Heterogeneity at the \textit{HLA-DRB1} locus and risk for multiple sclerosis

Lisa F. Barcellos\textsuperscript{1,2,3,*}, Stephen Sawcer\textsuperscript{4}, Patricia P. Ramsay\textsuperscript{1}, Sergio E. Baranzini\textsuperscript{2}, Glenys Thomson\textsuperscript{5}, Farren Briggs\textsuperscript{1}, Bruce C.A. Cree\textsuperscript{2}, Ann B. Begovich\textsuperscript{6}, Pablo Villoslada\textsuperscript{7}, Xavier Montalban\textsuperscript{8}, Antonio Ucelli\textsuperscript{9}, Giovanni Savettieri\textsuperscript{10}, Robin L. Lincoln\textsuperscript{7}, Carolyn DeLoa\textsuperscript{2}, Jonathan L. Haines\textsuperscript{11}, Margaret A. Pericak-Vance\textsuperscript{12}, Alastair Compston\textsuperscript{4}, Stephen L. Hauser\textsuperscript{2}, Jorge R. Oksenberg\textsuperscript{2}

\textsuperscript{1}Division of Epidemiology, 140 Warren Hall MC #7360, School of Public Health, University of California, Berkeley, CA 94720, USA

\textsuperscript{2}Department of Neurology, University of California, San Francisco, CA 94143, USA

\textsuperscript{3}Kaiser Permanente Northern California, Division of Research, 2000 Broadway, Oakland, CA 94612, USA

\textsuperscript{4}University of Cambridge, Department of Clinical Neurosciences, Addenbrooke’s Hospital, BOX 165 Cambridge, CB2 2QQ, UK

\textsuperscript{5}Department of Integrative Biology, University of California, Berkeley, CA 94720, USA

\textsuperscript{6}Celera Diagnostics, Alameda, CA 94502, USA

\textsuperscript{7}Department of Neurology, University of Navarra, Pamplona, Spain

\textsuperscript{8}Neuroimmunology Unit, Hospital Vall d’Hebron, Barcelona, Spain

\textsuperscript{9}Department of Neurology, University of Genova, Genova, Italy

© The Author 2006. Published by Oxford University Press. All rights reserved
10 Department of Neurology, Ophthalmology, Otolaryngology and Psychiatry, University of Palermo, Palermo, Italy

11 Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, TN 37232, USA

12 Center for Human Genetics, Duke University Medical Center, Durham, NC 27710, USA

*Corresponding Author; barcello@genepi.berkeley.edu; Tel: (510) 642-7814, Fax: (510) 643-5163
Abstract

Variation in major histocompatibility complex (MHC) genes on chromosome 6p21.3, specifically the HLA-DR2 or DRB1*1501-DQB1*0602 extended haplotype, confers risk for multiple sclerosis (MS). Previous studies of DRB1 variation and both MS susceptibility and phenotypic expression have lacked statistical power to detect modest genotypic influences, and have demonstrated conflicting results. Results derived from analyses of 1,339 MS families indicate DRB1 variation influences MS susceptibility in a complex manner. DRB1*15 was strongly associated in families (p=7.8 x 10^{-31}), and a dominant DRB1*15 dose effect was confirmed (OR=7.5, 95% CI=4.4, 13.0, p<0.0001). A modest dose effect was also detected for DRB1*03; however, in contrast to DRB1*15, this risk was recessive (OR=1.8, 95% CI=1.1, 2.9, p=0.03). Strong evidence for under-transmission of DRB1*14 (p=5.7 x 10^{-6}), even after accounting for DRB1*15 (p=0.03) was present, confirming a protective effect. In addition, a high risk DRB1*15 genotype bearing DRB1*08 was identified (OR=7.7, 95% CI=4.1, 14.4, p<0.0001), providing additional evidence for trans DRB1 allelic interactions in MS. Further, a significant DRB1*15 association observed in primary progressive (PPMS) families (p=0.0004), similar to relapsing-remitting (RRMS) families, suggests that DRB1 related mechanisms are contributing to both phenotypes. In contrast, results obtained from 2,201 MS cases argue convincingly that DRB1*15 genotypes do not modulate age of onset, or significantly influence disease severity measured using Expanded Disease Disability Score (EDSS) and disease duration. These results contribute substantially to our understanding of the DRB1 locus and MS, and underscore the importance of using large sample sizes to detect modest genetic effects, particularly in studies of genotype-phenotype relationships.
Introduction

Multiple sclerosis (MS) is a chronic inflammatory disorder of the central nervous system (CNS) characterized by demyelination, astrogliosis, varying degrees of axonal pathology, and a relapsing or progressive course (1). Although etiologic mechanisms are uncertain, MS risk is determined by an underlying complex genetic component likely acting in concert with undefined environmental exposures. Genetic susceptibility to MS is associated with the HLA region located on the short arm of chromosome 6. There are two major classes of HLA genes. The telomeric stretch of the locus contains the class I genes, whereas the centromere proximal region encodes HLA-class II genes. HLA class I and class II encoded molecules are highly polymorphic cell surface glycoproteins playing a fundamental role in self/non-self immune recognition. The association of MS with HLA class II genes, specifically the HLA-DR2 or DRB1*15 haplotype (DQB1*0602, DQA1*0102, DRB1*1501, DRB5*0101), has been a consistent finding across nearly all populations (2). The exact mechanism(s) by which the DRB1*15 haplotype influences susceptibility to MS remain undefined, but are likely related to the physiological function of HLA molecules in immune responses, including antigen binding and presentation, and T cell repertoire determination by negative selection of high-avidity autoreactive T cells within the embryonic thymic microenvironment (3, 4).

In multi-case MS families, linkage and association to the HLA-DRB1/DQB1 locus have also been unambiguously manifested (5-7), but remarkably, all linkage information and evidence for association have been derived from families in which DRB1*15 was present in at least one nuclear member (7, 8). There are no reports of statistically significant evidence for linkage in DRB1*15 negative families, although some researchers have observed modest linkage (9).
underscoring the complex genetic nature of this disease. The debate concerning the role of non-HLA class II genes mapping to this region continues, with some data suggesting that additional susceptibility genes lie within the central class III (10, 11) and/or telomeric to the class I HLA regions (12-15). However, a recent high-density SNP study covering the MHC region assigns the entire association signal to the HLA-class II region (16), and DRB1 itself remains the strongest candidate for the disease-associated locus.

