Mult Scler OnlineFirst, published on April 8, 2010 as doi:10.1177/1352458510364633

Research Paper

Multiple Sclerosis

Multiple Sclerosis
0(00) 1–12

DOI: 10.1177/1352458510364633
msj.sagepub.com

Analysis of multiple candidate genes in association with phenotypes of multiple sclerosis

Madeleine H Sombekke[1]*, David Arteta[2]*, Mark A van de Wiel[3,4], J Bart A Crusius[5], Diego Tejedor[2], Joep Killestein[1], Antonio Martínez[2], A Salvador Peña[5], Chris H Polman[1] and Bernard MJ Uitdehaag[1,3]

Abstract

Multiple sclerosis is a heterogeneous neurological disease with varying degrees of severity. The common hypothesis is that susceptibility to multiple sclerosis and its phenotype are caused by a combination of environmental and genetic factors. The genetic part exerts its effect through several genes, each having modest effects.

We evaluated whether disease severity could be predicted by a model based on clinical data and data from a DNA chip. The DNA chip was designed containing several single nucleotide polymorphisms in 44 genes, previously described to be associated with multiple sclerosis.

A total of 605 patients with multiple sclerosis were included in this analysis, using gender, onset type and age at onset as clinical covariates. We correlated 80 single nucleotide polymorphisms to the degree of disease severity using the following three outcome measures: linear Multiple Sclerosis Severity Score, dichotomous Multiple Sclerosis Severity Score (using a cut-off point of 2.5) and time to reach Expanded Disability Status Scale score 6.

Sixty-nine single nucleotide polymorphisms were included in the analysis. No individual single nucleotide polymorphism showed a significant association; however, a combination of single nucleotide polymorphisms significantly improved the prediction of disease severity in addition to the clinical variables. In all three models the Interleukin 2 gene was included, confirming a previously reported modest effect on disease severity. The highest power was obtained using the dichotomized Multiple Sclerosis Severity Score as outcome.

Several single nucleotide polymorphisms showed their added predictive value over the clinical data in the predictive models. These results support our hypothesis that disease severity is determined by clinical variables and genetic influences (through several genes with small effects) in concert.

Keywords

genetics, multiple sclerosis, outcome, Interleukin 2

Date received: 26th November 2009; accepted: 3rd February 2010

Introduction

Multiple sclerosis (MS) is a presumed autoimmune disease affecting the central nervous system, characterized by demyelination and neurodegeneration. The clinical disease course is highly variable; some patients remain without significant functional loss for many years, while others become wheelchair-bound within a short period of time.[1] As different treatment options have become available (with varying efficacy and side-effects), the identification of patients prone to develop high disability within a short period has become highly relevant. However, predictors for future disability are scarce.

[1]Department of Neurology, VU University Medical Center, Amsterdam, The Netherlands.
[2]Progenika Biopharma, S.A., Derio, Spain.
[3]Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands.
[4]Department of Mathematics, VU University, Amsterdam, The Netherlands.
[5]Laboratory of Immunogenetics, Department of Pathology, VU University Medical Center, Amsterdam, The Netherlands.

*These authors contributed equally to this study.

Corresponding author:
MH Sombekke, Department of Neurology, VU University Medical Center, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands.
Email: m.sombekke@vumc.nl

Susceptibility and disease progression in MS is believed to be conferred by the interplay of genetic and environmental factors. The major histocompatibility complex (MHC) class II region on chromosome 6 was found to be highly associated with susceptibility to MS. In addition, a few single nucleotide polymorphisms (SNPs) showed genome-wide significance of association with small effect sizes to MS susceptibility.[2–4] The current opinion is that in addition to these genes several other genes, each exerting a small effect on susceptibility to MS, remain to be discovered.[5] Much effort has been put into unravelling the genetic and environmental influences on disease susceptibility, but less attention has been given to causes of disease variability and severity. The current hypothesis is that disease progression and other phenotypic variability of MS are influenced by genetic and non-genetic factors.[6]

Earlier studies have reported intrafamilial concordance for certain phenotypes: disease course, disease severity and age at onset.[7–11] Several genes (APOE, IL1RN, MHC class II and IL1B) have been studied in relation to disease severity; however, effect sizes were small and results were ambiguous.[12–14] Recently, alleles at the HLA-DRB1 locus have quite convincingly been shown to affect disease severity.[15,16] HLA-DRB1 alleles are also found to be associated with the development of antibodies against interferon-beta therapy and magnetic resonance imaging (MRI) parameters, possibly exerting an effect on disease course.[17–19]

Determining the genetic influence on the disease process is important to gain insight into the pathophysiological mechanisms involved, and may suggest therapeutic approaches more readily than identifying genes involved in disease susceptibility.[12] Clinical predictors of disease severity have already been identified, the most important being onset type (relapsing versus progressive), age at onset and gender. We hypothesize that a combination of genes might improve the prediction of expected disease severity over purely clinical variables, and designed a DNA chip to address this theory.

In this study we evaluated the additional prognostic value of genetic information of a DNA chip, containing a set of candidate genes, previously correlated to MS (either susceptibility or phenotypes) over available demographics and clinical characteristics, aiming to improve the prediction of the expected disease severity for future patients.

