis Journal, 25(6)

ISSN

1352-4585

Authors

Creary, Lisa E
Mallempati, Kalyan C
Gangavarapu, Sridevi
et al.

Publication Date

2019-05-01

DOI

10.1177/1352458518770019

Peer reviewed

Published in final edited form as:

Mult Scler. 2019 May ; 25(6): 772–782. doi:10.1177/1352458518770019.

Deconstruction of *HLA-DRB1*04:01:01* and *HLA-DRB1*15:01:01* class II haplotypes using next generation sequencing in European Americans with multiple sclerosis.

Lisa E. Creary[1], **Kalyan C. Mallempati**[2], **Sridevi Gangavarapu**[2], **Stacy J. Caillier**[3], **Jorge R. Oksenberg**[3], and **Marcelo A. Fernández-Viña**[1]

[1]:Department of Pathology, Stanford University School of Medicine, Palo Alto, CA, USA

[2]:Histocompatibility, Immunogenetics and Disease Profiling Laboratory, Stanford Blood, Center, Palo Alto CA, USA

[3]:Department of Neurology, University of California, San Francisco CA, USA

Abstract

Background—The association between *HLA-DRB1*15:01* with multiple sclerosis (MS) susceptibility is well established but the contribution of the tightly associated *HLA-DRB5*01:01* allele has not yet been completely ascertained. Similarly, the effects of *HLA-DRB1*04:01* alleles and haplotypes, defined at the full gene resolution level, with MS risk remains to be elucidated.

Objectives—To characterize the molecular architecture of class II *HLA-DR15* and *HLA-DR4* haplotypes associated with MS.

Methods—Next-Generation Sequencing was used to determine *HLA-DQB1*, *HLA-DQA1*, and *HLA-DRB1/4/5* alleles in 1403 unrelated European-American patients and 1425 healthy unrelated controls. Effect sizes of HLA alleles and haplotypes on MS risk were measured by odds ratio with 95% confidence intervals.

Results—*HLA-DRB1*15:01:01:01SG* (OR=3.20, *P*<2.2E-16), *HLA-DRB5*01:01:01* (OR=2.96, *P*<2.2E-16) and *HLA-DRB5*01:01:01v1_STR1* (OR=8.18, *P*=4.3E-05) alleles all occurred at significantly higher frequencies in MS-patients compared to controls. The most significant predisposing haplotypes were *HLA-DQB1*06:02:01~HLA-DQA1*01:02:01:01SG~HLA-DRB1*15:01:01:01SG~HLA-DRB5*01:01:01* and *HLA-DQB1*06:02:01~HLA-DQA1*01:02:01:01SG~HLA-DRB1*15:01:01:01SG~HLA-DRB5*01:01:01v1_STR1* (OR=3.19, *P*<2.2E-16; OR=9.30, *P*=9.7E-05 respectively). Analyses of the *HLA-DRB1*04* cohort in the absence of *HLA-DRB1*15:01* haplotypes revealed that the *HLA-DQB1*03:01:01:01~HLA-DQA1*03:03:01:01~HLA-DRB1*04:01:01:01SG~HLA-DRB4*01:03:01:01* haplotype was protective (OR=0.64, *P*=0.028) whereas the *HLA-DQB1*03:02:01~HLA-DQA1*03:01:01~HLA-DRB1*04:01:01:01SG~HLA-DRB4*01:03:01:01* haplotype was associated with MS susceptibility (OR=1.66, *P*=4.9E-03).

Contact: Lisa E. Creary and Marcelo A. Fernández-Viña, Department of Pathology, Stanford University School of Medicine, 3155 Porter Drive, Palo Alto, CA 94304, USA, Tel: +1 650 498 1248, +1 650 723 7968, lcreary@stanford.edu, marcelof@stanford.edu.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Conclusions—*HLA-DR15* haplotypes, including genomic variants of *HLA-DRB5*, and *HLA-DR4* haplotypes affect MS risk.

Keywords

Human leukocyte antigen DRB1; Haplotype; European Americans; Next-generation sequencing; Multiple Sclerosis; Susceptibility; Protection

Introduction

Genome-wide association studies have identified over 200 independent genetic associations with multiple sclerosis (MS) susceptibility (1,2) As expected, all variants are relatively common in the population and have modest individual statistical effects on risk. The strongest genome-wide susceptibility locus maps to the major histocompatibility complex (MHC, 6p21.3). This association, which was first described several decades ago (3,4) is consistent with the long-held view that MS is fundamentally an antigen-specific autoimmune disease, but also reflects the complexity of risk inheritance by harboring multiple statistically independent allelic and haplotypic effects (5,6), including protective signals in the class I region (7) and specific class II genes allelic interactions (2,8) Recently, the largest meta-analysis of GWAS-derived data from 47,351 MS subjects and 68,284 controls identified 32 independent associations in the extended MHC including non-classical HLA effects (2). These new susceptibility variants provide powerful information for understanding the architecture of genetic susceptibility in the MHC region, including gene x gene interactions, and may be further investigated to examine their impact on gene function.

The specific association of MS with the class II *HLA-DRB1*15:01* allele has been the strongest genetic risk factor and consistently reported for nearly all studies (9). The most prudent explanation for this observation appears to be related to singularities in the molecular structure of the HLA heterodimer pockets, namely the large predominantly hydrophobic P4 pocket of the peptide-binding domain of *HLA-DRß1*15:01*, determining peptide ligands specificity and composition of the responding T-cell receptor repertoire (10). For the putative MS autoantigen myelin basic protein (MBP), this P4 pocket may be occupied by the aromatic side chain Phe92, responsible for its high-affinity binding to the HLA-DRα01:01/HLA-DRß15:01 heterodimer. The presence of alanine at the polymorphic HLA-DRβ71 position is critical in creating the necessary space for Phe92 of MBP, and among the common *HLA-DRB1* alleles the uncharged Ala at this position is only observed in the *HLA-DRB1*15* allele group (*HLA-DRB1*15:01–HLA-DRB1*15:06*).

*HLA-DRB1*15* allelic variants are in strong linkage disequilibrium (LD) with a second functional DRB locus of restricted polymorphism, *HLA-DRB5*. Remarkably, the crystal structure of a HLA-DRα/HLA-DRß5*01:01-Epstein-Barr virus peptide complex revealed a striking structural similarity to the HLA-DRß1*15:01-MBP peptide complex at the surface presented for T-cell receptor recognition (11), suggesting that peptides with only limited sequence identity with a myelin peptide could activate autoreactive T-cells. Not surprisingly, HLA-DRαHLA-DRß5 heterodimers appear to be effective MBP antigen presenting molecules (12,13). In addition, *HLA-DRB5*01:01* humanized transgenic mice crossed with

mice expressing T-cell receptors specific for MBP 83–99 showed spontaneous experimental autoimmune encephalitis (EAE) (14). Of great interest is the report of *HLA-DRB1~HLA-DRB5* epistatic functional interactions elegantly shown in an HLA humanized EAE model, whereby *HLA-DRB1* alleles underlie susceptibility and severity is mediated by *HLA-DRB5* (15). The prediction derived from the murine model that *HLA-DRB5* acted as a modifier, was confirmed in *HLA-DRB5*-null African-American MS patients, showing a faster progression to the progressive phase of the disease (16). However, the full spectrum of *HLA-DRB5* coding and non-coding variation in the context of MS susceptibility is unknown.