Recent studies of the HLA class II MS association have refined the location of the primary effect to DRB1 *15 (17, 18) and detected an unexpected DR2 (DRB1*15) allele dose effect on susceptibility (19). Further, strong evidence for DRB1 allelic and genotypic heterogeneity (13, 17, 20, 21) has been reported. Together these data are helping to refine the conceptual model of MS pathogenesis and suggest the possibility that complex trans DRB1 allelic interactions may determine the balance between susceptibility and resistance. Recent reports of disease protective DRB1 alleles require confirmation (21), and additional questions about DRB1 variation and clinical manifestations also persist; for example, HLA-DRB1*15 has been associated with an earlier age of disease onset, female gender, severe, relapsing-remitting, and mild MS courses (19, 22-24), or has shown a very minor or no influence on disease course (25-27). Complete resolution of these important issues is crucial to fully understand the role of DRB1 in disease pathogenesis. In this study, one of the largest familial MS datasets, to date, was assembled, including 1,339 families and a total of 2,294 cases, to comprehensively study the HLA-DRB1 locus in MS. The results provide strong evidence for complex genotypic DRB1 patterns associated with susceptibility, but not age of onset or disease progression, and demonstrate the
need to study large and well-characterized datasets for complex phenotypes even in the evaluation of primary genetic determinants.

Results

To better understand the role of *HLA-DRB1* variation in MS, and in particular, the risk and phenotype associated with *DRB1*15 genotypes, we comprehensively analyzed the *DRB1* locus in 1,339 well-characterized MS families (total=4,669 individuals; 1,571 MS cases and 3,098 unaffected family members). Approximately 85% of families were of northern European origin residing in the US or UK; the remaining families were southern Europeans, specifically from Northern Italy, Sicily and the Mediterranean Spanish basin. An additional 721 MS cases, (414 and 307 from the UK and US, respectively), without parental HLA data were also available for analyses of *DRB1* and clinical phenotypes. A total of 2,201 (of 2,294) MS cases had complete clinical and HLA data for analysis of phenotypes (96% of the dataset); see Table 1.

**HLA-DRB1 allele associations with disease susceptibility**

As expected, over transmission of the *DRB1*15 allele and highly significant association between MS and *DRB1* were observed (global test in all families, \(p=6.7 \times 10^{-33}\); Table 2). The *DRB1*15 association (\(p=7.8 \times 10^{-31}\)) was prominent in both northern and southern European datasets considered separately (data not shown). To identify predisposing or protective effects attributable to other *DRB1* alleles, the confounding influence of *DRB1*15 was excluded by repeating the analysis in just those families (n=494) without *DRB1*15. No significant evidence for any effects attributable to other *DRB1* alleles was observed. Most notably, no evidence for any effect of the more common *DRB1*03 or less common *DRB1*04 alleles associated with MS.
susceptibility in some Mediterranean (13, 20, 28-30), European (31), and African American populations (17) was found in this dataset. The DRB1*14 allele was protective (p=0.03), although this association would not remain significant after a conservative correction for multiple tests. Ignoring discordant sib pairs and only including trios from MS families did not change results. When southern European families were removed from the analyses, results were very similar (data not shown).

**HLA-DRB1 genotype associations with disease susceptibility**

A comprehensive strategy was employed to investigate potential DRB1 genotypic effects and MS risk. Using the genoPDT, we considered a total of 115 unique DRB1 genotypes and, in line with the results observed for DRB1 alleles, found highly significant evidence for association with MS (global p=8.7 x 10^{-17}; 1,294 families including 1,094 triads and 1,215 discordant sib pairs). Transmissions for each individual genotype were examined, and significant evidence for over-transmission was present only for those genotypes carrying the DRB1*15 allele (data not shown). When families carrying at least one DRB1*15 allele were removed from the analysis, a total of 89 unique genotypes were observed. Similarly, no genotypes were associated with MS after correction for multiple testing (494 families including 394 triads and 338 discordant sib pairs; data not shown).

To determine whether other DRB1 alleles influence the magnitude of risk conferred by DRB1*15, conditional logistic regression (CLR) modeling of DRB1 genotypes in MS families was then utilized. Two CLR analyses were performed. First, all family members were categorized by disease status (affected and unaffected) and DRB1 genotype. Second,
independent trios were derived from the larger family dataset and matched ‘pseudocontrols’ were constructed from non-transmitted parental alleles as previously described (32). Similarly, cases from trio families and matched pseudocontrols were categorized by \textit{DRB1} genotype for CLR modeling. Two reference groups were designated based on results obtained from initial genoPDT screening of \textit{DRB1} genotypes: the “low-risk” \textit{DRB1}*\textit{X/X} (where \textit{X}=non \textit{DRB1}*15 genotypes) and the “high-risk” \textit{DRB1}*15/15 genotype groups. Odds ratios and 95% confidence intervals were determined and are shown in Figure 1. Results for all family members and cases with matched pseudocontrols using both low (Fig. 1A) and high (Fig. 1B) risk reference \textit{DRB1} genotypes were similar. As expected, almost all \textit{DRB1}*15 genotypes demonstrated significantly increased odds of disease risk when compared to the \textit{DRB1}*\textit{X/X} genotype reference group (Fig. 1A). The \textit{DRB1}*15/15 genotype conferred highest risk for disease (OR=9.8, 95% CI=6.6, 14.6, \textit{p}<0.0001) while the observed odds ratios for \textit{DRB1}*15/1, \textit{DRB1}*15/3, \textit{DRB1}*15/4, \textit{DRB1}*15/7, \textit{DRB1}*15/11 and \textit{DRB1}*15/13 genotypes ranged between 3.5 and 5.0 (p-values <0.0001; based upon analyses of MS cases and pseudocontrols; see Figure 1 legend). When UK families (n=492) were considered independently, the previously reported \textit{DRB1}*15 dose effect on disease risk was confirmed (19): \textit{DRB1}*15/15 vs. \textit{DRB1}*\textit{X/X}: OR=7.5, 95% CI=4.4, 13.0, \textit{p}<0.0001 and \textit{DRB1}*15/\textit{X} vs. \textit{DRB1}*\textit{X/X}: OR=3.4, 95% CI=2.4, 4.8, \textit{p}<0.0001.