Material and methods

Study participants

A total of 605 unrelated Dutch Caucasian patients were selected retrospectively from natural history studies conducted at the MS Center at the VU University Medical Center (VUmc) in Amsterdam. The selection was based on the availability of DNA, clinical assessment of disability and the confirmed diagnosis of MS. No inclusion criteria for disability status, age, gender or onset type were applied during selection of data for analysis. This study was carried out with the approval of the Medical Ethics Committee of the VUmc and informed consent was obtained from all participants. Patients were all diagnosed with MS according to Poser or McDonald criteria.[20,21] Clinical data were collected retrospectively, including age, gender, onset type, disease course, age at onset and duration of the disease. Disability status was determined for all subjects by using Kurtzke's Expanded Disability Status Scale (EDSS).[22] These scores were not acquired during relapses.

Since no single golden standard for disease severity exists, we considered two outcomes for the patients: the Multiple Sclerosis Severity Score (MSSS) and time to reach EDSS 6, a clinically relevant endpoint, indicating that assistance for walking is required. We calculated the global MSSS from EDSS scores and disease duration at the time of patients' last visit as described earlier.[23] The global MSSS denotes the speed of disability accumulation of an individual patient compared with a large patient cohort.

Selection of SNPs

SNPs were selected based on involvement in MS pathogenesis, prognosis or response to treatment, published in the literature before July 2007. The polymorphisms were confirmed and associated to an identifier by using dbSNP database (www.ncbi.nlm.nih.gov/SNP). Nucleotide sequences for the design of allele-specific probes and PCR primers were retrieved from the SNPper database (http://snpper.chip.org/bio) and NCBI database (www.ncbi.nlm.nih.gov/SNP). Sequence-specific probes and primers were designed by using the software Primer3, which is freely available at http://frodo.wi.mit.edu/. Gene names were applied according to the HUGO Gene Nomenclature Committee.

Genotyping

Genomic DNA was isolated from anti-coagulated blood with DNAzol reagent (Molecular Research Center, Inc., Cincinnati, OH, USA). Genotyping was carried out using a newly developed low-density DNA microarray based on allele-specific probes. The design, fabrication, validation and analysis of the arrays were performed following the procedure described earlier with minor modifications.[24]

Variable selection

We employed feature selection to identify the most important and predictive features in the models to be analysed. This approach of variable filtering is based on the marginal association between each variable (SNP or clinical variable) and phenotype, as variables are typically filtered on the basis of a *p*-value cut-off from a univariate analysis.

The following outcome measures for disease severity were applied in our study: the time to reach EDSS 6 and MSSS. The MSSS outcome was employed in two ways: as a continuous outcome measure and after dichotomization, using a cut-off point of MSSS 2.5. The purpose of the latter approach was to compare relatively benign patients with more severely affected patients. For the continuous MSSS we used the non-parametric Kruskal-Wallis (KW) test. For the dichotomized MSSS, the Chi-Square test was used. For time to reach EDSS 6 groups were compared by logrank tests.

To correct for multiple testing, False Discovery Rate (FDR) according to Benjamini and Hockberg was applied.[25] The corrected number represents the expected proportion false discoveries for a given *p*-value cut-off. We used cut-off 0.05 after FDR correction.

In addition, we tested the association of the SNPs included on the chip with onset type (relapsing versus progressive onset) with a chi-squared statistic on contingency tables. *p*-values were adjusted using the FDR as described above.

Predictive models

HelixTree® software (Golden Helix, Inc., Bozeman, MT, USA) was used to calculate allelic association between different groups and deviations from Hardy–Weinberg Equilibrium (HWE). The same software calculated linkage disequilibrium among SNPs. In order to filter the SNPs included in the analyses, those SNPs that were monomorphic and those with minor allele frequency below 5% in our studied population were removed from the analysis. In addition, when a complete linkage disequilibrium among SNPs was observed ($r^2 > 0.8$) only the one with the lowest *p*-value for allelic association between SNPs and phenotypes was included in the regression model. Multivariate prognostic models were constructed for the following outcomes: continuous MSSS, dichotomous MSSS with the cut-off point of 2.5 and time to reach EDSS 6, using linear regression model, logistic regression model and Cox-regression model, respectively, using SPSS version 15.0 (SPSS Inc. Headquarters, Chicago, IL, USA) and R packages Design (Harrell, 2001) and Stats (R Development Core Team, 2008).

First, clinical variables (age at onset, gender and onset type) were included as independent variables in all models. Next, backward selection was applied to select the SNPs that contributed to the model to increase the predictive power.

The goodness of fit of the model was evaluated using the Hosmer–Lemeshow statistic, and the accuracy was assessed by calculating the bootstrap area under the Receiver Operating Characteristic (ROC) curve (AUC) with 95% confidence intervals (CI). Bootstrapping on the model's AUC was conducted 500 times using 100% random sampling by replacement, using R package Boot. To measure the impact of the SNPs and variables included in the logistic regression model of the analysed phenotypes, the sensitivity, specificity, and positive likelihood ratio (LR+ = sensitivity/(1 – specificity)) were computed by means of ROC curves. AUC–ROC curves of the models based on clinical variables alone and clinical variables plus SNPs were compared by using the method by Delong et al.[26] implemented in the software Analyse-it (Analyse-it Software Ltd. Leeds, UK). The goodness of fit of the models, based on clinical variables alone and clinical variables plus SNPs, was compared using a Likelihood Ratio Test and Akaike's Information Criteria (AIC).[27] Both analyses were performed using R statistical software.

Results

Patient characteristics

Our patient group ($n = 605$) reflects a representative MS population, with approximately 35% male and 17% primary progressive MS patients (see Table 1). Eighty-six out of 605 patients (14.2%) from the study population had MSSS values <2.5, indicating a relatively benign disease course.

