Burden of risk variants correlates with phenotype of multiple sclerosis.

Kelly Hilven[1], Nikolaos A Patsopoulos[2,3,4], Bénédicte Dubois[1,5] and An Goris[1]

1. Laboratory for Neuroimmunology, Department of Neurosciences, Experimental Neurology, KU Leuven - University of Leuven, B-3000 Leuven, Belgium

2. Department of Neurology, Brigham & Women's Hospital, Boston, MA, USA

3. Harvard Medical School, Boston, MA, USA

4. Broad Institute, Cambridge, MA, USA

5. Department of Neurology, University Hospitals Leuven, B-3000 Leuven, Belgium

Keywords

Multiple sclerosis, genetic risk, genetic association, disease course, oligoclonal bands, IgG index, relapse rate

Corresponding author:

An Goris, Laboratory for Neuroimmunology, Department of Neurosciences, Experimental Neurology, KU Leuven - University of Leuven, Herestraat 49 bus 1022, B-3000 Leuven, Belgium.

Email: an.goris@med.kuleuven.be

Phone: +32-16-330772

Fax: +32-16-330770

Abstract

Background: More than 100 common variants underlying multiple sclerosis (MS) susceptibility have been identified, but their effect on disease phenotype is still largely unknown.

Objective: To assess whether the cumulative genetic risk score of currently known susceptibility variants affects clinical presentation.

Methods: A cumulative genetic risk score was based on four Human Leukocyte Antigen (HLA) and 106 non-HLA risk loci genotyped or imputed in 842 Belgian MS patients and 321 controls. Non-parametric analyses were applied.

Results: An increased genetic risk is observed for MS patients, including subsets such as oligoclonal band negative and primary progressive MS patients, compared to controls. Within the patient group, a stronger association between HLA risk variants and the presence of oligoclonal bands, an increased IgG index and female gender was apparent. Results suggest an association between a higher accumulation of non-HLA risk variants and increased relapse rate as well as shorter relapse-free intervals after disease onset.

Conclusion: MS patients display a significantly increased genetic risk compared to controls, irrespective of disease course or presence of oligoclonal bands. Whereas the cumulative burden of non-HLA risk variants appears to be reflected in the relapses of MS patients, the HLA region influences intrathecal immunoglobulin G levels.

Introduction

Multiple sclerosis (MS) is a common neuroinflammatory disorder[1] that mainly affects young adults and causes cognitive and physical impairment. Epidemiological studies have shown the importance of both genetic and environmental factors in the development of the disease. Following the association between the human leukocyte antigen (HLA) region and MS susceptibility discovered in the early 1970s, joint efforts during the last decade have led to an enormous progress in the identification of additional genetic risk variants.[2-4] Large-scale screens in international study populations, of which the Belgian cohort was a part, have identified four classical HLA and 110 non-HLA susceptibility loci.[2-4] These variants point to an important role of the immune system in the pathogenesis of MS, with approximately one third of the associated genes having known roles in the immune system and one in five signals overlapping with at least one other autoimmune disease.[3] These variants are common in the general population and, with the exception of the HLA region (e.g. HLA-DRB1*15:01: odds ratio (OR) = 3.1),[2] individually exert modest effects (OR: 1.03-1.34).[3] Since the identification of single polymorphisms has proven not to be informative to predict disease outcome, several studies have investigated the cumulative effect of combinations of identified risk variants.[5-11] Despite results showing it is not applicable for case-control prediction,[5,6,11] significant associations with clinical presentation have been suggested.[8-10] In this study, we elaborate on these previous efforts by calculating MS risk based on the current extended list of MS risk variants[2-4] and investigate the effect on MS phenotype.

Methods

Sample recruitment

The study cohort consisted of 842 unrelated MS patients and 321 control subjects of Caucasian descent. Patients, diagnosed based on Poser[12] or McDonald[13] criteria, were recruited from the University Hospitals Leuven and the National Multiple Sclerosis Centre Melsbroek. Control individuals were spouses of neurological patients, recruited from the same population. The study was approved by the Ethics Committee of the University Hospitals Leuven and appropriate informed consent was obtained from all participants. Peripheral blood samples were taken and DNA was extracted according to standard methods. Clinical characteristics (gender, disease course, age at onset, severity, relapse rate, relapse-free interval after disease onset, oligoclonal band (OCB) status and count, immunoglobulin G (IgG) index) of the study cohort were collected and are presented in Table 1. The annualized baseline relapse rate for bout onset MS patients was calculated by dividing the number of relapses that occurred before the start of any treatment by the time between the first relapse and the initiation of treatment. In order to obtain a representative relapse rate, only patients that were followed for more than three months before treatment were included (median follow-up time was 6.46 years, ranging from 0.26 to 57.14 years). The relapse-free interval for bout onset MS patients was determined by the time between the first and second relapse that occurred before the start of any treatment. Severity is measured by the Multiple Sclerosis Severity Score (MSSS). Patients were classified as OCB positive when at least two OCBs were present in the cerebrospinal fluid (CSF) that were not seen in the serum, otherwise they were classified as OCB negative. OCB count represents the number of CSF specific OCBs. The IgG index represents the ratio of IgG in the CSF compared to serum, corrected for the presence of albumin in both compartments.

Table 1. Clinical description of the study cohort.

Clinical data	MS patients	N	Controls	N
Male : Female, %	35.51 : 64.49	842	47.66 : 52.34	321
BOMS : PPMS : unknown, %	86.34 : 11.64 : 2.02	842	-	-
AAO (y), median (IQR)	33 (25-41)	822	-	-
MSSS, median (IQR)	6.24 (2.6-8.5)	666	-	-
Relapse rate, median (IQR)	0.54 (0.21-1.12)	252	-	-
Relapse-free interval (y), median (IQR)	2 (0.83-5.08)	231	-	-
OCB positive : OCB negative, %	88.01 : 11.99	559	-	-
OCB count, median (IQR)	11 (5.75-15)	260	-	-
IgG index, median (IQR)	0.88 (0.65-1.32)	262	-	-

BOMS: bout onset MS; PPMS: primary progressive MS; AAO: age at onset; y: year; IQR: interquartile range; MSSS: MS severity score; OCB: oligoclonal band; IgG index: immunoglobulin G index; N: number of individuals.

Single nucleotide polymorphism (SNP) genotyping and imputation

Cases and controls were genotyped as part of previously published studies.[2,3] In short, for all controls and 309 MS patients, 109/110 non-HLA variants were directly genotyped on the Immunochip microarray (Illumina), previously described in detail.[3] For 538 MS patients genotyped on the Human660-Quad chip (Illumina) as part of the 2011 genome-wide association study (GWAS),[2] 53/110 SNPs were directly genotyped and all other risk variants were imputed with the 1000 Genomes Project European phase I (a) panel using BEAGLE[14] as part of the replication phase of the Immunochip project.[3] Five case samples were genotyped on both platforms and served as an internal control. Variants that were not genotyped or imputed in the GWAS dataset were accordingly replaced by the best proxy ($r^2>0.98$; Appendix 1) that was also present in the Immunochip dataset, in order to obtain a uniform set of SNPs for MS risk calculation. Based on the post-imputation genotype probabilities, genotypes were assigned if the probability was higher than 0.6, otherwise they were set to missing. SNPs not available for both platforms (rs2150702 in *MLANA*) or with genotyping call rate below 95% (rs2256814 in *SLC2A4RG*, rs4976646 in *RGS14*, rs716719 (proxy of rs201847125; $r^2=1$) intergenic between *C7orf72* and *IKZF1*) were excluded, leaving 106 non-HLA risk variants for MS risk calculation (Appendix 1). All variants passed Hardy-Weinberg

Equilibrium testing and total genotyping rate in all individuals was 99.83% for the remaining non-HLA SNPs.