*HLA-DRB1*04* has been reported to be associated with MS disease susceptibility (17,18) but this correlation appears to be limited by the ancestry of the population. In contrast to *HLA-DRB1*15:01* that is associated with MS risk regardless of the clinical phenotype, carriage of *HLA-DRB1*04* alleles has also been shown to be associated with the development of primary progressive disease MS (PP-MS) course (19). However, this observation was not found consistently and several other studies could only either suggest a non-significant trend to a positive association of *HLA-DRB1*04* alleles with PP-MS (20) or no effect at all (21). On the other hand, a few studies have shown a protective effect of *HLA-DRB1*04* with MS (22,23). The major limitation of previous studies, examining *HLA-DRB1*04* with MS, is that most *HLA-DRB1*04* genotypes are ambiguous due to serological or molecular low ‘first-field’ resolution typing. The DR4 serotype represents a large heterogeneous group of more than 300 *HLA-DRB1*04* alleles, many of which have substantially different amino acid sequences forming the antigen-binding site of the HLA molecules. Definitive *HLA-DRB1*04* alleles were not assigned in many of the previous studies and therefore it is very likely that different *HLA-DRB1*04* alleles, as well as the disease heterogeneity of the patients studied, may have been responsible for the opposing effects on MS risk and outcome. In addition *HLA-DRB1*04* alleles have various predisposing effects in Rheumatoid Arthritis (RA) suggesting balance in autoimmune effects of ancestral haplotypes; this concept should also be considered in the study of MS.

The application of next-generation sequencing (NGS) to the study of highly polymorphic and structurally complex regions of the human genome increases the throughput, accuracy, and resolution of genetic analysis by several orders of magnitude, presenting an opportunity to better understand the biological mechanisms underlying HLA disease associations. Many studies of HLA disease association have imputed HLA alleles from SNP typing (1,2,8). These approaches are useful for large-scale association studies but present some limitations. For instance HLA imputation methods usually only generate two-field resolution HLA types because the reference dataset does not include alleles differing at four fields. Association testing is restricted only to variations at the peptide-binding region of the HLA molecule omitting examination of non-coding variants that may influence expression. In addition, HLA imputation methods are unable to identify novel variants. To circumvent such limitations, we employed a NGS method developed to type complete and/or extended regions of HLA defined at three and four allele resolution, pertaining to molecular variations in coding and noncoding regions. With our NGS approach new HLA alleles at any resolution can be readily detected and reveal new distinctive haplotypic associations. In this study, we used NGS to genotype HLA class II loci in a cohort of European-American MS

patients and ethnically matched unrelated controls to assess the role of *HLA-DRB1*04:01* and *HLA-DRB1*15:01* bearing haplotypes on MS risk.

Materials and Methods

Study Population

The dataset consisted of 2828 de-identified DNA samples from 1403 MS patients (1016 females, 72.4%) and 1425 healthy unrelated controls (791 females, 55.5%). All MS subjects met established diagnostic criteria (24) and are non-Hispanic white. Control subjects were also white of European ancestry and reported no self-history of personal chronic diseases or in their nuclear family. This study was approved by the University of California, San Francisco Institutional Review Board.

HLA genotyping

DNA samples were retrospectively typed for HLA class II loci (*HLA-DQB1*, *HLA-DQA1*, *HLA-DRB1*, *HLA-DRB4*, and *HLA-DRB5*) using the MIA FORA NGS high-throughput semi-automated typing protocol (Immucor, Inc., Norcross, GA) and performed following the manufacturer's instructions. The coverage for class II loci amplified by long range PCR are: *HLA-DQB1* 5'UTR to 3'UTR; *HLA-DQA1* 5'UTR to 3'UTR; *HLA-DRB1/4/5* gene fragments were amplified in two separate reactions, 5' UTR to the first ~270 bp of intron-1, and ~250 bp at the 3' end of intron-1 to exon-6. Sample libraries were prepared and sequenced at a final concentration of 12 pM spiked with 5% PhiX on the Illumina HiSeq-2500 or 1.3 pM with 2% PhiX on NextSeq-500 instruments using 150 cycle paired-end kits (Illumina, Inc., San Diego, CA).

HLA sequence data analysis and genotype assignment

NGS sequence data stored as FASTQ files were uploaded into the MIA FORA FLEX v3.0 alignment software (Immucor, Norcross, GA) for analyses and assignment of HLA genotypes. The MIA FORA software demultiplexes FASTQ files according to each unique index and uses two complementary informatic strategies; competitive mapping of paired-end sequence reads and *de novo* assembly of paired-end reads to construct one or two phased consensus sequences. Paired-end reads and consensus sequences are compared with three sources of HLA reference sequences; (i) the IPD-IMGT/HLA database v3.25.0 (https://www.ebi.ac.uk/ipd/imgt/hla/), (ii) internal MIA FORA HLA references generated by cloning with sequencing and (iii) computational filled *in silico* HLA sequences, to assign genotypes. Final correct HLA genotype calls were assigned after manual review.

STR ambiguous groups

The MIA FORA software permits detailed examination of all sequence segments, as well as unambiguous allele assignment, with the exception of short tandem repeat (STR) enriched regions located within introns of some class II genes. These STR regions consists of 'A' or 'T' mononucleotides and/or 'GT' and 'GA' dinucleotides that are repeated typically ~10 to ≥20 times and cannot be assessed accurately by the sequencing methodology. Allele groups that are indistinguishable due to STRs are given the suffix SG (STR group) to the lowest numbered allele in that group. Characteristics of STRs for each SG and ambiguities that

occur due to unsequenced regions within the HLA genes are described in supplementary Table 1.

DRB5 reference sequences

There are two *HLA-DRB5*01:01:01* genomic reference sequences included in the MIA FORA HLA software; *HLA-DRB5*01:01:01* (GenBank accession number AL713966), and an intronic variant of *HLA*-DRB5*01:01:01, but lacking intron-1 sequence, denoted as *HLA-DRB5*01:01:01v1_STR1* (#KU593576) that was generated by cloning and sequencing experiments (Barsakis *et. al.* manuscript in preparation). Genomic sequences for both alleles are located in the European Nucleotide Archive repository (https://www.ebi.ac.uk/ena).

Statistical analyses

Allele carrier frequencies were determined by direct counting and were calculated as the number of individuals carrying a specific allele (either at the homozygous or heterozygous state) divided by the total number of individuals. In control subjects allele frequencies at all loci were tested for deviations from Hardy-Weinberg equilibrium (HWE) proportions using the exact method of Guo and Thompson implemented in the PyPop software (25).

The BIGDAWG software (26) was used to estimate six locus haplotypes. The effect sizes of HLA alleles,*HLA-DRB1*04:01* and *HLA-DRB1*15:01* bearing haplotypes on MS risk were measured by odds ratio (OR) with 95% confidence intervals (CI), and associated probability (*P)* values derived from a two-tailed Fischer's exact test. A *P* value of 0.05 (α) or less was considered statistically significant. Statistical analyses were performed using the *R* statistical program v3.3.2. Power calculations were performed using the PS power-calculator program (27) (Supplementary Table 2).