SNP selection

Eighty validated polymorphisms located in 44 different genes were studied on a DNA chip (see Supplementary Table 1 for the complete list). Five SNPs were monomorphic and six had a minor allele frequency below 5%. These 11 SNPs were excluded prior to univariate analysis. Finally, 69 SNPs were included in the analyses.

Univariate analysis on SNPs in relation to severity outcome-measures

We first determined the correlations of the individual SNPs to the outcome measures: MSSS, time to reach EDSS 6 and dichotomous MSSS using MSSS 2.5 as a

Table 1. Patient characteristics. Patient demographics and clinical characteristics divided in multiple sclerosis subtypes

	All	RR	SP	PP
Total number of patients	605	310	190	105
Gender (*n*; % Male)	219 (36.2)	96 (31.0)	78 (41.1)	45 (42.9)
Mean age at onset (SD)	32.4 (9.5)	30.4 (8.0)	30.6 (8.8)	41.6 (9.5)
Mean disease duration (SD)	13.1 (8.3)	9.7 (5.9)	18.2 (9.3)	13.8 (7.6)
Median EDSS (Interquartile range)	4.0 (3.5)	3.0 (2.0)	6.5 (1.5)	6.5 (3.5)
Median MSSS (Interquartile range)	5.6 (4.7)	3.9 (3.7)	7.9 (3.7)	7.3 (3.3)
Number of patients that reached EDSS 6 (%)	234 (38.7)	27 (8.7)	147 (77.4)	60 (57.1)
Median time to reach EDSS 6 in months (Interquartile range)	103 (105)	102 (83)	114 (121)	84 (91)

EDSS, Expanded Disability Status Scale; MSSS, Multiple Sclerosis Severity Score; PP, primary progressive; RR, SD, standard deviation; SP, secondary progressive.

Table 2. Predictive model using Multiple Sclerosis Severity Score as a continuous outcome variable (linear regression). Indicator variables have been created for the categorical predictors. The reference group for gender is 'female', the reference group for onset type is 'relapsing', and the reference group for IL2 (rs2069763) is 'GG'

	Indicator variable	β-coefficient (unstandardized)	β-coefficient (standardized)	95% confidence interval for β-coefficient (unstandardized)	Significance
Age at onset		0.05	0.17	0.02,–0.07	0.0001
Gender	Male	0.51	0.09	0.09,–0.93	0.0183
Onset type	Progressive	0.94	0.13	0.34,–1.53	0.0021
IL2 (rs2069763)	GT	−0.51	−0.10	−0.95,–0.08	0.02
	TT	−1.00	−0.13	−1.65,–0.08	0.0027

cut-off point. The raw *p*-values and corrected *p*-values of the SNPs are noted in supplementary Table 2. No SNP remained significantly associated after correcting for multiple testing.

Effect of SNPs on onset type

The analysis of the effect of the different SNPs on the onset type did not reveal any significant association after correcting for multiple testing.

Multivariate prognostic models on disease severity

The linear model (MSSS) returns two non-zero coefficients for the SNPs, both originating from the IL2 SNP rs2069763, in addition to the clinical variables (see Table 2). The Cox-regression analysis returns nine non-zero coefficients for the SNPs. However, all coefficients are close to zero or the 95% CI include 1 (see Table 3).

When using the model for the dichotomous outcome of the MSSS (with a cut-off point of 2.5), three clinical covariates and six SNPs were included in the model (see Table 4). The linear regression corrected R^2 is 0.088 and the Cox regression R^2 is 0.02. The Nagelkerke R^2 for the logistic regression is 0.219, indicating a higher predictive power for the logistic regression model using the dichotomous outcome, although these are not directly comparable. Therefore, only for the dichotomous model ROC curves are obtained to test the additive value of the SNPs over the clinical data relevant to disability accumulation.

Figure 1 shows ROC curves obtained with the 'clinical model' and the 'clinical–genotypic model'. The clinical model includes age at onset, onset type and gender. While the clinical–genotypic model includes the selected SNPs (see Table 4) in addition to the clinical parameters. The curve obtained from the clinical–genotypic model shows a sensitivity of 37% with specificity of 95.3% and an LR+ of 7.9 (AUC = 0.78), whilst the clinical model had a sensitivity of 27.4% with specificity of 95.3% and an LR+ of 5.8 (AUC = 0.68).

By including genetic information (SNPs) in the model, a significant improvement in predictive power was obtained as calculated by means of the AIC and the Likelihood Ratio Test. The Likelihood Ratio Test showed that the model including SNPs fits the data significantly better than the model based on clinical

Table 3. Predictive model using survival analysis on time to reach Expanded Disability Status Scale = 6 (Cox-regression)

		β-coefficient	Odds Ratio (OR)	95% Confidence interval for OR	Significance
Age at onset		0.03	1.033	1.02–1.05	<0.001
Gender					
	Male versus female	0.28	1.32	1.01–1.73	0.046
NDUFS7 (rs2074897)					0.030
	AG vs GG	−0.05	0.96	0.69–1.33	0.786
	AA vs GG	0.36	1.44	0.99–2.08	0.054
ADAMTS14 (rs4747075)					0.001
	AG vs GG	−0.54	0.58	0.42–0.80	0.001
	AA vs GG	−0.53	0.59	0.39–0.89	0.011
FAS (rs2234978)					0.016
	CT vs CC	0.25	1.28	0.97–1.70	0.086
	TT vs CC	0.60	1.83	1.19–2.82	0.006
IL2 (rs2069762)					0.008
	TG vs TT	0.28	1.33	1.00–1.75	0.044
	GG vs TT	0.70	2.02	1.23–3.31	0.005
SPP1 (rs2853744)					
	GT vsGG	0.58	1.78	1.18–2.67	0.006