Unexpected results were observed, however, for two \textit{DRB1}*15 genotypes: \textit{DRB1}*15/08 and \textit{DRB1}*15/14. \textit{DRB1}*15/08 emerged as a high-risk genotype (Figure 1A; OR=7.7, 95% CI=4.1, 14.4, \textit{p}<0.0001), and could not be statistically distinguished from \textit{DRB1}*15/15 (Figure 1B; OR=0.8, 95% CI=0.4, 1.5, \textit{p}=0.47). The odds for disease was greater for carriers of \textit{DRB1}*15/08 when compared to \textit{DRB1}*15/\textit{X} genotypes (OR=1.9, 95% CI=1.1, 3.1, \textit{p}=0.016; Table 3).
Further, the DRB1*08 allele was more frequent in DRB1*15/X MS cases compared to non-transmitted parental (control) chromosomes, (5.4% vs. 3.2%; p=0.02; Table 3), which provided additional evidence for an influence of DRB1*08 on MS risk, but only in the presence of DRB1*15.

DRB1*15/14 demonstrated strong evidence for a protective effect when contrasted with DRB1*15/X genotypes (OR=0.2, 95% CI=0.1, 0.5, p=0.0015, Table 3). In line with these results, DRB1*14 was also present at a lower frequency in DRB1*15/X cases when compared to non-transmitted parental (control) chromosomes (0.2% vs. 3.3%; p=0.0002; Table 3). Considering transmission from non-DRB1*15 parents (present in 350 trio families) to DRB1*15 positive MS cases also revealed evidence for a protective DRB1*14 effect, even in the presence of DRB1*15. Specifically, the rare DRB1*14 allele was transmitted less often from non-DRB1*15 parents to DRB1*15 MS cases (T:NT = 1:10, p=0.01, data not shown). Collectively, results indicate that DRB1*14 significantly reduces MS risk, even for carriers of DRB1*15.

Similar to previous results from a Swedish MS dataset (33), association between MS and DRB1*03/X genotypes was not observed (OR=0.9, 95% CI=0.7-1.1, p=0.20); however, when individuals carried two copies of DRB1*03, evidence for a modest but increased risk for MS was present (OR=1.8, 95% CI=1.1-2.9, p=0.03), even when accounting for DRB1*15 genotypic effects (data not shown). The DRB1*03/03 genotype was present in only 2.2% of MS cases. Results indicate that CLR modeling had more power to detect this less common genotypic effect compared to the genoPDT.
**HLA-DRB1*15 genotypes and clinical phenotype**

The influence of *HLA-DRB1*15 genotypes on age at onset and disease severity was examined (total n= 2,201 MS cases; see Table 1). Here, age of onset was defined as the first episode of focal neurological dysfunction suggestive of CNS demyelinating disease (8). This information was obtained via individual recall and verified through review of medical records. The mean age of onset in male MS cases, overall, was very similar to female cases (28.8 +/- 9.2 vs. 28.2 +/- 8.2, respectively, p=0.47), even when analyses were restricted to cases with an initial relapsing-remitting course (p=0.70). Age of onset distributions for MS cases grouped according to *DRB1*15 genotype status were compared to identify mean differences. When each mean age of onset in cases based on *DRB1*15 genotypic categorization was compared to the mean value derived from *DRB1*X/X cases, no significant differences were observed, even after adjustment for gender and country of origin (data not shown). Furthermore, no differences in distribution of the *DRB1*15 genotypes were present when male and female cases were considered separately (data not shown).

The Global MS severity score (MSSS) was used as a measure of disability (34). This algorithm adjusts disability as measured by the ‘expanded disability status score’ or EDSS (35) for disease duration. Disease duration was measured as the number of years between the year of onset of first symptom and year of last exam with EDSS assessment. The mean MSSS in male MS cases was significantly higher when compared to female cases, even when analyses were restricted to cases with an initial relapsing-remitting course (5.2 +/- 2.7 vs. 4.7 +/- 2.6, p=0.0009). However, when cases were grouped according to *DRB1*15 genotype and compared to *DRB1*X/X...
individuals, no statistically significant differences in MSSS distributions were observed (Figure 3), even after adjustment for gender and country of origin (data not shown).

Analyses also utilized ‘mild’ and ‘severe’ designations (most extreme clinical phenotypes) to examine, specifically, the HLA-DRB1*15 dose effect as previously reported (19). Mild disease (the ability to walk normally or have only mild gait disability) was defined for cases with an EDSS ≤3 after 15 years. Severe disease (the need for bilateral assistance to walk or wheelchair dependency) included cases who reached an EDSS >6, within 10 years of disease duration. When mild (n=146, 6.6% of dataset) and severe (n=97, 4.4% of dataset) MS cases derived from the larger dataset were considered separately, the presence of DRB1*15 (either one or two copies) did not distinguish one phenotype from the other (DRB1*15/15 vs. DRB1*X/X, OR=1.8, 95% CI=0.6-5.0, p=0.29; DRB1*15/X vs. DRB1*X/X, OR=0.8, 95% CI=0.4-1.5, p=0.54; data not shown).