HLA genotypes were imputed with HLA*IMP[15] as part of the original studies (previously described in detail).[2,3] Again, genotypes were assigned when post-imputation genotype probabilities were above 0.6 and set to missing if probabilities were below this threshold.

MS risk calculation

MS risk calculation was based on four classical HLA loci[2] and 106 of the 110[3] known non-HLA susceptibility variants (Appendix 1). To calculate the cumulative genetic risk score, corrected odds ratio (OR_C) values relative to the population lifetime odds for each genotype were determined as previously described in detail by Sawcer and Wason.[16]

The OR_C was calculated by taking into account the frequencies of the risk allele B (frequency: p) and the alternate allele b (frequency: q=1-p). OR_C for genotypes BB, Bb, bb is given by

$$OR_{BB} = \frac{OR^2}{D}; \quad OR_{Bb} = \frac{OR}{D}; \quad OR_{bb} = \frac{1}{D}$$

where $D = OR^2 p^2 + OR*2pq + q^2$ and represents the odds ratio between the odds of disease in the population as a whole and the odds of disease in the lowest risk genotype.[16] OR and risk allele frequency (RAF) were obtained from the 2011 GWAS[2] (HLA loci) or Immunochip study[3] (non-HLA loci). RAF were based on the frequencies observed in the UK control cohort,[2,3] since this was the largest and most closely related control population available. OR point estimates were obtained from these previous studies[2,3] and were determined based on the entire study population. The cumulative MS risk score was calculated per sample by summing up the natural logarithm of the corrected odds ratio values for each genotype for each of the four HLA and 106 non-HLA risk loci. To assess the effect of the HLA region, MS risk was calculated based on all risk variants combined or the HLA and non-HLA variants separately.

$$\text{Cumulative MS risk score} = \sum_{n=1}^{4} \ln(OR_{C(HLA)})_n + \sum_{n=1}^{106} \ln(OR_{C(non-HLA)})_n$$

In the rare event that an individual's genotype for a SNP was missing, the score for that SNP was set to zero. Next, the exponent of this summed score was multiplied by the risk of disease in all individuals prior to genotyping, i.e. the population prevalence (0.001). This absolute risk was used to investigate the association with MS phenotype.

Statistical analysis

Statistical analysis was performed using R statistical software (version 2.15.2). Correlations between MS risk and phenotype were assessed by using the nonparametric Wilcoxon rank-sum test, Kruskal-Wallis chi-squared or Spearman's rank correlation coefficient (linear regression). For the relapse-free interval after disease onset, additionally an extremes of outcome analysis using the log-rank test was performed, allowing the inclusion of censored untreated patients without a second relapse before the end of the follow-up period. P-values are represented as p_{All} when non-HLA and HLA risk variants were combined and $p_{non-HLA}$ or p_{HLA} when MS risk was based solely on the non-HLA or HLA component, respectively. The significance threshold, corrected for multiple testing (45 tests) was set at 0.0011.

Results

The distribution of MS risk was significantly higher in MS cases compared to controls (Figure 1(a) and (b)), either when MS risk calculation was based on all risk variants combined or when non-HLA and HLA risk variants were regarded separately ($p<10^{-15}$) (Table 2). When including all risk variants, 3.68% of cases showed an absolute risk above 1%, compared to only 0.31% of controls.

Next, we classified patients according to disease course and observed no difference in MS risk in patients with bout onset (BOMS) compared to primary progressive MS (PPMS) (Table 2; Figure 1(c) and 1(d)). Both groups of patients still presented with significantly higher genetic risk scores compared to the control group based on all risk variants (BOMS versus controls: $p_{All}<10^{-15}$; PPMS versus controls: $p_{All}=5.45*10^{-15}$) as well as non-HLA (BOMS versus controls: $p_{non-HLA}<10^{-15}$; PPMS versus controls: $p_{non-HLA}=4*10^{-9}$) or HLA risk variants alone (BOMS versus controls: $p_{HLA}<10^{-15}$; PPMS versus controls: $p_{HLA}=3.097*10^{-7}$).
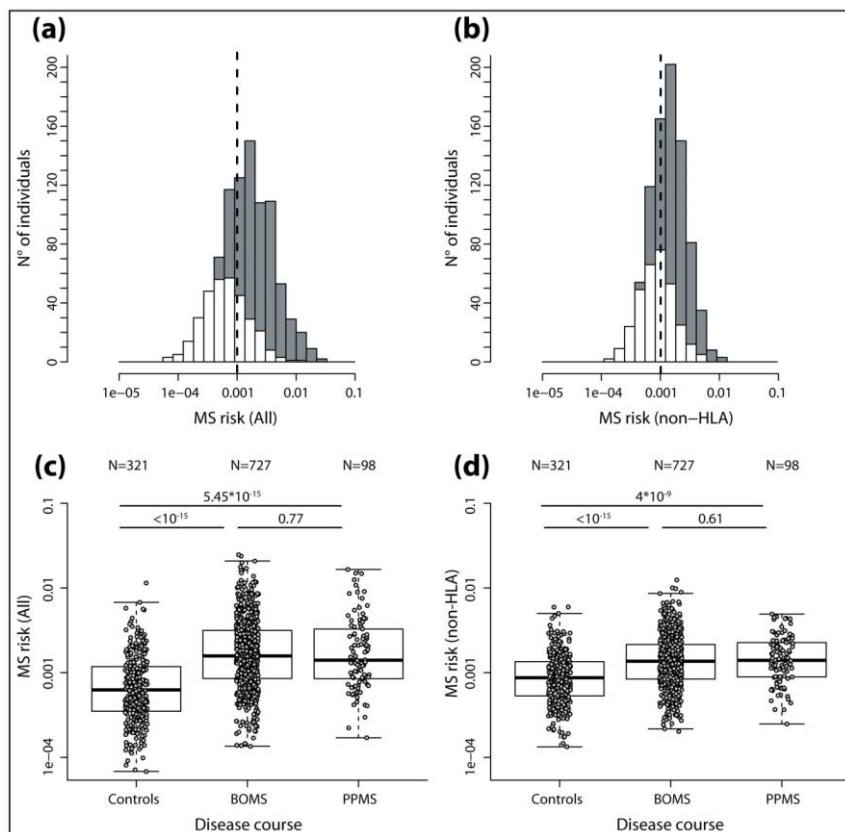
**Figure 1.** Distribution of MS risk according to disease course.