Table 4. Predictive model using dichotomous Multiple Sclerosis Severity Score (MSSS) as outcome measure (threshold is set on MSSS <2.5) (logistic regression)

		β-coefficient	Odds Ratio (OR)	95% Confidence interval for OR	Significance
Age at onset		0.05	1.05	1.02–1.08	0.004
Gender					
	Male versus female	0.70	2.02	1.14–3.57	0.015
Onset type					
	Progressive versus relapsing	1.55	4.69	1.32–16.63	0.017
NOS2 (rs1137933)					0.005
	AG vs GG	−0.63	0.53	0.32–0.89	0.016
	AA vs GG	−1.41	0.24	0.09–0.67	0.006
PITPNC1 (rs1318)					0.009
	AG vs AA	−0.81	0.45	0.27–0.75	0.002
	GG vs AA	−0.53	0.59	0.18–1.95	0.387
IL2 (rs2069763)					0.004
	GT vs GG	−0.94	0.39	0.22–0.70	0.001
	TT vs GG	−0.98	0.38	0.17–0.84	0.016
CCL5 (rs2107538)					0.062
	CT vs CC	0.71	2.04	1.12–3.70	0.020
	TT vs CC	0.38	1.47	0.38–5.67	0.576
IL1RN (rs423904)					
	CT/TT vs CC	−0.52	0.60	0.36–0.99	0.047
PNMT (rs876493)					
	GG vs AA/AG	−0.65	0.52	0.29–0.92	0.025

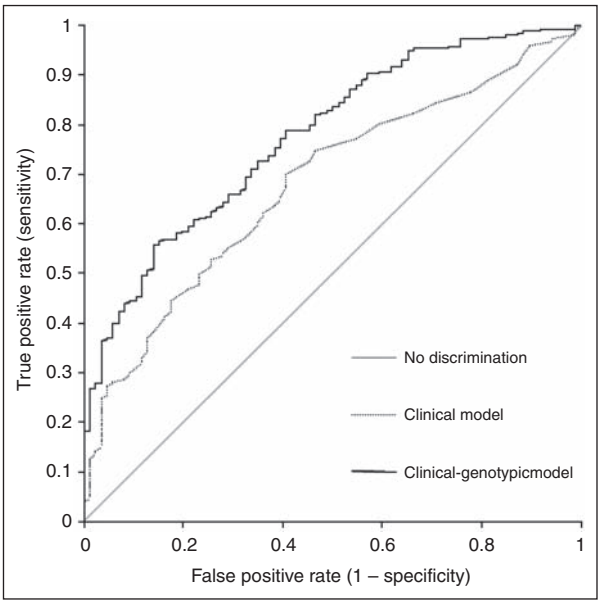

Figure 1. Receiver Operating Characteristic curves for the 'clinical model' and the 'clinical–genotypic model'. The clinical model includes the following variables: age at onset, gender and onset type. The clinical-genotypic model includes the variables mentioned in Table 4. The 'clinical model' had an area under the curve (AUC) of 0.68 and the clinical–genotypic model had an AUC of 0.78. The Likelihood Ratio Test showed that the model including single nucleotide polymorphisms fits the data significantly better than the model based on clinical variables only ($\chi^2 = 46.89$, df = 10, $p < 0.0001$).

variables only ($\chi^2 = 46.89$, df = 10, $p < 0.0001$). In addition to this, the model including SNPs showed a lower AIC than that based on clinical variables only (443.9 versus 470.7). The multivariate model combining clinical data and genotypic data significantly predicted the severity of the disease (model $\chi^2 = 78.8$, $p < 0.001$). The model discriminated well between patients who had mild and more severe forms of the disease (AUC = 0.78, bootstrap 95% CI 0.75–0.84).

Discussion

As expected, we could not detect one major SNP/gene related to disease severity in our study. The current hypothesis concerning susceptibility genes and MS is based on the assumption that multiple genes exert a small effect on developing MS on top of the major influence of HLA-DRB1*1501. Our hypothesis was based on transferring this assumption to genes involved in disease progression. We therefore developed an instrument to detect these effects, and built predictive models including several SNPs in addition to relevant clinical parameters. We tested three different predictive models based on different outcome measures, as no single golden standard has been described in literature. Our results show that a combination of SNPs increases the predictive value of the models.

Different genes were included in the different models, showing in general small individual effects on the outcome. The most prominent gene, included in all three models with moderately high coefficients encodes for interleukin-2 (IL-2). IL-2 (T-cell growth factor) is an immunoregulatory cytokine important for T-cell homeostasis and is involved in the regulation of autoimmunity.[28–30] Previously this gene was found to be related to susceptibility to MS,[31] and genetic differences in this gene were found between patients with relapsing remitting MS and patients with secondary progressive MS.[31,32] Interestingly, daclizumab, a humanized monoclonal antibody (mAb) that blocks the binding of IL-2 to the interleukin-2 receptor alpha unit (IL-2R-alpha chain; CD25), has shown to be effective in patients who experienced persistent MS disease activity with first-line therapy.[33] Moreover, in patients with secondary progressive MS heightened levels of IL-2 were reported, and cerebrospinal fluid (CSF) concentrations of IL-2 were correlated with the degree of disability in patients with clinically active patients.[34] Genetic variation within the IL-2 gene is likely to modify disease progression.