Results

Allele level association analyses

We tested in this population *HLA-DRB1*04*, *HLA-DRB1*15*, and *HLA-DRB5*01* for allelic heterogeneity and susceptibility to MS risk. Analyses of HLA NGS data from MS patients and controls revealed no new deleterious mutations that could affect expression in any of the alleles detected. The levels of discrimination by the NGS method and testing multiple loci, were able to estimate accurately the OR for susceptibility and resistance of alleles and haplotypes compared to low-resolution typing methods PCR sequence-specific oligonucleotide probe (PCR-SSOP) and sequence specific primer (SSP) or imputations of HLA alleles. The distribution of allele frequencies at all loci in controls did not deviate significantly from HWE. Allele frequencies were compared in the MS cases and controls (Table 1). As expected there was a significant increased frequency of *HLA-DRB1*15:01:01:01SG* in the MS group compared to controls and was highly associated with increased susceptibility to MS (50.5% cases vs 24.1% controls, OR=3.20, CI=2.63–3.91, *P*<2.2E-16,) whereas the alternative intronic variant *HLA-DRB1*15:01:01:04* allele, described initially in Asians, was not detected in either the MS group or control group.

The two intronic variants of *HLA-DRB5*01:01* also occurred at higher frequencies in MS cases than controls although *HLA-DRB5*01:01:01* was more frequent: 48.2% vs 23.9 %, OR=2.96, CI=2.43–3.61, *P*<2.2E-16; *HLA-DRB5*01:01:01v1_STR1* 2.3% vs 0.3%, OR=8.18, CI=2.68–24.95, *P*=4.3E-05. In contrast, the frequency of *HLA-DRB1*04:01:01:01SG* was significantly decreased in the MS group compared to controls and was protective even in the presence of the highly predisposing *HLA-DRB1*15:01:01:01SG* alleles: 6.6% vs 16.6%, OR=0.36, CI=0.28–0.46, *P*<2.2E-16.

HLA-DQB1~DQA1~DRB1~DRB4/5 haplotype level association analyses: *HLA-DRB1*15:01* haplotypes

The distribution of *HLA-DQB1~DQA1~DRB1~DRB4/5* haplotype blocks in MS cases and controls were compared. We first assessed which *HLA-DRB1*15:01* bearing haplotypes were associated with susceptibility to MS (Table 2). Of the overall 17 *HLA-DRB1*15:01* bearing haplotypes observed in the dataset, 11 different haplotypes were found in both MS patients and control subjects, whereas 2 haplotypes, which are *HLA-DRB5*01:01:01* variants of the 'classic' MS-positively associated haplotype, were found at significant increased frequencies in the MS group compared to controls: the common haplotype *HLA-DQB1*06:02:01~HLA-DQA1*01:02:01:01SG~HLA-DRB1*15:01:01:01SG~HLA-DRB5*01:01:01* (30.1% vs 11.9%, OR=3.19, CI=2.70–3.77, *P*<2.2E-16) and the rare haplotype *HLA-DQB1*06:02:01~HLA-DQA1*01:02:01:01SG~HLA-DRB1*15:01:01:01SG~HLA-DRB5*01:01:01v1_STR1* (1.0% vs 0.11%, OR=9.30, CI=2.62–33.02, *P*=9.7E-05), consistent with the involvement of *HLA-DRB1*15:01* in disease susceptibility. In addition, the *HLA-DQB1*06:03:01~HLA-DQA1*01:02:01:01SG~HLA-DRB1*15:01:01:01SG~HLA-DRB5*01:01:01* haplotype was also more frequent in MS cases than controls: 0.8% vs 0.3%, OR=2.90, CI=1.14–7.36, *P*=0.0353. The frequencies of the remaining fourteen haplotypes in both cases and controls were too low to deduce any plausible association results. Figure 1 depicts the *HLA-DRB1*15:01:01:01SG* bearing haplotypes positively associated with MS risk.

*HLA-DRB1*04:01:01:01SG* haplotypes in *HLA-DRB1*15:01* - positive patients and controls

As shown above, *HLA-DRB1*04:01:01:01SG* appears to be associated with disease protection. In the presence of *HLA-DRB1*15:01:01* the distribution of *HLA-DRB1*04:01:01:01SG* haplotypes blocks were slightly more homogenous in MS cases compared to controls, 5 vs 6 haplotypes respectively (Table 3). Two haplotypes bearing *HLA-DRB1*04:01:01:01SG* were significantly more frequent in the control group compared to the MS patients: *HLA-DQB1*03:01:01:01~HLA-DQA1*03:03:01:01~HLA-DRB1*04:01:01:01SG~HLA-DRB4*01:03:01:01* (1.1% vs 4.8%, OR=0.21, CI=0.14–0.32, *P*<2.2E-16) and *HLA-DQB1*03:02:01~HLA-DQA1*03:01:01~HLA-DRB1*04:01:01:01SG~HLA-DRB4*01:03:01:01* (1.9% vs 3.4%, OR=0.54, CI=0.35–0.76, *P*=4.3E-04).

A third less common haplotype bearing *HLA-DRB1*04:01:01:01SG* was also over represented in controls but the association was not statistically significant: *HLA-DQB1*03:02:01~HLA-DQA1*03:03:01:01~HLA-DRB1*04:01:01:01SG~HLA-DRB4*01:03:01:01* (0.4% vs 0.6%, OR=0.56, CI=0.26–1.22, *P*=0.11).

*HLA-DRB1*04:01:01:01SG* haplotypes after removal of *HLA-DRB1*15:01* haplotypes

Due to the dominant effect of *HLA-DRB1*15:01* with MS susceptibility we performed further analyses to examine the associations of *HLA-DRB1*04:01:01:01SG* haplotypes after exclusion of *HLA-DRB1*15:01* positive haplotypes in both MS cases and control groups (Table 4). Interestingly the *HLA-DQB1*03:02:01~HLA-DQA1*03:01:01~HLA-DRB1*04:01:01:01SG~HLA-DRB4*01:03:01:01* haplotype occurs at higher frequency in the MS group compared to controls and is significantly associated with MS susceptibility: 6.2% vs 3.9%, OR=1.66, CI=1.17–2.35, *P*=4.9E-03. Whereas the *HLA-DQB1*03:01:01:01~HLA-DQA1*03:03:01:01~HLA-DRB1*04:01:01:01SG~HLA-DRB4*01:03:01:01* haplotype frequencies remained elevated in controls and was moderately protective: 3.6% vs 5.6%, OR=0.64, CI=0.43–0.95, *P*=0.028.

The *HLA-DQB1*03:02:01~HLA-DQA1*03:03:01:01~HLA-DRB1*04:01:01:01SG~HLA-DRB4*01:03:01:01* haplotype frequencies were slightly higher in cases than controls but the difference was not statistically significant: 1.2% vs 0.7%, OR=1.67, CI=0.77–3.63, *P*=0.19. The *HLA-DRB1*04:01:01:01SG* haplotypes significantly associated with MS susceptibility and protection are illustrated in Figure 2.