(a) and (b) Histograms representing MS risk distribution in patients (grey) and control subjects (white) when (a) including all or (b) only non-HLA loci in MS risk calculation. MS patients have higher cumulative MS risk scores compared to controls. The population prevalence of disease (0.001) is indicated by the dotted vertical line.

(c) and (d) Boxplots representing the MS risk in patients, classified according to disease course, and controls based on (c) all risk variants or (d) non-HLA alone. Non-parametric statistical testing confirmed that MS risk in bout onset and primary progressive MS patients did not differ, but both groups showed significantly higher MS risk scores compared to controls. P-values from the Wilcoxon rank-sum test are reported.

MS risk is plotted on a logarithmic axis.

BOMS: bout onset MS; PPMS: primary progressive MS; N: number of individuals.

The presence of OCBs in the CSF of MS patients showed, compared to absence of OCBs, a trend for association with a higher MS risk score based on all known risk variants (p_{All}=0.0050) (Figure 2(a); Table 2). This appears due in particular to the HLA region (p_{HLA}=0.012). Despite OCB negative patients having a lower genetic risk compared to OCB positive patients, both groups still presented with a significantly heightened genetic burden compared to controls either overall (OCB+ versus controls: $p_{All}<10^{-15}$; OCB- versus controls: $p_{All}=2.77*10^{-5}$), for non-HLA (OCB+ versus controls: $p_{non-HLA}<10^{-15}$; OCB- versus controls: $p_{non-HLA}=5.19*10^{-4}$), and HLA risk variants (OCB+ versus controls: $p_{HLA}<10^{-15}$; OCB- versus controls: $p_{HLA}=0.0018$). The number of OCBs present in the CSF, but not in the serum of an MS patient, showed a trend for association with MS risk based on HLA risk variants as well (p_{HLA}=0.0028) (Table 3). Similarly, we analyzed the quantitative measure of intrathecal IgG synthesis and found that this IgG index is significantly associated with MS risk when only the HLA component is included ($p_{HLA}=3.56*10^{-6}$), but not when non-HLA risk variants are taken into account (Figure 2(c) and (d); Table 3).
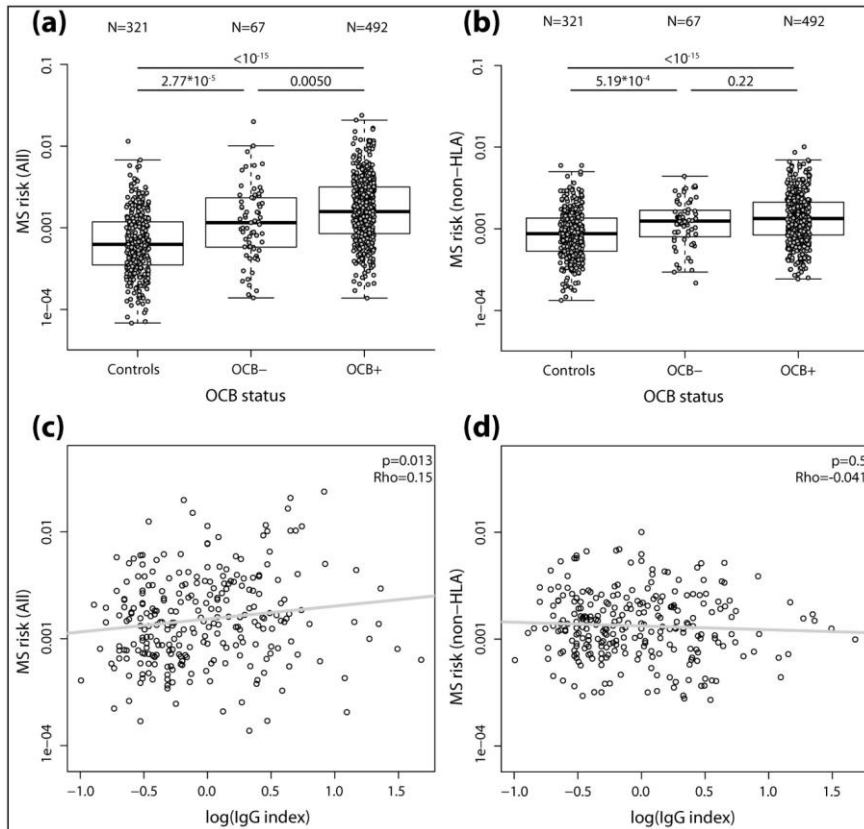
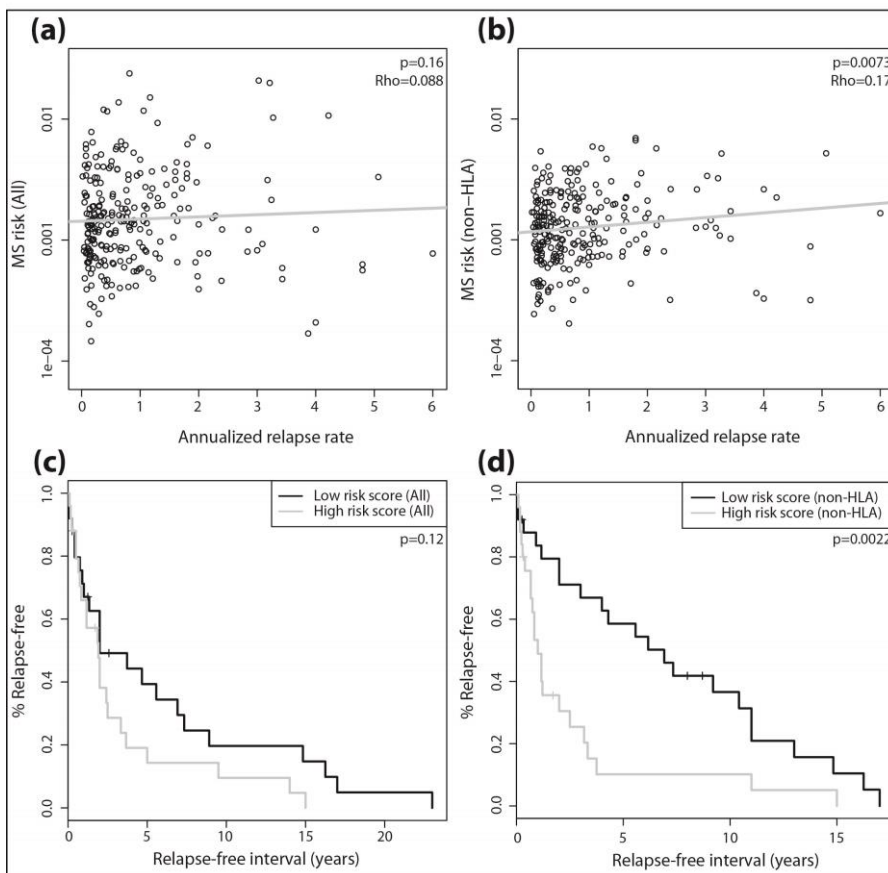**Figure 2.** Association between MS risk and CSF characteristics.