Interestingly, rs3135388 (a surrogate marker for HLA-DRB1*1501) did not show any effect on disease severity in any of the models. This contrasts with other studies that have shown that carriers of the HLA-DRB1*1501 had a more severe disease course using an extreme outcome strategy.[16] When applying the same strategy and definitions to our patient group, sample sizes of the benign and severe MS patient groups were too small to test this hypothesis.

The most convincing model with the highest predictive power and the highest coefficients was the dichotomous model on the MSSS. The dichotomous model on the MSSS discriminates benign patients from the more severely disabled patients by using a cut-off point of MSSS 2.5 (equivalent to an EDSS of three or lower at 15 years of disease duration, the common definition of benign MS). A dichotomous model might be more sensitive in detecting the small effects of the SNPs; however, the survival model (time to reach EDSS 6) and the linear MSSS model provide more detailed information on the speed of disability accumulation and on the severity of the disability. In our study only 38.7% of the patients reached an EDSS of 6, thereby limiting the power of survival analysis. Although we tried to lower the chance of overfitting of our predictive model, our findings should be further explored and confirmed in a different cohort of MS patients, preferably with a longer disease duration.

To improve the prediction on disability, selection of the right SNPs is essential. Our selection of SNPs was based on a heterogeneously reported association between this gene and multiple sclerosis susceptibility or phenotypes. The selection of SNPs was performed in the 'pre- Genome-Wide Association Study (GWAS) era'. Using information from recently published data of genes influencing disease phenotype[35] could substantially improve our SNP selection in search of predictive models. Baranzini et al.[35] showed that MSSS was associated with genes relatively new to the MS literature. These genes are involved in cellular mechanisms such as protein amino acid N-linked glycosylation, cellular respiration and embryonic development. This study illustrates that disease-modifying genes are not necessarily identical to disease susceptibility genes.[35]

More and more evidence points towards a predictive value of MRI parameters early in the disease for disability later on.[36] Including these early MRI parameters in the predictive model could improve the predictive value on future disability accumulation. Moreover, inclusion of yet-to-be-discovered genes, and maybe environmental factors, might increase the explained variance in disease severity.

In summary, we showed in a relatively large sample of well-characterized patients that in addition to clinical variables, genetic information is valuable to improve the prediction of disease severity in MS. However, to more precisely estimate the true genetic influence on disability accumulation in MS, replication of our results is key.

Acknowledgements

We would like to thank all participants for their participation.

Funding

The study was supported by a corporate sponsor (Progenika Biopharma S.A.). Design and conduct study, including selection of SNPs for MSchip: academia and corporate sponsor jointly. MSchip intellectually protected by corporate sponsor. Genetic analysis performed by academia and corporate sponsor jointly. Collection management analysis and interpretation of the data: academia. Preparation review and approval of the manuscript: first draft manuscript, academia; corporate sponsor reviewed and approved current manuscript.

Conflict of interest

Dr D Arteta, Dr D Tejedor and Dr A Martínez are current employees of the sponsor.

Professor Dr AS Peña is a member of the scientific advisory board of Progenika Biopharma.

Dr J Killestein has been involved in clinical trials of companies that market drugs for MS (Schering, AG, Biogen Idec, Serono, Teva) and with some companies that have development programmes for future drugs in MS.

Professor Dr CH Polman has accepted consulting fees or speaking fees from Actelion, Biogen Idec, Bayer Schering, Teva, Merck-Serono, Novartis, GlaxoSmithKline, UCB, Roche and Antisense Therapeutics; and grant support from Biogen Idec, Bayer Schering, GlaxoSmithKline, Novartis, UCB, Merck-Serono and Teva.

Dr BMJ Uitdehaag has received honoraria for consultancy from Novartis and Merck Serono.

References

1. Compston A, Coles A. Multiple sclerosis. *Lancet* 2002; 359: 1221–1231.
2. Gregory SG, Schmidt S, Seth P, et al. Interleukin 7 receptor alpha chain (IL7R) shows allelic and functional association with multiple sclerosis. *Nat Genet* 2007; 39: 1083–1091.
3. Hafler DA, Compston A, Sawcer S, et al. Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med* 2007; 357: 851–862.
4. Lundmark F, Duvefelt K, Iacobaeus E, et al. Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. *Nat Genet* 2007; 39: 1108–1113.
5. De Jager PL, Jia X, Wang J, et al. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat Genet* 2009; 41: 776–782.
6. Frohman EM, Filippi M, Stuve O, et al. Characterizing the mechanisms of progression in multiple sclerosis: Evidence and new hypotheses for future directions. *Arch Neurol* 2005; 62: 1345–1356.
7. Doolittle TH, Myers RH, Lehrich JR, et al. Multiple sclerosis sibling pairs: Clustered onset and familial predisposition. *Neurology* 1990; 40: 1546–1552.
8. Bulman DE, Sadovnick AD, Ebers GC. Age of onset in siblings concordant for multiple sclerosis. *Brain* 1991; 114: 937–950.
9. Brassat D, Azais-Vuillemin C, Yaouanq J, et al. Familial factors influence disability in MS multiplex families. French Multiple Sclerosis Genetics Group. *Neurology* 1999; 52: 1632–1636.
10. Barcellos LF, Oksenberg JR, Green AJ, et al. Genetic basis for clinical expression in multiple sclerosis. *Brain* 2002; 125: 150–158.
11. Hensiek AE, Seaman SR, Barcellos LF, et al. Familial effects on the clinical course of multiple sclerosis. *Neurology* 2007; 68: 376–383.
12. Kantarci OH, de Andrade M, Weinshenker BG. Identifying disease modifying genes in multiple sclerosis. *J Neuroimmunol* 2002; 123: 144–159.
13. Ramagopalan SV, DeLuca GC, Degenhardt A, Ebers GC. The genetics of clinical outcome in multiple sclerosis. *J Neuroimmunol* 2008; 201–202: 183–199.