Discussion

In this study we performed high-resolution NGS typing of the HLA class II genes, *HLA-DQB1*, *HLA-DQA1*, *HLA-DRB1*, *HLA-DRB4*, and *HLA-DRB5* to refine our current understanding of the effect of HLA alleles and haplotypes on MS susceptibility and protection in the European-American population. We exploited the powerful combination of long-range PCR (amplification of the entire locus or wide coverage of key regions of the gene) with NGS to generate accurate non-ambiguous 6 to 8 digit HLA genotypes. These ‘ultra’ high-resolution genotypes represent a more detailed description of HLA allelic variants, differing in coding and non-coding gene segments, than 2 or 4 digit typing. Furthermore, differences in intronic polymorphisms is useful for breaking down MHC LD patterns, which may be pertinent for locating the casual variant(s) in disease association studies.

Specifically, we focused our investigation on *HLA-DRB5*01~HLA-DRB1*15:01~HLA-DQA1~HLA-DQB1* and *HLA-DRB4*01~HLA-DRB1*04:01~HLA-DQA1~HLA-DQB1* alleles and haplotypes. We have confirmed the numerous previous findings of a strong association of *HLA-DRB1*15:01* with MS susceptibility especially in white populations such as European-Americans (28,29), and groups from Europe (30), and Australia (23,31). In addition, the allele frequency of *HLA-DRB1*15:01:01:01SG* in MS cases (50.5%) and the strength of association (OR=3.20) found in our study is similar to those reported in other white datasets (29,30,31). The *HLA-DRB1*15:01* MS association have also been found in non-European populations such as African-Americans (16), Latin-Americans (32), and Chinese cohorts (33) but the strength of association was weaker than reported in European-ancestry populations also other class II alleles were found to be of greater effect than *HLA-DRB1*15:01*.

In all MS cases and controls that carried a *HLA-DRB1*15* allele, *HLA-DRB1*15:01:01:01SG* was exclusively associated with a functional *HLA-DRB5* gene, therefore, unsurprisingly analyses of *HLA-DRB5*01:01:01* alleles also revealed a strong association with MS risk. Analyses of the DR15 haplotypes showed that two different 6–8 digit defined DR15 haplotypes were positively associated with MS: *HLA-DRB5*01:01:01* and *HLA-DRB5*01:01:01v1_STR1* bearing haplotypes OR=3.19 and 9.30 respectively, effect sizes that are very similar to those of *HLA-DRB1*15:01:01:01SG*, and *HLA-DRB5*01:01* alleles analyzed independently. Due to the intense LD between *HLA-DRB1*15:01:01* and *HLA-DRB5*01:01:01* alleles, as well as the similarity of the effect sizes of alleles analyzed individually or by haplotypes, it was not possible to discern the individual contributions of these alleles on the DR15 haplotypes with MS susceptibility. Although the strong positive association finding of the DR15 haplotype harboring the *HLA-DRB5*01:01:01_v1_STR* allele is of interest, the result should be viewed with caution due to the very low frequencies of this haplotype observed in both MS cases (0.97%) and controls (0.11%), and the relatively high OR of 9.30 is most likely an artifact due to low statistical power. A larger dataset would need to be examined to confirm the contribution, if any, of *HLA-DRB5*01:01:01v1_STR1* to MS. In a previous study of African-Americans, the haplotypes bearing *HLA-DRB1*15* in which *HLA-DRB5* is deleted were associated with increased risk for developing SPMS, suggesting that DRB5 attenuates MS severity (16). Our findings in European-American MS patients and controls do not show the presence of sequence variants within the gene that could affect gene expression of *HLA-DRB5* and therefore could either mitigate or increase disease susceptibility or severity.

Association results of *HLA-DRB1*04* alleles and haplotypes revealed a more complex picture about MS compared to our *HLA-DRB1*15* findings. We found that *HLA-DRB1*04:01:01:01SG* alleles were strongly negatively associated with MS (OR=0.36). At the haplotypic level, due to genetic diversity at the DQ loci, different haplotypes bearing *HLA-DRB1*04:01:01:01SG* have opposite effects on MS risk dependent and independent of *HLA-DRB1*15:01:01*. Two distinct *HLA-DRB1*04:01:01:01SG* haplotypes, bearing either *HLA-DQB1*03:01:01:01~HLA-DQA1*03:03:01:01* or *HLA-DQB1*03:02:01~HLA-DQA1*03:01:01* were highly protective when the overall cohort was analyzed. Further analyses of the haplotypes bearing *HLA-DRB1*04:01:01:01SG* controlling for potential confounding by *HLA-DRB1*15* showed that the *HLA-DQB1*03:01:01:01~HLA-DQA1*03:03:01:01~HLA-DRB1*04:01:01:01SG~HLA-DRB4*01:03:01:01* haplotype remained protective (OR=0.64, P=0.028) whereas the *HLA-DQB1*03:02:01~HLA-DQA1*03:01:01~HLA-DRB1*04:01:01:01SG~HLA-DRB4*01:03:01:01* haplotype exhibited a moderately predisposing effect (OR=1.66, P=0.0049). These findings clearly indicate that the *HLA-DRB1*04:01:01:01SG* allele alone does not affect MS risk but is part of a protective or susceptibility haplotype. To our knowledge, we are the first to report MS associations with *HLA-DRB1*04:01:01:01SG* typed at the 4-field allele level. Our findings that two different DR4 haplotypes associate with MS in opposite directions is not unexpected as it well known, since the 1980's, that two common haplotypes bearing *HLA-DRB1*04:01* (Dw4) associate with different DQB1 alleles (34,35), and these two haplotypes show distinct effects in susceptibility to two diseases. In insulin-dependent diabetes mellitus only *HLA-DRB1*04:01~HLA-DQB1*03:02* associates with susceptibility (34) while in RA

both *HLA-DRB1*04:01~HLA-DQB1*03:02* and *HLA-DRB1*04:01~HLA-DQB1*03:01* haplotypes associate with predisposition to the disease (35). In some diseases such as RA where specific motifs that determine function appear to have causative effects the application of NGS extended testing may allow for fine mapping of these factors. In the present study because of tight LD we were unable to map these structural features to only one allele of a given locus lending the contribution of different factors in several steps of the disease causing mechanism, such as peptide presentation and elimination of T-cell clones in the thymus.

HLA class I and II alleles exist on extended haplotypes encompassing multiple genes and specific alleles in tight LD, therefore it is possible that different loci and alleles act in synergy to confer susceptibility or protection to MS risk. Compared to the number of class II loci implicated in MS, a few class I alleles have been reported to be associated with susceptibility to MS (*HLA-A*03*, *HLA-B*07*), independent of class II molecules, and the class I protective effect has been reported to be mainly driven by *HLA-A*02:01* (30) and *HLA-B*44:02* (Bw4–80T) (29). One of the protective *HLA-DRB1*04:01:01:01SG* extended haplotypes found in this study includes the *HLA-B*44:02* allele. The mapping of protection may be elucidated when both class I and II HLA loci are examined simultaneously by ultra-high resolution typing. The NGS approach taken in the present study can be applied to examine systematically the effect of all classical class I and class II HLA alleles in diseases in which associations with HLA factors have been described.

In conclusion, our results support and extend previous findings of an association between *HLA-DRB1*15:01* and *HLA-DRB1*04:01* alleles with MS and demonstrate the advantages of high-resolution NGS to determine risk and protective haplotypes. Larger multi-ancestral studies using NGS combined with genotyping of a dense set of SNP markers are planned to further elucidate the contribution of HLA to MS susceptibility.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We are grateful to the subjects with MS and controls who have participated in this study.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the U.S. National Institutes of Health grants R01NS026799 and U19AI119350.