(a) and (b) Boxplots representing MS risk in patients, classified according to OCB status, and controls based on (a) all risk variants or (b) non-HLA risk variants alone. MS patients display higher cumulative MS risk scores compared to controls, irrespective of OCB status. OCB positive patients display a higher MS risk score compared to OCB negative patients only when HLA is included in the risk model. P-values from the Wilcoxon rank-sum test are reported.

OCB+: OCB positive MS patients; OCB-: OCB negative MS patients; N: number of individuals.

(c) and (d) Plots showing the distribution of MS risk according to the IgG index when (c) including all or (d) only non-HLA risk variants. A higher genetic risk score is associated with a heightened IgG index when HLA risk variants are included in the model. This association is not observed when MS risk is based solely on non-HLA risk variants. Rho estimates and p-values from the Spearman's rank correlation coefficient test are reported.

MS risk and IgG index are plotted on logarithmic axes.

Analysis showed no association with relapse rate based on all risk variants (Figure 3(a)). However, the burden of non-HLA risk variants showed a trend for a higher baseline relapse rate ($p_{non-HLA}$=0.0073) (Figure 3(b), Table 3) and was significantly associated with a shorter relapse-free interval after disease onset ($p_{non-HLA}$= 0.00083). Additionally, an extremes of outcome survival analysis was applied comparing patients with the 10% lowest risk scores based on non-HLA risk variants to those with the 10% highest risk scores. Whereas a median relapse-free interval of 1.00 year (95% CI: 0.75 - 3.17) was seen for the high-risk group, this interval extended to a median of 6.91 years (95% CI: 4.00 - 11.00) for the low-risk group (p=0.0022) (Figure 3(c)). This effect was no longer observed when HLA risk variants were taken into account (Figure 3(d)). No associations were identified between MS risk and age at onset or severity, measured by the Multiple Sclerosis Severity Score (MSSS) (Table 3).

Figure 3. Association between MS risk and clinical relapses.

(a) and (b) Plots showing the distribution of MS risk according to the baseline relapse rate when (a) including all or (b) only non-HLA risk variants. A higher genetic risk score is associated with an increased relapse rate when MS risk is based on non-HLA risk variants alone. Including HLA in the risk model attenuates this effect. Rho estimates and p-values from the Spearman's rank correlation test are reported.

MS risk is plotted on a logarithmic axis.

(c) and (d) Survival curves for the relapse-free interval after disease onset in MS patients with the 10% highest (grey) and 10% lowest (black) risk scores based on all risk variants (c) or non-HLA risk variants alone (d). A higher genetic risk score based on non-HLA risk variants is associated with a shortened relapse free survival after disease onset. Censored patients are indicated. P-values from the log-rank test are reported.

No association was observed between MS risk based on all risk variants and gender among MS patients or controls (Table 4). However, when analyzing the MS risk based on non-HLA or HLA risk variants separately, a trend for a higher HLA-based risk (p_{HLA}=0.0013) and lower non-HLA-based risk ($p_{non-HLA}$=0.018) was observed in female compared to male MS patients. This was mirrored by a somewhat higher frequency of the HLA-DRB1*15:01 haplotype amongst female compared to male patients (allele frequency: 32.32% versus 24.58%). Additionally, the data indicated departure from Hardy-Weinberg equilibrium for the HLA-DRB1*15:01 allele in female MS patients (p=0.035) These differences were not present in the control population .

Table 2. Statistical analysis of MS risk versus disease course and OCB status.

		MS risk (All)			MS risk (non-HLA)			MS risk (HLA)		
Disease course	N	median (IQR)	p vs HC	p vs PPMS	median (IQR)	p vs HC	p vs PPMS	median (IQR)	p vs HC	p vs PPMS
All patients	842	0.0016 (0.00085-0.0032)	$< 10^{-15}$	-	0.0014 (0.00085-0.0022)	$< 10^{-15}$	-	0.0014 (0.00064-0.0020)	$< 10^{-15}$	-
BOMS	727	0.0016 (0.00086-0.0032)	$< 10^{-15}$	0.77	0.0014 (0.00085-0.0022)	$< 10^{-15}$	0.61	0.0014 (0.00064-0.0020)	$< 10^{-15}$	0.48
PPMS	98	0.0014 (0.00086-0.0032)	$5.45*10^{-15}$	-	0.0014 (0.00090- 0.0023)	$4*10^{-9}$	-	0.0014 (0.00064-0.0020)	$3.097*10^{-7}$	-
HC	321	0.00063 (0.00035-0.0012)	-	-	0.00088 (0.00054-0.0014)	-	-	0.00064 (0.00047-0.0011)	-	-
OCB status	N	median (IQR)	p vs HC	p vs OCB-	median (IQR)	p vs HC	p vs OCB-	median (IQR)	p vs HC	p vs OCB-
OCB+	492	0.0016 (0.00086-0.0032)	$< 10^{-15}$	*0.0050*	0.0013 (0.00085-0.0021)	$< 10^{-15}$	0.22	0.0014 (0.00064-0.0020)	$< 10^{-15}$	*0.012*
OCB-	67	0.0012 (0.00059-0.0023)	$2.77*10^{-5}$	-	0.0012 (0.00080-0.0017)	$5.19*10^{-4}$	-	0.00081 (0.00064-0.0020)	*0.0018*	-
HC	321	0.00063 (0.00035-0.0012)	-	-	0.00088 (0.00054-0.0014)	-	-	0.00064 (0.00047-0.0011)	-	-

Table representing the mean MS risk scores and p-values obtained after statistical analyses (Wilcoxon rank-sum test) of MS risk versus disease course (top part of the table) or OCB status (bottom part of the table). To assess the effect of the HLA region, analyses were performed based on all risk variants (left), only non-HLA risk variants (middle) or only HLA risk variants (right); corresponding p-values are depicted as p_{All}, $p_{non-HLA}$ and p_{HLA}, respectively. Columns 4, 7 and 10 represent p-values when comparing patient groups with healthy controls, columns 5, 8 and 11 represent p-values when comparing bout onset with primary progressive MS (top part) or OCB positive with OCB negative MS patients (bottom part). Significant p-values are underlined, p-values depicting a trend for association are indicated in italic.

[a] The group 'All patients' includes both patients with bout onset and primary progressive disease course as well as patients for whom exact disease course is unknown (N=17).

BOMS: bout onset MS; PPMS: primary progressive MS; HC: healthy controls; OCB+: OCB positive MS patients; OCB-: OCB negative MS patients; N: number of individuals; IQR: interquartile range.

Table 3. Statistical analysis of MS risk and continuous variables.