14. Ramagopalan SV, DeLuca GC, Morrison KM, et al. Analysis of 45 candidate genes for disease modifying activity in multiple sclerosis. *J Neurol* 2008; 255: 1215–1219.
15. DeLuca GC, Ramagopalan SV, Herrera BM, et al. An extremes of outcome strategy provides evidence that multiple sclerosis severity is determined by alleles at the HLA-DRB1 locus. *Proc Natl Acad Sci U S A* 2007; 104: 20896–20901.
16. Barcellos LF, Oksenberg JR, Begovich AB, et al. HLA-DR2 dose effect on susceptibility to multiple sclerosis and influence on disease course. *Am J Hum Genet* 2003; 72: 710–716.
17. Hoffmann S, Cepok S, Grummel V, et al. HLA-DRB1*0401 and HLA-DRB1*0408 are strongly associated with the development of antibodies against interferon-beta therapy in multiple sclerosis. *Am J Hum Genet* 2008; 83: 219–227.
18. Okuda DT, Srinivasan R, Oksenberg JR, et al. Genotype-Phenotype correlations in multiple sclerosis: HLA genes influence disease severity inferred by 1HMR spectroscopy and MRI measures. *Brain* 2009; 132: 250–259.
19. Sombekke MH, Lukas C, Crusius JB, et al. HLA-DRB1*1501 and spinal cord magnetic resonance imaging lesions in multiple sclerosis. *Arch Neurol* 2009; 66: 1531–1536.
20. Polman CH, Reingold SC, Edan G, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the ''McDonald Criteria''. *Ann Neurol* 2005; 58: 840–846.
21. Poser CM, Paty DW, Scheinberg L, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol* 1983; 13: 227–231.
22. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an Expanded Disability Status Scale (EDSS). *Neurology* 1983; 33: 1444–1452.
23. Roxburgh RH, Seaman SR, Masterman T, et al. Multiple Sclerosis Severity Score: using disability and disease duration to rate disease severity. *Neurology* 2005; 64: 1144–1151.
24. Tejedor D, Castillo S, Mozas P, et al. Reliable low-density DNA array based on allele-specific probes for detection of 118 mutations causing familial hypercholesterolemia. *Clin Chem* 2005; 51: 1137–1144.
25. Benjamini Y, Hockberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 1995; 57: 289–300.
26. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; 44: 837–845.
27. van Houwelingen JC, Le Cessie S. Predictive value of statistical models. *Stat Med* 1990; 9: 1303–1325.
28. Horak I, Lohler J, Ma A, Smith KA. Interleukin-2 deficient mice: a new model to study autoimmunity and self-tolerance. *Immunol Rev* 1995; 148: 35–44.
29. Maier LM, Wicker LS. Genetic susceptibility to type 1 diabetes. *Curr Opin Immunol* 2005; 17: 601–608.
30. Schimpl A, Berberich I, Kneitz B, et al. IL-2 and autoimmune disease. *Cytokine Growth Factor Rev* 2002; 13: 369–378.
31. Matesanz F, Fedetz M, Collado-Romero M, et al. Allelic expression and interleukin-2 polymorphisms in multiple sclerosis. *J Neuroimmunol* 2001; 119: 101–105.
32. Fedetz M, Ndagire D, Fernandez O, et al. Multiple sclerosis association study with the TENR-IL2-IL21 region in a Spanish population. *Tissue Antigens* 2009; 74: 244–247.
33. Rose JW, Burns JB, Bjorklund J, Klein J, Watt HE, Carlson NG. Daclizumab phase II trial in relapsing and remitting multiple sclerosis: MRI and clinical results. *Neurology* 2007; 69: 785–789.
34. Killestein J, Den Drijver BF, Van der Graaff WL, Uitdehaag BM, Polman CH, Van Lier RA. Intracellular cytokine profile in T-cell subsets of multiple sclerosis patients: different features in primary progressive disease. *Mult Scler* 2001; 7: 145–150.
35. Baranzini SE, Wang J, Gibson RA, et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet* 2008; 18: 767–778.
36. Fisniku LK, Brex PA, Altmann DR, et al. Disability and T2 MRI lesions: a 20-year follow-up of patients with relapse onset of multiple sclerosis. *Brain* 2008; 131: 808–817.