References

1. Beecham AH, Patsopoulos NA, Xiafra DK, Davis MF, Kemppinen A, Cotsapas C, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nature Genetics. 2013; 45: p. 1353–60. [PubMed: 24076602]
2. Patsopoulos N, Baranzini SE, Santaniello A, Shoostari P, Cotsapas C, Wong G, et al. bioRxiv: The Multiple Sclerosis Genomic map:Role of peripheral immune cells and resident microglia in susceptibility. [Online].; 2017 [cited 2017 October 30 10.1101/143933.

3. Bertrams J, Kuwert E, Liedtke U. HL-A antigens and multiple sclerosis. Tissue Antigens. 1972; 2: p. 405–408. [PubMed: 4655776]
4. Naito S, Naerow N, Mickey MR, Terasaki PI. Multiple sclerosis association with HLA-A3. Tissue Antigens. 1972; 2: p. 1–4. [PubMed: 5077731]
5. Dyment DA, Herrera BM, Cader MZ, Willer CJ, Lincoln MR, D SA, et al. Complex interactions among MHC haplotypes in multiple sclerosis: susceptibility and resistance. Human Molecular Genetics. 2005; 14: p. 2019–26. [PubMed: 15930013]
6. Patsopoulos NA, Barcellos LF, Hintzen RQ, Schaefer C, van Duijin CM, Noble JA, et al. Fine-mapping the genetic association of the major histocompatibility complex in multiple sclerosis: HLA and non-HLA effects. PLoS Genet. 2013; 9: p. e1003926. [PubMed: 24278027]
7. Rioux JD, Goyette P, Vyse TJ, Hammarstrom L, Fernando MM, Green T, et al. Mapping of multiple susceptibility variants within the MHC region for 7 immune-mediated diseases. Proc Natl Acad Sci USA. 2009; 106: p. 18680–85. [PubMed: 19846760]
8. Moutsianas L, Jostins L, Beecham AH, Dilthey AT, Xiafara DK, Ban M, et al. Class II HLA interactions modulate genetic risk for multiple sclerosis. Nat. Genet 2015; 47(10): p. 1107–13. [PubMed: 26343388]
9. Hollenbach JA, Oksenberg JR. The immunogenetics of multiple sclerosis: A comprehensive review. J. Autoimmun. 2015; 64: p. 13–25. [PubMed: 26142251]
10. Smith KJ, Pyrdol J, Gauthier L, Wiley DC, Wucherpfenning KW. Crystal structure of HLA-DR2 (DRA*0101, DRB2*1501) complexed with a peptide form human myelin basic protein. J. Exp. Med. 1998; 188: p. 1511–20. [PubMed: 9782128]
11. Lang HL, Jacobsen H, Ikemizu S, Andersson C, Harlos K, Madsen L, et al. A functional and structural basis for TCR cross-reactivity in multiple sclerosis. Nat. Immunol. 2002; 3: p. 940–43. [PubMed: 12244309]
12. Prat E, Tomaru U, Sabater L, Park DM, Granger R, Kruse N, et al. HLA-DRB5*0101 and -DRB1*1501 expression in the multiple sclerosis-associated HLA-DR15 haplotype. J Neuroimmunol.. 2005; 167: p. 108–19. [PubMed: 16111772]
13. Sospedra M, Muraro PA, Stefanova I, Zhao Y, Chung K, Li Y, et al. Redundancy in antigen-antigen-presenting function of the HLA-DR and -DQ molecules in the multiple sclerosis-associated HLA-DR2 haplotype. J. Immunol. 2006; 176: p. 1951–61. [PubMed: 16424227]
14. Quandt JA, Huh J, Baig M, Yao K, Ito N, Bryant M, et al. Myelin basic protein-specific TCR/HLA-DRB5*01:01 transgenic mice support the etiologic role of DRB5*01:01 in multiple sclerosis. J. Immunol 2012; 189: p. 2897–908. [PubMed: 22888134]
15. Gregersen JW, Kranc KR, Ke X, Svendsen P, Medsen LS, Thomsen AR, et al. Functional epistasis on a common MHC haplotype associated with multiple sclerosis. Nature. 2006; 443: p. 574–77. [PubMed: 17006452]
16. Caillier SJ, Briggs F, Cree BA, Baranzini SE, Fernandez-Viña M, Ramsay PP, et al. Uncoupling the roles of HLA-DRB1 and HLA-DRB5 genes in multiple sclerosis. J. Immunol 2008; 181: p. 5473–80. [PubMed: 18832704]
17. Brassat D, Salemi G, Barcellos LF, McNeill G, Proia P, Hauser SL, et al. The HLA locus and multiple sclerosis in Sicily. Neurology. 2005; 64: p. 361–3. [PubMed: 15668443]
18. Saruhan-Direskeneli G, Esin S, Baykan-Kurt B, Ornek I, Vaughan R, Eraksoy M. HLA-DR and –DQ associations with multiple sclerosis in Turkey. Hum. Immunol 1997; 55: p. 59–65. [PubMed: 9328791]
19. Weinshenker BG, Santrach P, Bissonet AS, McDonnell SK, Schaid D, Moore SB, et al. Major histocompatibility complex class II alleles and the course and outcome of MS: a population-based study. Neurology. 1998; 51: p. 742–7. [PubMed: 9748020]
20. Smestad C, Brynedal B, Jonasdottir G, Lorentzen AR, Masterman T, Akesson E, et al. The impact of HLA-A and –DRB1 on age at onset, disease course and severity in Scandinavian multiple sclerosis patients. Eur. J. Neurol 2007; 14: p. 835–40. [PubMed: 17662002]
21. Greer JM, Pender MP. The presence of glutamic acid at positions 71 or 74 in pocket 4 of the HLA-DRbeta1 chain is associated with the clinical course of multiple sclerosis. Psychiatry. 2005; 76: p. 656–62.

22. Masterman T, Ligers A, Olsson T, Andersson M, Olerup O, Hillert J. HLA-DR15 is associated with lower age at onset in multiple sclerosis. Ann. Neurol. 2000; 48: p. 211–19. [PubMed: 10939572]

23. Stankovich J, Butzkueven H, Marriott M, Chapman C, Tubridy N, Tait BD, et al. HLA-DRB1 associations with disease susceptibility and clinical course in Australians with multiple sclerosis. Tissue Antigens. 2009; 74: p. 17–21. [PubMed: 19392788]

24. Polman CHRSC, Banwell B, Clanet M, Cohen JA, Filippi M, Fujihara K, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann. Neurol 2011; 69: p. 292–302. [PubMed: 21387374]

25. Lancaster AK, Single RM, Solberg OD, Nelson MP, Thomson G. PyPop update – a software pipeline for large-scale multilocus population genomics. Tissue Antigens. 2007; 69 Suppl 1: p. 192–7. [PubMed: 17445199]

26. Pappas DJ, Marin W, Hollenbach JA, Mack SJ. Bridging Immuno Genomic Data Analysis Workflow Gaps (BIGDAWG): An integrated case-control analysis pipeline.. Hum. Immunol. 2016; 77: p. 283–7. [PubMed: 26708359]