Variables	p_{All}	Rho_{All}	$p_{non\text{-}HLA}$	$Rho_{non\text{-}HLA}$	p_{HLA}	Rho_{HLA}
MS risk vs OCB count	0.069	0.11	0.34	-0.060	*0.0028*	0.18
MS risk vs IgG index	*0.013*	0.15	0.50	-0.041	<u>$3.56*10^{-6}$</u>	0.28
MS risk vs relapse rate	0.16	0.088	*0.0073*	0.17	0.44	-0.05
MS risk vs relapse-free interval	0.25	-0.076	<u>0.00083</u>	-0.22	0.092	0.11
MS risk vs AAO	0.34	-0.033	0.70	-0.014	0.30	-0.036
MS risk vs MSSS	0.18	0.053	0.35	0.036	0.49	0.027

Table representing p-values obtained after statistical analyses (Spearman's rank correlation) of the variables depicted in column 1. To assess the effect of the HLA region, analyses were performed including all risk variants (p_{All}, Rho_{All}), only non-HLA ($p_{non\text{-}HLA}$, $Rho_{non\text{-}HLA}$) or only HLA risk variants (p_{HLA}, Rho_{HLA}). Significant p-values are underlined; p-values depicting a trend for association are indicated in italic.

OCB count: number of oligoclonal bands present in the cerebrospinal fluid, but not in the serum of an MS patient; IgG index: immunoglobulin G index; AAO: age at onset; MSSS: MS severity score; Rho: Spearman's rank correlation coefficient.

Table 4. Comparison of MS risk between men and women.

Gender patients	N	MS risk (All) Median (IQR)	p	MS risk (non-HLA) Median (IQR)	p	MS risk (HLA) Median (IQR)	p
Men	299	0.0016 (0.00084-0.0031)	0.65	0.0015 (0.00089-0.0024)	*0.018*	0.00081 (0.00064-0.0020)	*0.0013*
Women	543	0.0016 (0.00086-0.0032)		0.0013 (0.00083-0.0020)		0.0014 (0.00064-0.0020)	
Gender controls	N	Median (IQR)	p	Median (IQR)	p	Median (IQR)	p
Men	153	0.00063 (0.00037-0.0012)	0.70	0.00090 (0.00050-0.0015)	0.69	0.00064 (0.00047-0.0011)	0.75
Women	168	0.00063 (0.00035-0.0011)		0.00085 (0.00054-0.0013)		0.00064 (0.00047-0.0011)	

Table representing the mean MS risk and p-values obtained after statistical analyses (Wilcoxon rank-sum test) of MS risk versus gender in MS patients (top part of the table) and controls (bottom part of the table). To assess the effect of the HLA region, analyses were performed based on all risk variants (left), only non-HLA risk variants (middle) and only HLA risk variants (right); corresponding p-values are depicted as p_{All}, $p_{non-HLA}$ and p_{HLA}, respectively. P-values depicting a trend for association are indicated in italic.

N: number of individuals IQR: interquartile range.

Discussion

In recent years, large-scale collaborations have led to the identification of four classical HLA and 110 non-HLA loci that are associated with susceptibility to MS.[2-4] Very few of these risk variants have been found to be individually associated with disease phenotype (i.e. disease course, severity, age at onset and CSF antibody production). Previous studies did show that patients carrying HLA-DRB1*15:01 are more likely to develop disease at an earlier age[2,17,18] and are more likely to be OCB positive.[19,20] This paucity of individual associations has led research groups to investigate whether the cumulative effect of these variants influences disease phenotype.[7-10] In the present study we elaborated on these studies by integrating currently known MS susceptibility loci[2,3] in the risk model. We applied a weighted genetic risk score to account for the different effect sizes. The model was based on that provided by Sawcer and Wason,[16] including allele frequencies and disease prevalence in order to estimate an individual's genetic risk to develop MS. The distribution of MS risk shows that even among MS patients, the genetic risk never exceeds 10% and only 3.68% of cases have risks above 1% when including all MS susceptibility loci. This underlines the findings from previous studies[5,6,11] indicating that the cumulative genetic risk score is not applicable to predict whether or not someone will develop MS.

The clinical course of MS can be classified as bout onset MS, i.e. relapsing-remitting or secondary progressive MS, or primary progressive MS.[1] Progressive onset is known to be associated with a later age at disease onset and an increased male to female ratio compared to bout onset.[21] We observed an increased MS risk in patients compared to controls, but not between patients with a different disease course. This is in agreement with previous studies based on a smaller number of risk loci in which no difference in MS risk was observed between bout onset and primary progressive MS patients,[2,7-10] except for a difference between these disease course groups in male patients only in one study.[10] Our findings also correspond

to what is seen in family studies, where it was shown that affected family members do not necessarily have the same clinical course,[22] but do present with similar genetic risk scores.[7,8]

A pathological hallmark of MS is intrathecal IgG synthesis, which can be measured in a qualitative (OCB status) or quantitative (IgG index) way. In this study we observed an increased MS risk in OCB positive compared to OCB negative patients and a positive association with IgG index. These associations were mainly driven by the HLA region. This is in line with previous results on a smaller set of risk variants[9,20] and with the reported association of HLA-DRB1*15:01 with antibody levels in the CSF[19,20]. Our data indicate that most of the genetic risk factors influence disease susceptibility without affecting intrathecal IgG levels. The contrary is also possible, with the IGHC locus which is found to be associated with the IgG index, without any association with MS susceptibility.[23]

The aforementioned results indicate that groups that were previously seen as possibly different MS subgroups, namely PPMS and OCB negative MS, share the same genetic risk variants as bout onset and OCB positive MS, with the exception of HLA-DRB1*15:01 that is more strongly associated with OCB positive MS.

Although not much is known about the effect of genetic variation on relapse rate in MS patients,[24-26] previous studies have shown a negative correlation with age at onset and disease duration[27,28] whereas no consistent effect on disability progression could be determined.[29] Our results suggest that a higher MS risk based on non-HLA risk variants is associated both with a higher relapse rate and a shorter relapse-free interval after disease onset. In agreement with previous studies that failed to show an association between relapse rate or the relapse-free interval after disease onset and HLA-DRB1*15:01,[25,26,30] this effect was no longer significant when HLA was included in the risk calculation. These results suggest a role for non-HLA risk variants in the inflammatory processes underlying clinical relapses. In the current study, we included patients who were followed up for a minimum of three months. This may have

limitations for the exact estimation of the relapse rate, but has the advantage to rank patients with high activity, needing early treatment, high in our non-parametric statistical test.

In the current study no association was observed between the cumulative genetic risk of MS and age at disease onset. This extends the studies by Gourraud et al.[7] and van Pelt et al..[31] The latter observed a comparable genetic risk score, based on 57 non-HLA risk variants and HLA-DRB1*15:01, in pediatric and adult onset MS. In contrast, the study of Harbo et al.[9] did observe an earlier age at onset in patients displaying a higher MS risk score, although this finding mainly reflected the known association of HLA-DRB1*15:01 with lower age at onset.[2]

The Multiple Sclerosis Severity Score[32] (MSSS) was preferred over the Expanded Disability Status Scale (EDSS) to measure severity, since it corrects EDSS for disease duration. In accordance with previous studies,[7,9] no association was observed in the present study. This seems to indicate that genetic risk variants that influence susceptibility to MS, do not affect severity.