Supplementary tables

Table 1. SNPs included on the DNA chip. MAF=minor allele frequency in our sample

Gene	rs-number	Chromosome	Polymorphism	MAF
ADAMTS14*	rs4747075	10q22	A/G	0.29
ADAMTS14	rs7081273	10q22	C/G	0.35
ADAMTS14	rs4746060	10q22	C/T	0.09
BTNL2*	rs2076530	6p21.3	A/G	0.23
CACNG4	rs4790896	17q24	C/T	0.41
CCDC46	rs987931	17q24	G/T	0.34
CCL5	rs2280788	17q11.2-q12	C/G	0.03**
CCL5	rs2107538	17q11.2-q12	C/T	0.19
CCR5	rs333	3p21	−/+	0.10
CD24	rs8734	6q21	C/T	0.00**
CIITA	rs3087456	16p13	A/G	0.26
CNTF	rs1800169	11q12	A/G	0.13
CRYAB	rs14133	11q21–q23	C/G	0.26
CRYAB	rs762550	11q21–q23	A/G	0.37
CTLA4	rs231775	2q33	A/G	0.37
CTLA4	rs5742909	2q33	C/T	0.08
EBF1	rs1368297	5q34	A/T	0.38
FAS	rs1800682	10q23	C/T	0.45
FAS	rs3781202	10q23	C/T	0.40
FAS	rs2234978	10q23	C/T	0.31
GABBR1	rs1805057	6p21.3	A/G	0.00**
HELZ	rs2363846	17q24	C/T	0.49
HLA*	rs2395166	6p21.3	C/T	0.47
HLA-DRA	rs2213584	6p21.3	C/T	0.41
HLA	rs2227139	6p21.3	C/T	0.41
HLA-DRA	rs3135388	6p21.3	C/T	0.31
HLA	rs9268458	6p21.3	A/C	0.18
HLA*	rs6457594	6p21.3	A/G	0.39
HLA-DRA	rs2395182	6p21.3	G/T	0.39
HLA-DRA	rs2239802	6p21.3	C/G	0.39
HSPB2	rs2234702	11q21–q23	C/G	0.00**
IFNAR1	rs2257167	21q22	C/G	0.11
IFNGR2	rs9808753	21q22	A/G	0.15
IL1B	rs1799916	2q14	G/T	0.00**
IL1B	rs1143627	2q14	C/T	0.34
IL1B	rs1143634	2q14	C/T	0.27
IL1RN	rs419598	2q14	C/T	0.27
IL1RN	2073 C/T	2q14	C/T	0.27
IL2	rs2069763	4q26–27	G/T	0.36
IL2	rs2069762	4q26–27	G/T	0.25
IL4R	rs1801275	16p12	A/G	0.23
IL7R	rs11567685	5p13	C/T	0.25
IL7R	rs7718919	5p13	G/T	0.14
IL7R	rs11567686	5p13	A/G	0.35
IL10*	rs1800896	1q32	A/G	0.50
MC1R	rs1805009	16q24	C/G	0.02**
MC1R	rs1805006	16q24	A/C	0.01**

(continued)

Table 1. *Continued*

Gene	rs-number	Chromosome	Polymorphism	MAF
MEFV	rs28940577	16p13.3	A/G	0.00**
MOG	rs3130250	6p22	A/G	0.20
MOG	rs3130253	6p22	A/G	0.11
NDUFA7	rs2288414	19p13.2	C/G	0.03**
NDUFA7	rs561	19p13.2	A/G	0.18
NDUFS5	rs2889683	1p34.2	C/T	0.30
NDUFS5	rs6981	1p34.2	A/G	0.06
NDUFS7	rs2074897	19p13.3	A/G	0.49
NFKBIL1	rs3130062	6p21.3	C/T	0.15
NOS2	rs1137933	17q11.2	C/T	0.24
NOS2	rs2779248	17q11.2	C/T	0.40
NOTCH4	rs367398	6p21.3	A/G	0.21
PDCD1	rs11568821	2q37	A/G	0.11
PITPNC1	rs1318	17q24	C/T	0.20
PITPNC1	rs2365403	17q24	C/G	0.17
PNMT	rs876493	17q11–q23	C/T	0.42
PRKCA	rs7220007	17q24	A/G	0.48
PRKCA	rs887797	17q24	C/T	0.29
PRKCA	rs2078153	17q24	C/G	0.23
PRKCA *	rs3890137	17q24	A/G	0.36
PTPN22	rs2476601	1p13	A/G	0.10
PTPRC	rs17612648	1q31	C/G	0.02**
PTPRC	rs4915154	1q31	A/G	0.01**
SPP1	rs1126616	4q21	C/T	0.23
SPP1	rs1126772	4q21	A/G	0.19
SPP1	rs2853744	4q21	G/T	0.05
SPP1	rs9138	4q21	A/C	0.24
SPP1	rs4754	4q21	C/T	0.24
TNF	rs1800629	6p21.3	A/G	0.17
TNFSF10	rs1131568***	3q26	A/G	0.34
UCP2	rs659366	11q13	C/T	0.36
VDR	rs1544410	12q13	A/G	0.40
VDR	rs731236	12q13	C/T	0.39

*Deviation from Hardy Weinberg proportions with $p < 0.05$.
**Excluded from analyses due to minor allele frequency < 0.05.
***Previously rs-number: rs9880164.
Abbreviations:
ADAMTS14 = a disintegrin and metalloproteinase with thrombospondin motif, type 1 motif 14; BTNL2 = butyrophilin-like 2; CACNG4 = calcium channel = voltage-dependent, gamma subunit 4; CCDC46 = coiled coil domain containing 46; CCL5 = chemokine (C-C motif) ligand 5; CCR5 = chemokine (C-C motif) receptor 5; CIITA = class II, major histocompatibility complex = transactivator; CNTF = ciliary neurotrophic factor; CRYAB = Alpha B crystallin; CTLA4 = cytotoxic T-lymphocyte-associated protein 4; EBF1 = early B-cell factor 1; FAS = TNF receptor superfamily, member 6; GABBR1 = gamma-aminobutyric acid (GABA) B receptor, 1; HELZ = helicase with zinc finger; HLA = human leucocyte antigen; HLA-DRA = human leucocyte antigen DR alpha; HSPB2 = heat shock 27kDa protein 2; IFNAR1 = interferon (alpha = beta and omega) receptor 1; IFNGR2 = interferon gamma receptor 2 (interferon gamma transducer 1); IL1B = interleukin 1, beta; IL1RN = interleukin 1 receptor antagonist; IL2 = interleukin 2; IL4R = interleukin 4 receptor; IL7R = interleukin 7 receptor; IL10 = interleukin 10; MC1R = melanocortin 1 receptor; MEFV = mediterranean fever; MOG = myelin oligodendrocyte glycoprotein; NDUFA7 = NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 7; NDUFS5 = NADH dehydrogenase (ubiquinone) Fe-S protein 5; NDUFS7 = NADH dehydrogenase (ubiquinone) Fe-S protein 7; NFKIBL1 = nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-like 1; NOS2 = nitric oxide synthase 2; NOTCH4 = Notch homolog 4; PDCD1 = programmed cell death 1; PITPNC1 = phosphatidylinositol transfer protein, cytoplasmic 1; PNMT = phenylethanolamine N-methyltransferase; PRKCA = protein kinase C, alpha; PTPN22 = protein tyrosine phosphatase, non-receptor type 22; PTPRC = protein tyrosine phosphatase, receptor type, C; SPP1 = secreted phosphoprotein 1/osteopontin; TNF = tumor necrosis factor; TNFSF10 = tumor necrosis factor (ligand) superfamily, member 10; UCP2 = uncoupling protein 2; VDR = vitamin D (1,25- dihydroxyvitamin D3) receptor.