HLA-DRB1*15 and primary progressive MS

DRB1*15 genotype frequencies for PPMS (total n=87 cases; 31 were from US and 56 from UK) and initial relapsing-remitting or RRMS (including SPMS); total n=1,036 cases; 544 from US and 492 from UK) were compared. Control DRB1*15 genotype frequencies were derived using non-transmitted allele frequencies from 433 parents of US MS cases, under assumptions of Hardy-Weinberg equilibrium(36). Similar DRB1*15 genotype frequencies were observed in both PPMS and RRMS/SPMS cases and were higher compared to controls, suggesting that DRB1*15 contributes to MS pathogenesis in both subgroups (Figure 3). These analyses were restricted to the UK and US datasets where overall DRB1*15 frequencies were similar. A total
of 48 PPMS families (27 complete trios and 38 discordant sib pairs) were also available for PDT analysis. The results were highly significant for DRB1 (global p=0.003) and for over-transmission of DRB1*15, specifically (p=0.0004). PDT analyses utilized all available PPMS families within the overall dataset.
Discussion

A strong genetic association between MS and the \textit{HLA-DRB1*1501, DQB1*0602} haplotype has been a recurrent finding across populations, the notable exceptions being the association with five different \textit{DRB1-DQB1} haplotypes in cases from Sardinia (13, 37), and other Mediterranean and Asian populations (20, 28, 29, 38). Recent studies in MS cases from Martinique and African Americans indicate an association with the \textit{DRB1} gene independent of \textit{DQB1} further supporting the hypothesis that the disease is \textit{DRB1}-peptide driven (17, 39). Antigenic peptides are bound to the \textit{DRB1} encoded molecules in an extended conformation anchoring in key pockets shaped by specific polymorphic residues. There is, however, discrepancy surrounding the occurrence and range of \textit{DRB1} allelic and genotypic heterogeneity in MS. Here, using a large, stringently ascertained and well-characterized familial MS dataset, a comprehensive investigation of \textit{DRB1} variation and disease risk was performed. Genetic association to \textit{HLA-DRB1} and a strong association with \textit{DRB1*15} were observed in the overall dataset, as previously reported for the individual populations. The \textit{DRB1*15} dose effect on susceptibility (19) was replicated, and the preferential transmission of \textit{DRB1*08} in \textit{trans} with \textit{DRB1*15} and under-transmission of \textit{DRB1*14} recently detected in MS families ascertained in Canada (21), were confirmed. \textit{DRB1*15/08} and \textit{DRB1*15/14} genotypes were clearly identified as high risk and protective, respectively, for MS, in comparison to other disease predisposing \textit{DRB1*15} genotypes. These results strongly support the concept that complex interactions among \textit{DRB1} alleles, and perhaps extended \textit{DRB1} haplotypes and unidentified gene(s) contained within, confer both susceptibility and resistance.
The robust DRB1 genotypic effect detected in European and North-American MS families is not consistent with a single locus dominant disease model proposed to explain certain HLA-disease associations, best represented by the association of HLA-B27 and ankylosing spondylitis (40). Dominantly acting MHC genes are thought to function via high-affinity binding to self-peptides derived from the target organs, which are then efficiently presented to pathogenic autoreactive T cells. In MS, however, complex DRα/DRβ heterodimer effects in trans were detected, showing a disease association gradient ranging from high vulnerability (DRB1*15 homozygotes) to low susceptibility or resistance (DRB1*15/*14 heterozygotes). Adding to the complexity of the HLA contribution to MS, a dose effect was also detected for DRB1*03 and disease risk; however, in contrast to DRB1*15, the risk conferred by DRB1*03 appears to be recessive, with two copies significantly increasing risk (almost two-fold) compared to no risk for individuals carrying one copy, even after accounting for DRB1*15. The results are similar to a recent report in Swedish MS cases (33), but differ from a Canadian MS study (21). MS cases from both of these populations demonstrate a strong DRB1*15 association. While a primary DRB1*03 allelic association has been noted in other populations where the DRB1*15 effect is less prominent or absent, this was not observed in the current study.

DRB1 alleles that confer susceptibility differ from non-associated alleles at only a few positions in the binding site, implying a high degree of specificity. Structural studies demonstrate that polymorphic residues which affect the shape and charge of the P4 pocket in the peptide-binding site are important determinants in DRB1 associated human autoimmune diseases. DRβ1*1501 is different from other DRβ molecules in that aromatic residues in the ligand are preferred by the large hydrophobic P4 pocket of the peptide binding domain (41). For the MS putative...
autoantigen myelin basic protein (MBP), this pocket is primarily occupied by the aromatic side chain Phe92, acting as an important primary anchor and accounting for its high-affinity binding to the HLA-DRα0101/DRβ1501 heterodimer. The polymorphic residue at DRβ71 is critically important in creating the necessary space for Phe92 of MBP, and the uncharged Ala at this position is only observed for DR15 alleles (DRB1*1501-DRB1*1506) and the rare DRB1*1309 allele. Full molecular typing was available for more than 70% of the DRB1*15 alleles observed in this study; > 99% were DRB1*1501. The two aromatic residues of the DRβ1501 P4 pocket, β26Phe and β78Tyr facilitate the binding of aromatic side chains. The molecular structures of the permisive alleles DRB1*08 and DRB1*0301, identified as MS-associated in this and other studies, are significantly divergent in the proximity of the bound peptide, suggesting either they present a different MBP epitope or perhaps a different autoantigen. Hence, in the case of DRB1*08, this allele may synergize with DRB1*1501 by extending the aberrant immune response to either a minor epitope, or a novel encrypted determinant uncovered through molecular spreading (42).

Interestingly, alignment of all MS susceptibility alleles identified to date, DRB1*1501, *1503, *0301, *0401, and *0801/3 (in this study) reveals sharing of an aromatic residue (Tyr) at position 60 whereas the most common forms of the resistance allele DRB1*14 (DRB1*140101, *140102, *140103, *1404) carry the basic residue His. DRB1 genotyping methods used in this study did not distinguish every rare DRB1*14 variant; however > 90% of alleles with high resolution typing (~30% of the dataset) were *1401 (82%) or *1404 (9%). Although the crystal structure of DRβ1*14 is not available, computer modeling suggests that position 60 affects the shape and charge of the P9 pocket at the binding site (Fig. 4). To accommodate the observed
dominant protective effect of this allele over DRB1*15, it can be speculated that sub-optimal engagement of the encephalitogenic peptide leads to a dominant protective immune-deviation similar to that observed upon injection of altered peptide ligands (APLs) (43). Despite the apparent failure of APL in modulating MS at non-allergenic doses, a new generation of therapeutic synthetic peptides developed based upon a better understanding of the molecular structure of HLA-DRB1 molecules and association to autoimmunity may be available for clinical trials in the near future (44). The observation of disparate DRB1 genotypic effects on MS susceptibility is consistent with a multi-locus effect model comprised of a dominantly acting susceptibility gene present on DRB1*15 haplotypes (or recessive in DRB1*03 haplotypes) plus the absence of a protective gene required for maintenance of peripheral tolerance present on non-DRB1*15 haplotypes. The proposed models are testable using single and double transgenic mice exposed to myelin epitopes (45). In addition, the imminent availability of genetic maps defining discrete haplotype bins in the HLA extended region (46, 47) will provide a useful reference and necessary tools to identify the true disease gene(s) for MS operating within this superlocus.