We observed a slightly higher burden of non-HLA risk variants in male and of HLA risk variants in female patients. These effects were not present in the control population. Since women tend to have an increased intrinsic susceptibility to MS, it has been suggested that men may require a greater number of risk factors to cross the clinical threshold to develop disease.[33,34] This has been demonstrated by an increased transmission of MS from affected fathers to their children compared to mothers, but this parent of origin effect has been subject of debate with different studies showing opposing results.[33-35] The increased HLA-based risk in female compared to male MS patients is due to the higher frequency of HLA-DRB1*15:01 in female (32.32%) compared to male (24.58%) patients observed in our study population. This is in agreement with a longstanding discussion in the literature, reporting an increased frequency of HLA-DRB1*15:01 in women,[2,17,36-38] a higher transmission of HLA-

DRB1*15:01 exclusively in female-female patient pairs[37] or a departure of Hardy-Weinberg equilibrium in female but not male patients.[2]

In this study, we included four HLA and 106 non-HLA risk variants for MS in order to calculate the cumulative genetic risk score and assess the relationship with disease phenotype. No terms for interaction or environmental risk factors were included, since knowledge of interactions in MS is limited and exposure to environmental risk factors is difficult to record in retrospective studies.

In conclusion, a cumulative genetic risk score based on known MS risk variants is increased in MS patients, irrespective of disease course or OCB status, compared to control subjects. Whereas a high accumulation of non-HLA risk variants is associated with an increased relapse rate and shortened relapse-free interval after disease onset, OCB status and IgG index seem to be influenced mainly by the presence of HLA risk loci. We propose that the combined effect of non-HLA risk variants favours inflammation that is associated with relapses. HLA risk variants on the other hand, are often disease-specific and may drive an inflammatory reaction to a particular target. The associations may aid us to further explore the role of genetic risk variants in clinical expression of MS.

Acknowledgments

Conflict of Interest Statement

 None declared.

Funding

References

1. Compston A and Coles A. Multiple sclerosis. *Lancet*. 2008; 372: 1502-1517.

2. International Multiple Sclerosis Genetics Consortium & Wellcome Trust Case Control Consortium 2. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature*. 2011; 476: 214-219.

3. International Multiple Sclerosis Genetics Consortium. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat Genet*. 2013; 45: 1353-1360.

4. Sawcer S, Franklin RJ and Ban M. Multiple sclerosis genetics. *Lancet Neurol*. 2014; 13: 700-709.

5. De Jager PL, Chibnik LB, Cui J, et al. Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: a weighted genetic risk score. *Lancet Neurol*. 2009; 8: 1111-1119.

6. Sawcer S, Ban M, Wason J, et al. What role for genetics in the prediction of multiple sclerosis? *Ann Neurol*. 2010; 67: 3-10.

7. Gourraud PA, McElroy JP, Caillier SJ, et al. Aggregation of multiple sclerosis genetic risk variants in multiple and single case families. *Ann Neurol*. 2011; 69: 65-74.

8. Isobe N, Damotte V, Lo Re V, et al. Genetic burden in multiple sclerosis families. *Genes Immun*. 2013; 14: 434-440.

9. Harbo HF, Isobe N, Berg-Hansen P, et al. Oligoclonal bands and age at onset correlate with genetic risk score in multiple sclerosis. *Mult Scler*. 2013; 20: 660-668.

10. Sorosina M, Brambilla P, Clarelli F, et al. Genetic burden of common variants in progressive and bout-onset multiple sclerosis. *Mult Scler*. 2013: Epub ahead of print.

11. Disanto G, Dobson R, Pakpoor J, et al. The refinement of genetic predictors of multiple sclerosis. *PLoS One*. 2014; 9: e96578.

12.   Poser CM, Paty DW, Scheinberg L, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol*. 1983; 13: 227-231.

13.   McDonald WI, Compston A, Edan G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol*. 2001; 50: 121-127.

14.   Browning BL and Yu Z. Simultaneous genotype calling and haplotype phasing improves genotype accuracy and reduces false-positive associations for genome-wide association studies. *Am J Hum Genet*. 2009; 85: 847-861.

15.   Dilthey AT, Moutsianas L, Leslie S, et al. HLA*IMP--an integrated framework for imputing classical HLA alleles from SNP genotypes. *Bioinformatics*. 2011; 27: 968-972.

16.   Sawcer S and Wason J. Risk in complex genetics: "all models are wrong but some are useful". *Ann Neurol*. 2012; 72: 502-509.

17.   Hensiek AE, Sawcer SJ, Feakes R, et al. HLA-DR 15 is associated with female sex and younger age at diagnosis in multiple sclerosis. *J Neurol Neurosurg Psychiatry*. 2002; 72: 184-187.

18.   Smestad C, Brynedal B, Jonasdottir G, 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: 835-840.

19.   Mero IL, Gustavsen MW, Saether HS, et al. Oligoclonal band status in Scandinavian multiple sclerosis patients is associated with specific genetic risk alleles. *PLoS One*. 2013; 8: e58352.

20.   Leone MA, Barizzone N, Esposito F, et al. Association of genetic markers with CSF oligoclonal bands in multiple sclerosis patients. *PLoS One*. 2013; 8: e64408.

21.   Ebers GC. Natural history of primary progressive multiple sclerosis. *Mult Scler*. 2004; 10 Suppl 1: S8-13; discussion S13-15.

22.   Hensiek AE, Seaman SR, Barcellos LF, et al. Familial effects on the clinical course of multiple sclerosis. *Neurology*. 2007; 68: 376-383.

23.   Buck D, Albrecht E, Aslam M, et al. Genetic variants in the immunoglobulin heavy chain locus are associated with the IgG index in multiple sclerosis. *Ann Neurol*. 2013; 73: 86-94.

24.   Lin R, Taylor BV, Simpson S, Jr., et al. Association between multiple sclerosis risk-associated SNPs and relapse and disability - a prospective cohort study. *Mult Scler*. 2014; 20: 313-321.

25.   Mowry EM, Carey RF, Blasco MR, et al. Multiple sclerosis susceptibility genes: associations with relapse severity and recovery. *PLoS One*. 2013; 8: e75416.

26.   Balnyte R, Rastenyte D, Vaitkus A, et al. The importance of HLA DRB1 gene allele to clinical features and disability in patients with multiple sclerosis in Lithuania. *BMC Neurol*. 2013; 13: 77.

27.   Held U, Heigenhauser L, Shang C, et al. Predictors of relapse rate in MS clinical trials. *Neurology*. 2005; 65: 1769-1773.

28.   Tremlett H, Zhao Y, Joseph J, et al. Relapses in multiple sclerosis are age- and time-dependent. *J Neurol Neurosurg Psychiatry*. 2008; 79: 1368-1374.

29.   Young PJ, Lederer C, Eder K, et al. Relapses and subsequent worsening of disability in relapsing-remitting multiple sclerosis. *Neurology*. 2006; 67: 804-808.

30.   Van der Walt A, Stankovich J, Bahlo M, et al. Heterogeneity at the HLA-DRB1 allelic variation locus does not influence multiple sclerosis disease severity, brain atrophy or cognition. *Mult Scler*. 2011; 17: 344-352.