27. Dupont W, WD P. Power and sample size calculations: A review and computer program. Controlled Clinical Trials. 1990 4; 11(2): p. 116–128. [PubMed: 2161310]

28. Barcellos LF, Sawcer S, Ramsay PP, Baranzini SE, Thomson G, Briggs F, et al. Heterogeneity at the HLA-DRB1 locus and risk for multiple sclerosis. Human Molecular Genetics. 2006; 15: p. 2813–24. [PubMed: 16905561]

29. Healy BC, Liguori M, Tran D, Chitnis T, Glanz B, Wolfish C, et al. HLA-B*44: protective effects in MS susceptibility and MRI outcome measures. Neurology. 2010; 75: p. 634–40. [PubMed: 20713950]

30. Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, Moutsianas L, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. 2011; 476: p. 214–9. [PubMed: 21833088]

31. Wu JS, James I, Qiu W, Castley A, Christiansen FT, Carroll WM, et al. HLA-DRB1 allele heterogeneity influences multiple sclerosis severity as well as risk in Western Australia. J. Neuroimmunol 2010; 219: p. 109–13. [PubMed: 20006387]

32. Rivera VM. Multiple sclerosis in Latin Americans: Genetic Aspects. Curr Neurol Neurosci Rep. 2017; 57: p. 1–7.

33. Qiu W, James I, Carroll WM, Mastaglia FL, Kermode AG. HLA-DR allele polymorphism and multiple sclerosis in Chinese populations: a meta-analysis. Muliple Sclerosis. 2011; 17: p. 382–8.

34. Sheehy MJ, Scharf SJ, Rowe JR, Neme de Gimenez MH, Meske LM, Erlich HA, et al. A diabetes-susceptible HLA haplotype is best defined by a combination of HLA-DR and -DQ alleles. J Clin Invest. 1989; 83: p. 830–5. [PubMed: 2784133]

35. Gao XJ, Olsen NJ, Pincus T, Stastny P. HLA-DR alleles with naturally occurring amino acid substitutions and risk for development of rheumatoid arthritis. Arthritis Rheum.. 1990; 33: p. 939–46. [PubMed: 2369430]

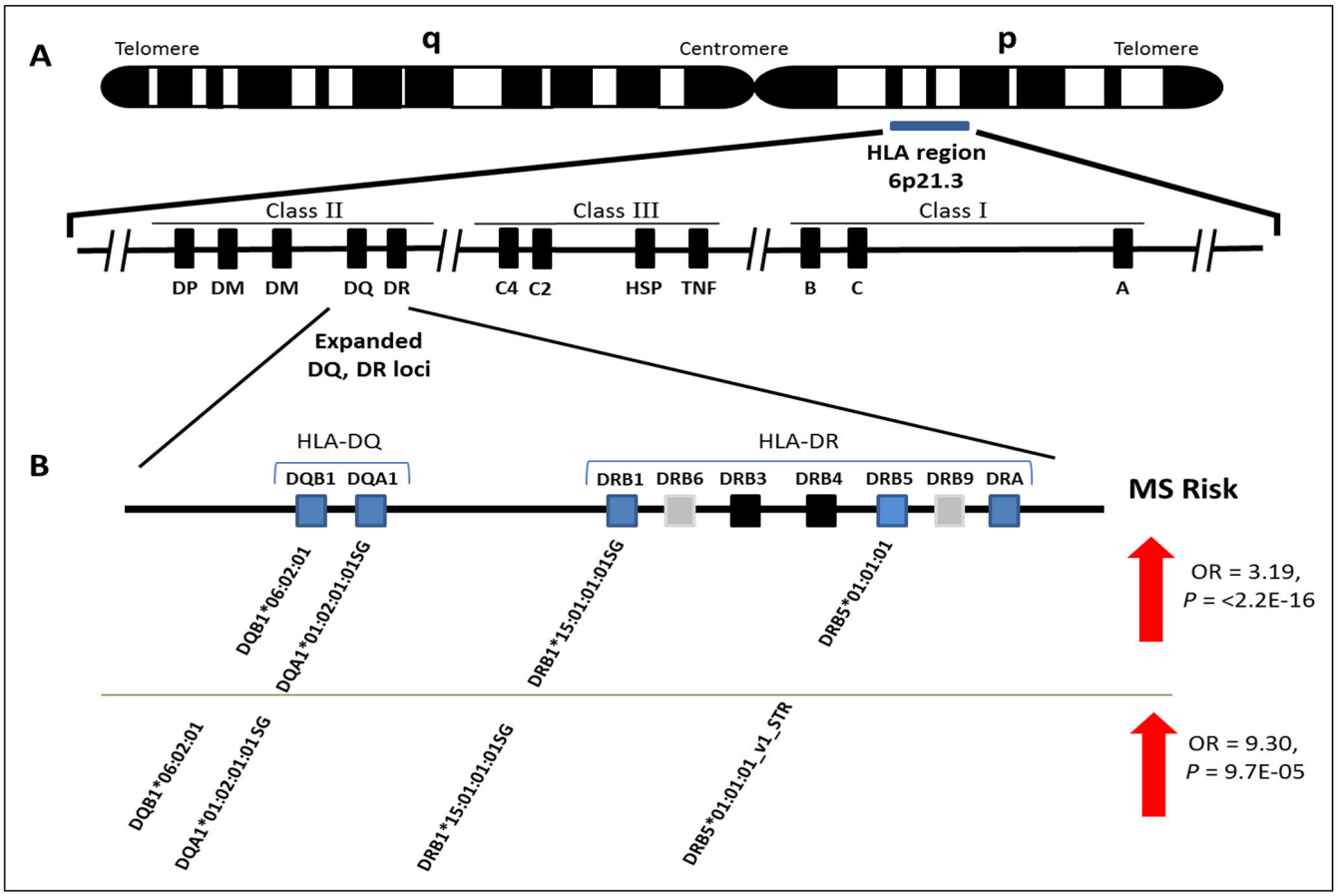

Figure 1.

*HLA-DRB1*15:01* haplotypes associated with Multiple Sclerosis in European Americans. (A) Map of HLA class I, III, and II genes, (B) MS associated *HLA-DRB1*15:01:01:01SG* haplotypes.