Whether DRB1*15 influences age of onset or disease severity in MS has been controversial. A new measure, the MSSS, was utilized in this study to evaluate the relationship between expanded DRB1*15 genotypes and disease severity (34). The MSSS algorithm adjusts the most validated and widely accepted measure of disability, the EDSS (35) for disease duration by comparing an individual’s disability with the distribution of scores in cases having equivalent disease duration. Simulations have revealed that the MSSS may be more powerful than other methods for measurement of severity in MS (34) including mild and severe categorizations and
two types of progression indices (48, 49). Using the Global MSSS approach, no significant
differences in score distributions for cases characterized by common \(DRB1*15\) genotypes were
detected. Further, no evidence for a \(DRB1*15\) influence on extreme clinical phenotypes such as
mild (6.6% of cases) or severe (4.4% of cases) courses defined by EDSS and disease duration (8,
19, 50) was identified. Neither age of onset nor gender in MS cases were associated with
\(DRB1*15\) status or \(DRB1*15\) genotypic variation (data not shown).

Common \(DRB1*15\) genotypes do not appear to modulate age of onset or severity in MS.
However, ‘radiological relapses’ using magnetic resonance imaging (MRI), defined as the
presence of gadolinium-enhancing lesions, may be a better indicator of disease burden compared
to a neurological examination that includes the EDSS. This is currently under debate (33), and
has important implications for genetic studies. Neither MRI nor drug therapy information were
available for all MS cases in this study. Treatment options for MS specifically target the
inflammatory phase and include immunomodulators such as interferon betas and glatiramer
acetate, and an immunosuppressant, mitoxantrone (51). Confounding in this study remains
possible, if MS cases with distinct \(DRB1\) genotypes respond differentially to particular
treatments. Interestingly, \(DRB1*15\) status does not appear to influence response to interferon
therapy (27), though larger studies are needed to fully address this question.

MS cases with a primary progressive disease course (PPMS), approximately 10% of all cases,
experience a gradual but insidious progression of disability from onset without superimposed
relapses, and demonstrate additional demographic, pathological and imaging features which
distinguish them from RRMS (52-54). It is not certain if these features result from different
etiologic mechanisms or, alternatively, represent opposite ends of a clinical spectrum. Whether *HLA-DRB1*15 is associated with PPMS, specifically, has not been previously established; several small studies failed to show any association between PPMS and *DRB1*15, whereas others appeared to show association (25, 48, 55). In this study, *DRB1*15 status did not distinguish PPMS from RR/SPMS, as *DRB1*15 frequencies were very similar in both subgroups. In addition, PPMS families also demonstrated significant evidence for over-transmission of *DRB1*15 to MS cases. Taken together, these results suggest PPMS and RR/SPMS share a common HLA-related pathoetiology, and indicate both phenotypes are within the same disease spectrum. If correct, this model has important implications for disease pathogenesis and therapy.

In summary, our results show that the *DRB1* contribution to MS risk is more complex than originally considered. Whether genotypic heterogeneity in MS is due to differences in affinity attributed to specific *DRB1* alleles for a number of autoantigenic peptides, or indicates the involvement of other nearby genes within the MHC remains unknown. In light of these findings, functional studies of MS-related *DRB1* alleles and genotypic combinations identified in this study are warranted.
Materials and Methods

Diagnostic criteria, ascertainment protocols and clinical and demographic characteristics for MS cases and family members are summarized elsewhere (19, 20, 27, 56). Appropriate institutional review boards approved all studies and written informed consent was obtained from all participants.

*HLA-DRB1* genotypes were determined as previously described (19, 57); both low (2 digit) and medium (4 digit) resolution typing methods were utilized. Medium resolution typing was available for approximately half of the entire dataset, or 50.3% of all alleles; the majority of these were *DRB1*15. More than 70% of the *DRB1*15 alleles observed in this study had 4 digit typing; > 99% were characterized as *DRB1*1501. *DRB1* genotyping methods used in this study did not distinguish every rare *DRB1*14 variant; however > 90% of alleles with medium resolution typing (30% of the dataset) were *1401* (82%) or *1404* (9%). Medium resolution typing was performed for 50% of *DRB1*03 alleles (99% were distinguished as *0301*) and 32% of *DRB1*08 alleles (78% were *0801* and 16% were *0803*). To maximize the statistical power of the large MS family dataset, low resolution allele grouping was used for all analyses.

All family genotypes were examined for Mendelian inconsistencies using PEDCHECK (58) and any discrepancies addressed. Family-based association analyses using the pedigree disequilibrium test (PDT) v. 5.1 (59) were performed. Two PDT statistics were used: the PDT-sum statistic (60), which examines allelic effects, and the genotype-PDT that examines genotypic effects (61). The PDT is a powerful analytical method that uses genetic data from related nuclear families and discordant sibships within extended pedigrees. Conditional logistic
regression (CLR) modeling of DRB1 genotypes in MS families was used to assess the magnitude of disease risk associated with DRB1*15 and DRB1*03 genotypes in the MS families using PROC THREG as implemented in SAS (v. 9.1; SAS Institute, Cary, NC). All affected and unaffected family members were utilized in this analysis with m:n matching (where a varying number of affected and unaffected members were present in each matched set or family). In addition, one MS case and three matched ‘pseudocontrols’ constructed from the non-transmitted parental alleles as previously described (32) were used in (CLR) analyses. One fully genotyped independent trio (both parents and one affected) from each MS family was selected for this analysis. Affected family-based controls (non-transmitted parental alleles or ‘AFBAC’) were derived as previously described (36). The chi-square test of heterogeneity was used to test the null hypothesis that the distribution of non-DRB1*15 or ‘X’ alleles in DRB1*15/X MS cases is similar to the distribution of non-DRB1*15 or ‘X’ alleles in controls. P-values, odds ratios and confidence intervals for all chi-square or Fisher’s exact test of allele or genotype case-control comparisons were derived using SAS (v. 9.1; SAS Institute, Cary, NC).