31.   van Pelt ED, Mescheriakova JY, Makhani N, et al. Risk genes associated with pediatric-onset MS but not with monophasic acquired CNS demyelination. *Neurology*. 2013; 81: 1996-2001.

32.   Roxburgh RHSR, Seaman SR, Masterman T, et al. Multiple sclerosis severity score - Using disability and disease duration to rate disease severity. *Neurology*. 2005; 64: 1144-1151.

33.   Kantarci OH, Barcellos LF, Atkinson EJ, et al. Men transmit MS more often to their children vs women: the Carter effect. *Neurology*. 2006; 67: 305-310.

34.   Sawcer S. The complex genetics of multiple sclerosis: pitfalls and prospects. *Brain*. 2008; 131: 3118-3131.

35.   Herrera BM, Ramagopalan SV, Orton S, et al. Parental transmission of MS in a population-based Canadian cohort. *Neurology*. 2007; 69: 1208-1212.

36.   Sadovnick AD. Differential effects of genetic susceptibility factors in males and females with multiple sclerosis. *Clinical Immunology*. 2013; 149: 170-175.

37.   Chao MJ, Ramagopalan SV, Herrera BM, et al. MHC transmission: insights into gender bias in MS susceptibility. *Neurology*. 2011; 76: 242-246.

38.   Irizar H, Munoz-Culla M, Zuriarrain O, et al. HLA-DRB1*15:01 and multiple sclerosis: a female association? *Mult Scler*. 2012; 18: 569-577.

Appendix 1. Description of variants used for MS risk calculation.

Chr	SNP ID	Gene	proxy	r²	RA	RAF	OR	OR_BB	OR_Bb	OR_bb
1	rs3748817	*MMEL1*			A	0.64	1.14	1.09	0.96	0.84
1	rs3007421	*PLEKHG5*			A	0.12	1.12	1.22	1.09	0.97
1	rs12087340	*BCL10, DDAH1*			A	0.09	1.22	1.43	1.17	0.96
1	rs11587876	*DDAH1*			A	0.79	1.12	1.05	0.93	0.83
1	rs41286801	*EVI5*			A	0.14	1.20	1.36	1.13	0.94
1	rs7552544	*VCAM1, EXTL2*			A	0.56	1.08	1.07	0.99	0.92
1	rs11581062	*VCAM1*			G	0.29	1.05	1.07	1.02	0.97
1	rs6677309	*CD58*			A	0.88	1.34	1.06	0.79	0.59
1	rs666930	*PHGDH*			G	0.53	1.09	1.08	0.99	0.91
1	rs2050568	*FCRL1*			G	0.53	1.08	1.07	0.99	0.92
1	rs35967351	*SLAMF7*			A	0.67	1.09	1.06	0.97	0.89
1	rs1359062	*RGS21, RGS1*			C	0.82	1.18	1.06	0.90	0.76
1	rs55838263	*C1orf106*			A	0.71	1.12	1.07	0.95	0.85
2	rs4665719	*CENPO*			G	0.25	1.09	1.14	1.04	0.96
2	rs2163226	*HAAO, ZFP36L2*			A	0.71	1.10	1.05	0.96	0.87
2	rs842639	*FLJ16341*			A	0.65	1.11	1.07	0.97	0.87
2	rs7595717	*CNRIP1, PLEK*			A	0.26	1.10	1.15	1.04	0.95
2	rs17174870	*MERTK*			G	0.76	1.03	1.01	0.98	0.96
2	rs9967792	*STAT4*			G	0.62	1.11	1.08	0.97	0.88
2	rs9989735	*SP140*			C	0.18	1.17	1.29	1.10	0.94
3	rs11719975	*SATB1, KCNH8*			C	0.27	1.09	1.13	1.04	0.95
3	rs2371108	*EOMES*			A	0.38	1.08	1.10	1.02	0.94
3	rs1813375	*EOMES, CMC1*			A	0.47	1.15	1.15	1.00	0.87
3	rs4679081	*CCR4, GLB1*			G	0.52	1.08	1.07	1.00	0.92
3	rs9828629	*FOXP1*			G	0.62	1.08	1.06	0.98	0.91
3	rs2028597	*CBLB*			G	0.92	1.04	1.01	0.97	0.93
3	rs1131265	*TIMMDC1*			C	0.80	1.19	1.07	0.90	0.75
3	rs1920296	*IQCB1*			C	0.64	1.14	1.09	0.96	0.84
3	rs2255214	*ILDR1, CD86*			C	0.52	1.11	1.10	0.99	0.90

3	rs9282641	CD86			G	0.92	1.12	1.02	0.91	0.81
3	rs1014486	IQCJ-SCHIP1, IL12A			G	0.43	1.11	1.12	1.01	0.91
4	rs7665090	NFKB1, MANBA			G	0.52	1.08	1.08	1.00	0.92
4	rs2726518	TET2			C	0.55	1.09	1.08	0.99	0.91
5	rs6881706	IL7R, CAPSL			C	0.72	1.12	1.06	0.95	0.85
5	rs6880778	DAB2, PTGER4			G	0.60	1.10	1.08	0.98	0.89
5	rs71624119	ANKRD55			G	0.76	1.12	1.05	0.94	0.84
5	rs756699	VDAC1, TCF7			A	0.87	1.12	1.03	0.92	0.82
5	imm_5_141486748[a]	NDFIP1	rs6874308	0.99	A[b]	0.61	1.07	1.05	0.98	0.92
5	rs2546890	LOC285626			A	0.52	1.06	1.06	1.00	0.94
6	rs17119	CD83, JARID2			A	0.81	1.11	1.04	0.94	0.84
6	rs941816	PXT1			G	0.18	1.13	1.22	1.08	0.95
6	rs72928038	BACH2			A	0.17	1.11	1.19	1.07	0.96
6	rs802734	THEMIS, PTPRK			A	0.69	1.03	1.02	0.99	0.96
6	rs11154801	AHI1			A	0.37	1.11	1.14	1.03	0.92
6	rs17066096	IL20RA, IL22RA2			G	0.23	1.14	1.22	1.07	0.94
6	rs7769192	OLIG3, TNFAIP3			G	0.55	1.08	1.07	0.99	0.92
6	rs67297943	TNFAIP3, PERP			A	0.78	1.12	1.05	0.94	0.84
6	rs212405	TAGAP, FNDC1			T	0.62	1.15	1.11	0.96	0.84
7	rs1843938	CARD11, SDK1			A	0.44	1.08	1.09	1.01	0.93
7	rs706015	SKAP2, HOXA1			C	0.18	1.14	1.24	1.08	0.95
7	rs917116	JAZF1			C	0.20	1.12	1.20	1.07	0.95
7	rs60600003	ELMO1			C	0.10	1.16	1.30	1.12	0.97
7	rs354033	ZNF767			G	0.74	1.03	1.02	0.99	0.96
8	rs1021156	PKIA, ZC2HC1A			A	0.24	1.12	1.18	1.06	0.94
8	rs2456449	PCAT1, POU5F1B			G	0.36	1.10	1.13	1.02	0.93
8	rs4410871	MIR1204, PVT1			G	0.72	1.12	1.06	0.95	0.85
8	rs759648	PVT1, MIR1208			C	0.31	1.09	1.13	1.03	0.95
10	rs2104286	IL2RA			A	0.72	1.21	1.10	0.91	0.75
10	rs793108	ZNF438, ZEB1-AS1			A	0.50	1.09	1.09	1.00	0.92
10	rs2688608	CAMK2G, C10orf55			A	0.55	1.07	1.06	0.99	0.93