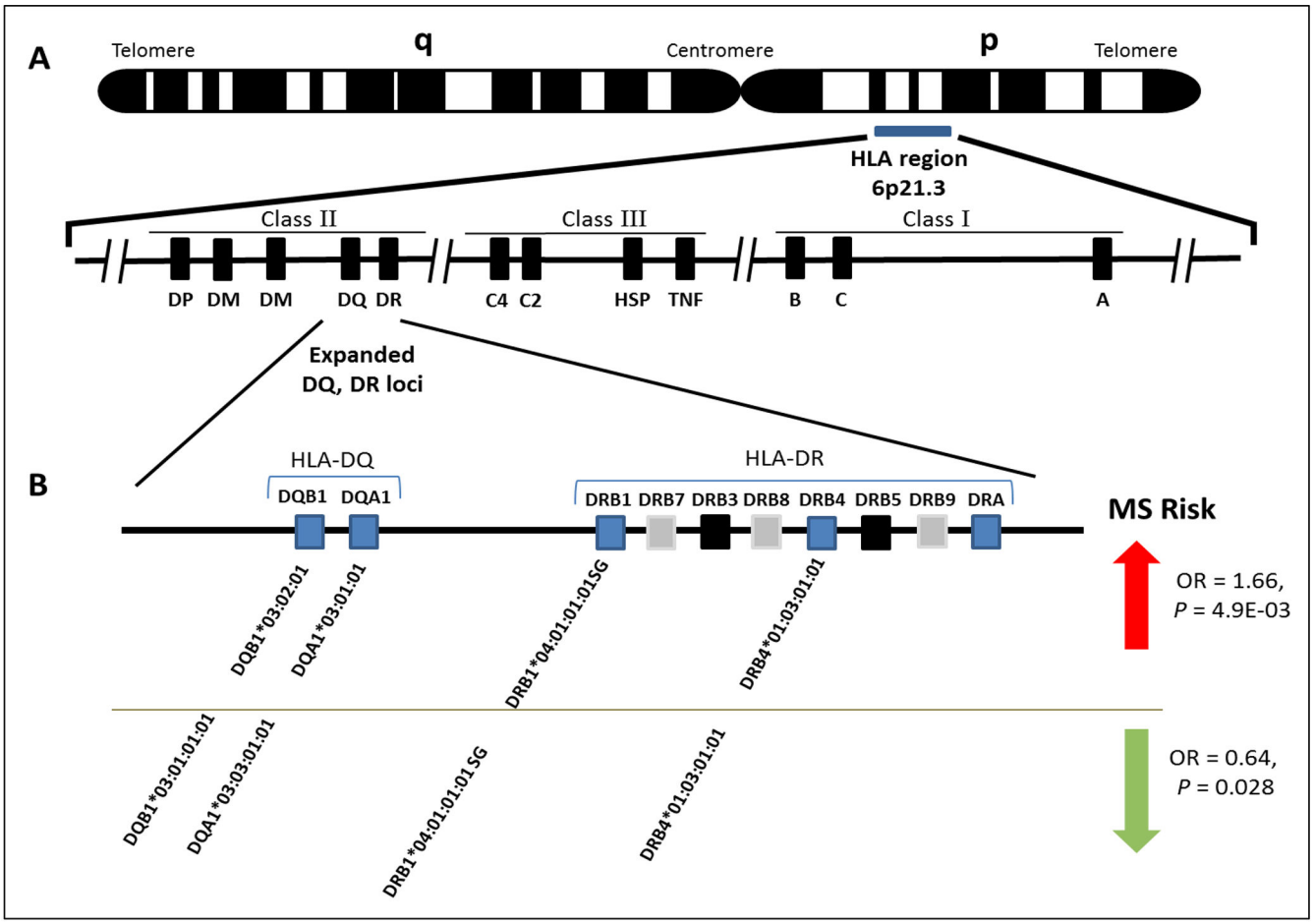

Figure 2.

*HLA-DRB1*04:01* haplotypes associated with Multiple Sclerosis in European Americans conditional on the absence of *HLA-DRB1*15:01* haplotypes. (A) Map of HLA class I, III, and II genes, (B) MS associated *HLA-DRB1*04:01:01:01SG* haplotypes.

Table 1.

Significant associations of *HLA-DRB1*15:01:01:01*, *HLA-DRB5*01:01:01*, and *HLA-DRB1*04:01:01:01* alleles in European American MS cases and controls

	MS Cases		Controls				
HLA allele	**Count**	**Fq**	**Count**	**Fq**	**OR**	**95% CI**	***P* value**
*HLA-DRB1*15:01:01:01SG*	314	0.5048	344	0.2414	3.20	2.63 – 3.91	<2.2E-16
*HLA-DRB5*01:01:01*	300	0.4823	341	0.2393	2.96	2.43 – 3.61	<2.2E-16
*HLA-DRB5*01:01:01v1_STR1*	14	0.0225	4	0.0028	8.18	2.68 – 24.95	4.3E-05
*HLA-DRB1*04:01:01:01SG*	93	0.0663	237	0.1663	0.36	0.28 – 0.46	<2.2E-16

Abbreviations: MS, Multiple Sclerosis; Fq, frequency; OR, odds ratio; CI, confidence interval

*HLA-DRB1*15:01:01:01SG* represents ambiguity group *HLA-DRB1*15:01:01:01/HLA-DRB1*15:01:01:02/HLA-DRB1*15:01:01:03*

*HLA-DRB1*04:01:01:01SG* represents ambiguity group *HLA-DRB1*04:01:01:01/HLA-DRB1*04:01:01:02*

Mult Scler. Author manuscript; available in PMC 2019 May 01.

Table 2.

Association of *HLA-DRB1*15:01* haplotypes in European American MS cases and controls

	MS Cases (2n=1236)		Controls (2n=2850)				
***HLA-DQB1~DQA1~DRB1*15:01~DRB5* haplotypes**	**Count**	**Fq**	**Count**	**Fq**	**OR**	**95% CI**	***P* value**
*DQB1*06:02:01~DQA1*01:02:01:01SG~DRB1*15:01:01:01SG~DRB5*01:01:01*	372	0.3010	339	0.1189	3.19	2.70 – 3.77	<2.2E-16
*DQB1*06:02:01~DQA1*01:02:01:01SG~DRB1*15:01:01:01SG~DRB5*01:01:01v1_STR1*	12	0.0097	3	0.0011	9.30	2.62 – 33.03	9.7E-05
*DQB1*06:03:01~DQA1*01:02:01:01SG~DRB1*15:01:01:01SG~DRB5*01:01:01*	10	0.0081	8	0.0028	2.90	1.14 – 7.36	0.0353
*DQB1*05:02:01~DQA1*01:02:01:01SG~DRB1*15:01:01:01SG~DRB5*01:01:01*	1	0.0008	4	0.0014	0.58	0.06 – 5.16	1.00
*DQB1*05:02:01~DQA1*01:02:02~DRB1*15:01:01:01SG~DRB5*01:01:01*	2	0.0016	2	0.0007	2.31	0.32 – 16.40	0.59
*DQB1*05:01:01:02~DQA1*01:02:01:01SG~DRB1*15:01:01:01SG~DRB5*01:01:01*	0	-	1	0.0004	0.77[+]	0.03 – 18.9	1.00
*DQB1*05:03:01:01~DQA1*01:02:01:01SG~DRB1*15:01:01:01SG~DRB5*01:01:01*	1	0.0008	0	-	6.92[+]	0.28 – 170.1	0.09
*DQB1*06:01:01~DQA1*01:02:01:01SG~DRB1*15:01:01:01SG~DRB5*01:01:01*	1	0.0008	0	-	6.92[+]	0.28 – 170.1	0.09
*DQB1*06:01:01~DQA1*01:02:01:01SG~DRB1*15:01:01:01SG~DRB5*01:01:01v1_STR1*	1	0.0008	0	-	6.92[+]	0.28 – 170.1	0.09
*DQB1*06:01:01~DQA1*01:03:01:01~DRB1*15:01:01:01SG~DRB5*01:01:01v1_STR1*	1	0.0008	0	-	6.92[+]	0.28 – 170.1	0.09
*DQB1*06:02:01~DQA1*01:02:01:01SG~DRB1*15:01:01:01SG~DRB5*01:01:02*	0	-	1	0.0004	0.77[+]	0.03 – 18.87	1.00
*DQB1*06:02:01~DQA1*01:02:01:01SG~DRB1*15:01:01:01SG~DRB5*01:05*	0	-	1	0.0004	0.77[+]	0.03 – 18.87	1.00
*DQB1*06:02:01~DQA1*01:02:01:01SG~DRB1*15:01:12~DRB5*01:01:01*	0	-	1	0.0004	0.77[+]	0.03 – 18.87	1.00
*DQB1*06:02:01~DQA1*01:02:02~DRB1*15:01:01:01SG~DRB5*01:01:01*	1	0.0008	0	-	6.92[+]	0.28 – 170.1	0.09
*DQB1*06:04:01~DQA1*01:02:01:01SG~DRB1*15:01:01:01SG~DRB5*01:01:01*	1	0.0008	0	-	6.92[+]	0.28 – 170.1	0.09
*DQB1*06:16~DQA1*01:02:01:01SG~DRB1*15:01:01:01SG~DRB5*01:01:01*	0	-	1	0.0004	0.77[+]	0.03 – 18.87	1.00
*DQB1*06:39~DQA1*01:02:01:01SG~DRB1*15:01:01:01SG~DRB5*01:01:01*	0	-	1	0.0004	0.77[+]	0.03 – 18.87	1.00