The MSSS (34) was used to examine DRB1 genotypic effects influencing severity. The MSSS algorithm is a simple method for adjusting disability for disease duration, and is based upon data obtained from a collection of approximately 10,000 MS cases across the world. The Global MSSS is the decile of the EDSS within the range of cases who have had the disease for the same duration. Cases in this study were assigned a global score and then tested for differences in median scores between cases grouped according to DRB1 genotypic categories using the non-parametric Wilcoxon Rank Sums test implemented in ranksum (Stata v. 8.2, StataCorp LP, College Station, TX). Two sample t-tests and linear regression were also used to compare mean
MSSS in cases grouped according to DRB1 genotype. These analyses were performed in Stata (regress, ttest, Stata v. 8.2), and results were adjusted for gender and country of origin. Cases for whom EDSS scores were designated as <3 or 3-5.5 were assigned an average score (total n=738 cases) based on the Global MSSS assigned for disease duration at both ends of the range. 

Results obtained from analyses including and excluding these individuals were identical (data not shown). Similarly, two sample t-tests and linear regression were used to compare age of onset distributions in cases grouped according to DRB1 genotype, as described above; results were adjusted for gender and country of origin. Age of onset and MSSS analyses using the entire dataset (n=2,201) or restricted to unrelated MS cases only (n=1,743) were very similar. Results obtained from unrelated MS cases are reported in the text.

EDSS scores were used in conjunction with disease duration to define mild or severe forms of MS. Disease duration was measured as the number of years between the year of onset of first symptom and year of last exam with EDSS assessment; in most cases this is at entry of study, given the cross sectional nature of these datasets. Mild disease (the ability to walk normally or have only mild gait disability) was defined as those cases with an EDSS ≤3 after 15 years. Severe disease (need for bilateral assistance to walk or wheelchair dependency), included cases who reached an EDSS >6, within 10 years of disease duration. Logistic regression models were used to determine the odds of DRB1*15 carrier status (copy number) in mild vs. severe MS case groups using PROC GENMOD (SAS v. 9.1, SAS Institute, Cary, NC). All odds ratios were adjusted for gender.
Sequence alignment and comparison of predisposing and protective HLA-DRB1 alleles to DRB1*1501 was performed using the Immunogenetics database (IMGT) webpage (www.ebi.ac.uk/imgt). Visualizations for HLA-DRB1 alleles were performed with the Chimera software (www.cgl.ucsf.edu/chimera).

Acknowledgements

We are grateful to the individuals with MS and their families for making this study possible. We thank the International MS Genetics Consortium for providing some of the HLA-DRB1 genotype data. We thank R. Gomez and C. Tong for recruitment of U.S. cases to the study and W. Chin, H. Mousavi, and R. Guerrero for sample preparation and repository management. We also thank G. Artim and J. Penko for technical support. This work was funded by grants of the National Institutes of Health, National Multiple Sclerosis Society, and Nancy Davis Foundation.

Conflict of Interest Statement

None
References


Legends to Figures

Figure 1. *HLA-DRB1*15 genotype associations with disease susceptibility. Odds Ratios (OR) and 95% confidence intervals from conditional logistic regression (CLR) analyses using *low risk* (A) and *high risk* (B) genotype reference groups. (*DRB1*X/X, where X excludes *DRB1*15 genotypes). The analyses included all affected and unaffected individuals from 1,339 MS families (total N = 4,669 individuals; 1,571 MS cases and 3,098 unaffected individuals; m:n matching) and 958 individual trio families (total N = 3,832; 958 MS cases and 2,874 pseudocontrols; 1:3 matching). Unaffected family members included all parents and siblings of affected individuals. Pseudocontrols were derived from non-transmitted parental alleles (32). All analyses were performed using CLR modeling as implemented in PROC TPHREG (SAS v. 9.1; SAS Institute, Cary, NC). All p-values for *DRB1* genotype comparisons using *low risk* *DRB1*X/X as reference group were highly significant (p<0.0001), with the exception of *DRB1*14 (p>0.25). All p-values for *DRB1* genotype comparisons using *high risk* *DRB1*15/15 as reference group were significant (ranging from p<0.0001 to p<0.02), with the exception of *DRB1*08 (p>0.30) Rare *HLA-DRB1* genotypes (< 1% frequency) were collapsed into the *DRB1*15/X (other) genotype category (*DRB1*09, 10, 12, 1502 and 16).

Figure 2. *HLA-DRB1*15 genotypes and MSSS. MS cases were assigned a Global MSSS (34). When cases were grouped according to *DRB1*15 genotype to test for differences in MSSS distributions, results were very similar across the different genotypes. When resulting median or mean MSSS for each *DRB1* genotype shown here was compared to the median or mean value derived from the *DRB1*X/X cases, the distributions were statistically indistinguishable. The
Wilcoxon Rank Sum test and t-test (*ranksum, ttest, rgress* in Stata v. 8.2, StataCorp LP, College Station, TX) was used to test for significant differences in MSSS distributions (see Methods). When analyses were restricted to unrelated cases, results were very similar (data not shown in Figure 2).