10	rs1782645	*ZMIZ1*			A	0.43	1.09	1.10	1.01	0.93
10	rs7923837	*HHEX, EXOC6*			G	0.61	1.11	1.08	0.97	0.88
11	rs7120737	*AGBL2*			G	0.15	1.13	1.23	1.09	0.96
11	rs34383631	*CD6, CD5*			A	0.40	1.11	1.13	1.02	0.92
11	rs694739	*PRDX5, CCDC88B*			A	0.62	1.08	1.06	0.98	0.91
11	rs533646	*TREH, DDX6*			G	0.68	1.10	1.06	0.96	0.88
11	rs9736016	*DDX6, CXCR5*			T	0.63	1.10	1.07	0.97	0.89
11	rs523604	*CXCR5*			A	0.53	1.09	1.08	0.99	0.91
12	rs1800693	*TNFRSF1A*			G	0.40	1.14	1.17	1.02	0.90
12	rs12296430	*LTBR, CD27-AS1*			C	0.19	1.14	1.23	1.08	0.95
12	rs11052877	*CD69*			G	0.36	1.10	1.13	1.02	0.93
12	rs201202118	*TSFM*	rs10431552	1.00	G[b]	0.67	1.14	1.09	0.95	0.84
12	rs7132277	*PITPNM2*			A	0.19	1.10	1.17	1.06	0.96
13	rs4772201	*MIR548AN, TM9SF2*			A	0.82	1.12	1.04	0.93	0.83
14	rs2236262	*ZFP36L1*			A	0.50	1.08	1.08	1.00	0.92
14	rs4903324	*JDP2, BATF*			A	0.19	1.10	1.17	1.06	0.96
14	rs74796499	*GALC*			C	0.95	1.31	1.02	0.78	0.60
14	rs12148050	*TRAF3*			A	0.35	1.08	1.10	1.02	0.95
15	rs59772922	*MORF4L1, CTSH*			A	0.83	1.11	1.03	0.93	0.84
15	rs8042861	*IQGAP1*			A	0.44	1.08	1.09	1.01	0.93
16	rs2744148	*SOX8, SSTR5-AS1*			G	0.18	1.09	1.15	1.06	0.97
16	rs12927355	*CLEC16A*			G	0.68	1.21	1.12	0.93	0.77
16	rs4780346	*CLEC16A, SOCS1*			A	0.23	1.09	1.14	1.05	0.96
16	rs6498184	*PRM1, RMI2*			G	0.81	1.15	1.05	0.91	0.79
16	rs7204270	*MAPK3, CORO1A*			G	0.50	1.09	1.09	1.00	0.91
16	rs1886700	*CDH3*			A	0.14	1.11	1.19	1.08	0.97
16	rs12149527	*WWOX*			A	0.47	1.08	1.08	1.00	0.93
16	rs7196953	*MAF, DYNLRB2*			A	0.29	1.08	1.11	1.03	0.96
16	rs35929052	*IRF8, LOC146513*			G	0.89	1.14	1.03	0.90	0.79
17	rs12946510	*GRB7, IKZF3*			A	0.47	1.08	1.08	1.00	0.93
17	rs4796791	*STAT3*			A	0.36	1.10	1.13	1.02	0.93

Chr	SNP ID	Gene	RA	RAF	OR	OR$_c$_BB	OR$_c$_Bb	OR$_c$_bb
17	rs4794058	*MRPL45P2, NPEPPS*	A	0.50	1.07	1.07	1.00	0.93
17	rs8070345	*VMP1*	A	0.45	1.14	1.15	1.01	0.88
18	rs7238078	*MALT1*	A	0.77	1.05	1.02	0.97	0.93
19	rs1077667	*TNFSF14*	G	0.79	1.16	1.06	0.92	0.79
19	rs34536443	*TYK2*	C	0.95	1.28	1.02	0.80	0.62
19	rs2288904	*SLC44A2*	G	0.77	1.14	1.06	0.93	0.82
19	rs1870071	*EPS15L1*	G	0.29	1.12	1.17	1.05	0.93
19	rs11554159	*IFI30*	G	0.73	1.15	1.07	0.93	0.81
19	rs8107548	*DKKL1*	G	0.25	1.09	1.14	1.04	0.96
20	rs4810485	*CD40*	A	0.25	1.08	1.12	1.04	0.96
20	rs17785991	*SLC9A8*	A	0.35	1.09	1.12	1.02	0.94
20	rs2248359	*CYP24A1, PFDN4*	G	0.60	1.07	1.05	0.99	0.92
20	rs6062314	*ZBTB46*	A	0.92	1.10	1.01	0.92	0.84
22	rs2283792	*MAPK1*	C	0.51	1.08	1.08	1.00	0.92
22	rs470119	*TYMP*	A	0.39	1.07	1.08	1.01	0.95
6	HLA-DRB1*15:01	*HLA-DRB1*	15:01	0.13	3.10	5.87	1.89	0.61
6	HLA-DRB1*03:01	*HLA-DRB1*	03:01	0.15	1.26	1.47	1.17	0.93
6	HLA-DRB1*13:03	*HLA-DRB1*	13:03	0.01	2.40	5.62	2.34	0.98
6	HLA-A*02:01	*HLA-A*	02:01	0.26	0.73	0.62	0.84	1.16

Table representing 106 non-HLA and 4 classical HLA loci associated with MS susceptibility, which were used to calculate MS risk.

When available rsIDs of the associated non-HLA variants are given in column 2. The SNPs are located in or near the genes listed in column 3. For intergenic SNPs, the nearest gene on each side of the variant is given.

[a] No rsID is available for imm_5_141486748, the ID for this SNP used in the Immunochip study[1] is given.

[b] In the event that a proxy was used, the risk allele given in column 6 is the risk allele of the proxy SNP. Uncorrected ORs (column 8) are obtained from the Immunochip project[1] for the non-HLA risk variants and from the 2011 GWAS[2] for the HLA region. ORs are corrected according to the formula presented in the methods section.

Chr: chromosome; SNP ID: identifier of risk single nucleotide polymorphism; RA: risk allele; RAF: risk allele frequency; OR: odds ratio, OR$_c$_BB: corrected odds ratio for being homozygous for the risk allele; OR$_c$_Bb: corrected odds ratio for being heterozygous for the risk and alternate allele; OR$_c$_bb: corrected odds ratio for being homozygous for the alternate allele.

1.   International Multiple Sclerosis Genetics Consortium. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nature genetics*. 2013; 45: 1353-60.

2.   International Multiple Sclerosis Genetics Consortium and Wellcome Trust Case Control Consortium 2. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature*. 2011; 476: 214-9.