Abbreviations: MS, Multiple Sclerosis; Fq, frequency; OR, odds ratio; CI, confidence interval

*HLA-DQB1*05:03:01:01* represents ambiguity group *HLA-DQB1*05:03:01:01/HLA-DQB1*05:03:01:02*

*HLA-DQA1*01:02:01:01SG* represents ambiguity group *HLA-DQA1*01:02:01:01/HLA-DQA1*01:02:01:03/HLA-DQA1*01:02:01:05*

*HLA-DRB1*15:01:01:01SG* represents ambiguity group *HLA-DRB1*15:01:01:01/HLA-DRB1*15:01:01:02/HLA-DRB1*15:01:01:03*

[+] Haldane correction applied

Table 3.

Association of *HLA-DRB1*04:01:01* haplotypes in European American MS cases and controls

	MS Cases (2n=2806)		Controls (2n=2850)				
***HLA-DQB1~DQA1~DRB1*04:01:01:01~DRB4* haplotypes**	**Count**	**Fq**	**Count**	**Fq**	**OR**	**95% CI**	***P* value**
*DQB1*03:01:01:01~DQA1*03:03:01:01~DRB1*04:01:01:01SG~DRB4*01:03:01:01*	30	0.0107	138	0.0484	0.21	0.14 – 0.32	<2.2E-16
*DQB1*03:02:01~DQA1*03:01:01~DRB1*04:01:01:01SG~DRB4*01:03:01:01*	52	0.0185	96	0.0337	0.54	0.35 – 0.76	4.3E-04
*DQB1*03:02:01~DQA1*03:03:01:01~DRB1*04:01:01:01SG~DRB4*01:03:01:01*	10	0.0036	18	0.0063	0.56	0.26 – 1.22	0.11
*DQB1*03:01:01:01~DQA1*03:03:01:01~DRB1*04:01:01:01SG~DRB4*01:03:01:02N*	0	-	1	0.0004	0.34[+]	0.01 – 8.31	0.99
*DQB1*03:01:01:03~DQA1*03:03:01:01~DRB1*04:01:01:01SG~DRB4*01:03:01:01*	1	0.0004	0	-	3.05[+]	0.12 – 74.86	0.20
*DQB1*03:02:01~DQA1*03:01:01~DRB1*04:01:01:01SG~DRB4*01:01:01:01*	0	-	1	0.0004	0.34[+]	0.01 – 8.31	0.51
*DQB1*03:02:01~DQA1*03:01:01~DRB1*04:01:01:01SG~DRB4*01:03:01:02N*	1	0.0004	0	-	3.05[+]	0.12 – 74.86	0.31
*DQB1*03:02:01~DQA1*03:01:01~DRB1*04:01:01:01SG~DRB4*01:03:02*	0	-	1	0.0004	0.34[+]	0.01 – 8.31	0.51

Abbreviations: MS, Multiple Sclerosis; Fq, frequency; OR, odds ratio; CI, confidence interval

*HLA-DRB1*04:01:01:01SG* represents ambiguity group *HLA-DRB1*04:01:01:01/HLA-DRB1*04:01:01:02*

*HLA-DRB4*01:03:01:01* represents ambiguity group *HLA-DRB4*01:03:01:01/HLA-DRB4*01:03:01:03*

[+]Haldane correction applied

Table 4.

Association of *HLA-DRB1*04:01:01* haplotypes after removal of *HLA-DRB1*15:01* haplotypes

	MS Cases (2n=833)		Controls (2n=2488)				
***HLA-DQB1~DQA1~DRB1*04:01:01:01~DRB4* haplotypes**	**Count**	**Fq**	**Count**	**Fq**	**OR**	**95% CI**	***P* value**
*DQB1*03:02:01~DQA1*03:01:01~DRB1*04:01:01:01SG~DRB4*01:03:01:01*	52	0.0624	96	0.0386	1.66	1.17 – 2.35	4.9E-03
*DQB1*03:01:01:01~DQA1*03:03:01:01~DRB1*04:01:01:01SG~DRB4*01:03:01:01*	30	0.0360	138	0.0555	0.64	0.43 – 0.95	0.028
*DQB1*03:02:01~DQA1*03:03:01:01~DRB1*04:01:01:01SG~DRB4*01:03:01:01*	10	0.0120	18	0.0072	1.67	0.77 – 3.63	0.19
*DQB1*03:01:01:01~DQA1*03:03:01:01~DRB1*04:01:01:01SG~DRB4*01:03:01:02N*	0	-	1	0.0004	1.00[+]	0.04 – 24.45	1.00
*DQB1*03:01:01:03~DQA1*03:03:01:01~DRB1*04:01:01:01SG~DRB4*01:03:01:01*	1	0.0012	0	-	8.97[+]	0.37 – 220.43	0.06
*DQB1*03:02:01~DQA1*03:01:01~DRB1*04:01:01:01SG~DRB4*01:01:01:01*	0	-	1	0.0004	1.00[+]	0.04 – 24.45	1.00
*DQB1*03:02:01~DQA1*03:01:01~DRB1*04:01:01:01SG~DRB4*01:03:01:02N*	1	0.0012	0	-	8.97[+]	0.37 – 220.43	0.06
*DQB1*03:02:01~DQA1*03:01:01~DRB1*04:01:01:01SG~DRB4*01:03:02*	0	-	1	0.0004	1.00[+]	0.04 – 24.45	1.00

Abbreviations: MS, Multiple Sclerosis; Fq, frequency; OR, odds ratio; CI, confidence interval

*HLA-DRB1*04:01:01:01SG* represents ambiguity group *HLA-DRB1*04:01:01:01/HLA-DRB1*04:01:01:02*

*HLA-DRB4*01:03:01:01* represents ambiguity group *HLA-DRB4*01:03:01:01/HLA-DRB4*01:03:01:03*

[+] Haldane correction applied