**Figure 3. HLA-DRB1*15 genotype frequencies in MS cases grouped by disease course.**

Independent cases derived from the larger family dataset were used to characterize \( DRB1 \) genotypes in PPMS (total n=87; 31 cases were from US and 56 from UK) and RRMS/SPMS (n=1,036; 544 from US and 492 from UK). Control genotype frequencies were derived using non-transmitted allele frequencies from 433 parents of US MS cases, under assumptions of Hardy-Weinberg equilibrium(36). Similar \( DRB1*15 \) genotype frequencies were observed in PPMS and RRMS/SPMS cases suggesting the presence of similar \( DRB1*15 \) related disease mechanisms across phenotypes. \( DRB1*15 \) genotype frequencies in both MS subgroups were higher compared to controls. A total of 48 PPMS families (27 complete trios and 38 discordant sib pairs) were also available for PDT analysis (PDT v. 5.1 (59); results: global p-value for \( DRB1=0.003 \); and for over-transmission of \( DRB1*15 \), specifically, p=0.0004, data not shown.

**Figure 4. Molecular modeling of HLA-DRB1 susceptibility and resistance alleles in MS.**

A. Sequence alignment of various HLA-DRB1 alleles compared to \( DRB1*1501 \). Arrowheads encompass the sequence of the floor of the groove that mostly interacts with the peptide (p56-p87). While all permissive alleles exhibit a tyrosine (Y) in position 60 (boxed), the resistance allele \( DRB1*1401 \) displays a histidine (H). The alignment was performed using the Immunogenetics database (IMGT) webpage (www.ebi.ac.uk/imgt)
B-C. Crystallographic structure of HLA-DRβ1*1501 bound to MBP peptide 85-98. (B) Side view of the complex with the HLA molecule (white) in surface view and the peptide (green) in ball and stick. Allele discriminatory position β60 (Y/H) is colored in red to highlight its proximity to the peptide, potentially altering its binding properties. Residue βA71, critical for peptide binding, is shown in yellow. The small and uncharged βA71 creates the necessary room for the binding of large hydrophobic side antigen chain in the P4 pocket. (C) Close-up view of HLA-DRβ1 position β60 showing the Y residue (red) of allele *1501, and all possible rotamers of a modeled Y-H substitution. The OH⁻ terminus of the MBP peptide T97-P98 is shown in green. In the absence of a crystal structure for β*1401, all possible rotamers of a Y-H substitution were modeled using Swiss-PDB viewer (http://ca.expasy.org/spdbv). The most thermodynamically favorable rotamers are displayed in colors, while the least favorable ones are grayed. Visualizations were performed with the software Chimera (www.cgl.ucsf.edu/chimera).
Table 1. Description of MS family datasets for *HLA-DRB1* risk and phenotype analyses

<table>
<thead>
<tr>
<th></th>
<th>US families&lt;sup&gt;1&lt;/sup&gt;</th>
<th>UK families</th>
<th>UK MS ASP&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Spanish families</th>
<th>Italian families</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total no. families</td>
<td>527</td>
<td>492</td>
<td>204</td>
<td>193</td>
<td>127</td>
</tr>
<tr>
<td>Total no. individuals</td>
<td>2,534</td>
<td>1,475</td>
<td>414</td>
<td>570</td>
<td>399</td>
</tr>
<tr>
<td>Total no. MS cases&lt;sup&gt;3&lt;/sup&gt;</td>
<td>1,063</td>
<td>492</td>
<td>414</td>
<td>196</td>
<td>129</td>
</tr>
<tr>
<td>Female:male ratio</td>
<td>3.1:1</td>
<td>3.2:1</td>
<td>2.4:1</td>
<td>1.9:1</td>
<td>1.9:1</td>
</tr>
<tr>
<td>Mean age in years (SD)</td>
<td>43.4 (10.7)</td>
<td>37.9 (7.3)</td>
<td>46.2 (11.0)</td>
<td>36.8 (10.4)</td>
<td>35.9 (9.7)</td>
</tr>
<tr>
<td>Mean age of onset, years (SD)</td>
<td>30.3 (8.7)</td>
<td>25.4 (6.7)</td>
<td>29.3 (9.2)</td>
<td>26.9 (8.8)</td>
<td>27.6 (8.3)</td>
</tr>
<tr>
<td>Mean disease duration, years (SD)</td>
<td>13.1 (9.8)</td>
<td>12.5 (7.3)</td>
<td>16.9 (10.4)</td>
<td>7.9 (7.3)</td>
<td>9.8 (7.1)</td>
</tr>
<tr>
<td>RR (RR/SP)&lt;sup&gt;4&lt;/sup&gt; n (%)</td>
<td>947 (89.1)</td>
<td>492 (100.0)</td>
<td>351 (84.8)</td>
<td>167 (85.2)</td>
<td>115 (90.6)</td>
</tr>
<tr>
<td><em>DRB1</em>&lt;sup&gt;*&lt;/sup&gt;15 n (%)</td>
<td>617 (58.0)</td>
<td>304 (61.8)</td>
<td>268 (64.7)</td>
<td>71 (36.2)</td>
<td>30 (23.3)</td>
</tr>
</tbody>
</table>

<sup>1</sup> 307 MS cases had no available parental *HLA-DRB1* data, and therefore were not used in family-based association analyses.

<sup>2</sup> 414 MS cases (affected sib pairs or ASP) had no available parental *HLA-DRB1* data, and therefore were not used in family-based association analyses.

<sup>3</sup> Complete clinical data was available for 2,201 MS cases (of total n= 2,294) or 96% of the dataset.

<sup>4</sup> RR=relapsing-remitting; SP=secondary progressive. Together these refer to a relapsing at onset disease subtype.
Table 2. P-values from pedigree disequilibrium test (PDT) analyses of *DRB1* in MS families

<table>
<thead>
<tr>
<th>Allele</th>
<th>All Families</th>
<th>Families w/o <em>DRB1</em>15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trios T:NT</td>
<td>DSP A:UA</td>
</tr>
<tr>
<td><em>DRB1</em>15</td>
<td>698:278</td>
<td>439:421</td>
</tr>
<tr>
<td><em>DRB1</em>01</td>
<td>163:248</td>
<td>78:130</td>
</tr>
<tr>
<td><em>DRB1</em>03</td>
<td>265:305</td>
<td>149:236</td>
</tr>
<tr>
<td><em>DRB1</em>04</td>
<td>295:387</td>
<td>169:215</td>
</tr>
<tr>
<td><em>DRB1</em>07</td>
<td>233:296</td>
<td>141:159</td>
</tr>
<tr>
<td><em>DRB1</em>08</td>
<td>68:62</td>
<td>41:40</td>
</tr>
<tr>
<td><em>DRB1</em>09</td>
<td>20:28</td>
<td>6:14</td>
</tr>
<tr>
<td><em>DRB1</em>10</td>
<td>10:13</td>
<td>6:5</td>
</tr>
<tr>
<td><em>DRB1</em>11</td>
<td>145:186</td>
<td>89:140</td>
</tr>
<tr>
<td><em>DRB1</em>12</td>
<td>27:29</td>
<td>12:12</td>
</tr>
<tr>
<td><em>DRB1</em>13</td>
<td>186:221</td>
<td>126:182</td>
</tr>
<tr>
<td><em>DRB1</em>14</td>
<td>25:62</td>
<td>12:44</td>
</tr>
<tr>
<td><em>DRB1</em>16</td>
<td>28:38</td>
<td>18:13</td>
</tr>
</tbody>
</table>

¹MS families used in PDT analyses are as follows: All (1,294 families comprised of 1,094 trios and 1,215 discordant sib pairs or DSP) and families w/o *DRB1*15 (494 families comprised of 394 trios and 338 DSP). The pedigree disequilibrium test (PDT v. 5.1) was used (59). Strong
association overall with DRB1 was observed in the combined MS family dataset (global test 
p=6.7 x 10^{-33}); whereas, the global test in DRB1*15 negative families was not significant 
(p=0.52).

\( ^2 \)T=transmitted, NT=non-transmitted, A=affected, and UA=unaffected.

\( ^3 \)P-values derived from the PDT for each allele are based upon analyses of all trios and DSP 
when present in MS families, and are uncorrected for multiple tests. When southern European 
families were removed from the analyses, results were very similar (data not shown). Very rare 
DRB1 alleles (< 0.5% in all MS cases) are not shown in the table.
Table 3. Summary of evidence for trans *DRB1* allele interactions and risk for MS

<table>
<thead>
<tr>
<th><em>DRB1</em> allele distributions derived from <em>DRB1</em>15/X MS cases and NT parental chromosomes&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>DRB1</em> genotype</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td><em>DRB1</em>15/08</td>
</tr>
<tr>
<td><em>DRB1</em>15/14</td>
</tr>
</tbody>
</table>

<sup>1</sup>Odds Ratios (OR) and 95% confidence intervals from CLR analyses. *DRB1*15/X was used as the reference group, where X=all other *DRB1* alleles except *08, *14 and *15. The analyses included all affected and unaffected individuals from 1,339 MS families (total N = 4,669 individuals; 1,571 MS cases and 3,098 unaffected individuals, shown in Table 3) and 958 individual trio families (total N = 3,832; 958 MS cases and 2,874 pseudocontrols); results from both comparisons were very similar. Unaffected family members included all parents and siblings of affected individuals. Pseudocontrols were derived...
from non-transmitted parental alleles (one case per family) (32). All analyses were performed using CLR modeling as implemented in PROC TPHREG (SAS v. 9.1; SAS Institute, Cary, NC).

2One case was derived from each independent trio family. Non-transmitted parental control alleles were identified using AFBAC (affected family based control) software (36). A total of 461 ‘X’ alleles from DRB1*15/X MS cases were compared to 1,764 non-transmitted parental control ‘X’ alleles, where X=all non-DRB1*15 alleles. P-values were determined using SAS (v. 9.1; SAS Institute, Cary, NC).

3Considering transmission from non-DRB1*15 parents (present in 350 trio families) to DRB1*15 positive MS cases also revealed evidence for a protective DRB1*14 effect, even in the presence of DRB1*15. Specifically, the rare DRB1*14 allele was transmitted less often from non-DRB1*15 parents to DRB1*15 MS cases (T:NT = 1:10, p=0.01 and was derived using PDT v. 5.1 (59).
Figure 1B

The figure shows a bar chart with odds ratios for different alleles of the DRB1 gene. The chart compares CASE/PSEUDO and ALL categories. Each bar represents an allele, and the height of the bar indicates the odds ratio. The error bars show the variability of the data. The chart includes alleles such as DRB1*15/14, DRB1*15/13, DRB1*15/11, DRB1*15/10, and DRB1*15/06.
Figure 2.
Figure 3.

- DRB1*15/15
- DRB1*15/x
- DRB1*15+
- DRB1*x/x
**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFBAC</td>
<td>Affected Family BAse Control</td>
</tr>
<tr>
<td>APL</td>
<td>Altered Peptide Ligands</td>
</tr>
<tr>
<td>ASP</td>
<td>Affected Sib Pair</td>
</tr>
<tr>
<td>CLR</td>
<td>Conditional Logistic Regression</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>DSP</td>
<td>Discordant Sib Pair</td>
</tr>
<tr>
<td>EDSS</td>
<td>Expanded Disability Status Score</td>
</tr>
<tr>
<td>HLA</td>
<td>Human Leukocyte Antigen</td>
</tr>
<tr>
<td>MHC</td>
<td>Major Histocompatibility Complex</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnet Resonance Imaging</td>
</tr>
<tr>
<td>MS</td>
<td>Multiple Sclerosis</td>
</tr>
<tr>
<td>MSSS</td>
<td>Multiple Sclerosis Severity Score</td>
</tr>
<tr>
<td>MBP</td>
<td>Myelin Basic Protein</td>
</tr>
<tr>
<td>RRMS</td>
<td>Relapsing-Remitting Multiple Sclerosis</td>
</tr>
<tr>
<td>OR</td>
<td>Odds Ratio</td>
</tr>
<tr>
<td>PDT</td>
<td>Pedigree Disequilibrium Test</td>
</tr>
<tr>
<td>PPMS</td>
<td>Primary Progressive Multiple Sclerosis</td>
</tr>
<tr>
<td>SPMS</td>
<td>Secondary Progressive Multiple Sclerosis</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>US</td>
<td>United States</td>
</tr>
</tbody>
</table>