Table 2. Results of univariate analysis before and after correction for multiple testing. Significant values are printed in bold

Gene	rs-Number	p-value logrank test on time to reach EDDS6	Corrected p-value (FDR) logrank test on time to reach EDDS6	p-value (Kruskal Wallis) on linear MSSS	Corrected p-value (FDR) on linear MSSS	p-value (ChiSquare) on dichotomous MSSS	Corrected p-value on ChiSquare on dichotomous MSSS
ADAMTS14	rs4747075	**0.02**	0.52	**0.02**	0.60	0.09	0.86
ADAMTS14	rs4746060	0.11	0.64	0.38	0.84	0.16	0.87
ADAMTS14	rs7081273	**0.03**	0.55	0.19	0.68	0.30	0.87
BTNL2	rs2076530	0.58	0.97	0.50	0.88	0.30	0.87
CACNG4	rs4790896	0.78	0.97	0.61	0.91	0.55	0.97
CCDC46	rs987931	0.17	0.66	0.62	0.92	0.46	0.97
CIITA	rs3087456	0.87	0.97	0.87	0.93	0.94	0.97
CCL5	rs2107538	0.19	0.66	0.21	0.68	**0.04**	0.47
CCR5	rs333	0.33	0.82	0.42	0.84	0.61	0.97
CNTF	rs1800169	0.83	0.97	0.72	0.93	0.59	0.97
CRYAB	rs762550	0.30	0.80	0.83	0.93	0.66	0.97
CRYAB	rs14133	0.98	1.00	0.71	0.93	0.74	0.97
CTLA4	rs5742909	0.85	0.97	0.18	0.68	0.11	0.87
CTLA4	rs231775	0.27	0.80	0.44	0.85	0.64	0.97
EBF1	rs1368297	0.86	0.97	0.93	0.96	0.71	0.97
FAS	rs2234978	**0.01**	0.52	0.47	0.85	0.49	0.97
FAS	rs1800682	0.99	1.00	0.81	0.93	0.64	0.97
FAS	rs3781202	0.83	0.97	0.97	0.97	0.85	0.97
HELZ	rs2363846	0.52	0.97	0.21	0.68	0.55	0.97
HLA	rs3135388	0.32	0.82	0.12	0.65	0.25	0.87
HLA-DRA	rs2395166	0.12	0.64	0.20	0.68	0.96	0.97
HLA	rs9268458	0.29	0.80	0.36	0.84	0.51	0.97
HLA-DRA	rs2213584	0.66	0.97	0.38	0.84	0.54	0.97
HLA	rs2227139	0.70	0.97	0.43	0.84	0.65	0.97
HLA	rs6457594	0.69	0.97	0.47	0.85	0.71	0.97
HLA-DRA	rs2395182	0.30	0.80	0.06	0.65	0.16	0.87
HLA-DRA	rs2239802	0.24	0.76	0.06	0.65	0.16	0.87
IFNAR1	rs2257167	0.15	0.64	0.10	0.65	<**0.01**	0.10
IFNGR2	rs9808753	0.11	0.64	0.08	0.65	0.12	0.87
IKBL	rs3130062	0.71	0.97	0.41	0.84	0.30	0.87
IL1B	rs1143627	0.06	0.55	0.10	0.65	0.83	0.97
IL1B	rs1143634	0.94	1.00	0.76	0.93	0.64	0.97
IL1RN	@2073CTIntron2	0.78	0.97	0.89	0.93	0.18	0.87
IL1RN	rs419598	0.72	0.97	0.96	0.97	0.20	0.87
IL2	rs2069763	**0.05**	0.55	**0.01**	0.49	**0.01**	0.12
IL2	rs2069762	0.07	0.55	0.17	0.68	0.97	0.97
IL4	rs1801275	0.06	0.55	0.42	0.84	0.67	0.97
IL7R	rs11567685	**0.05**	0.55	0.08	0.65	0.32	0.87
IL7R	rs7718919	0.61	0.97	0.25	0.69	0.75	0.97
IL7R	rs11567686	0.67	0.97	0.66	0.93	0.90	0.97
IL10	rs1800896	0.14	0.64	0.22	0.69	0.83	0.97
MOG	rs3130250	0.18	0.66	0.11	0.65	0.61	0.97
MOG	rs3130253	0.36	0.86	0.12	0.65	0.91	0.97
NDUFA7	rs561	0.80	0.97	0.86	0.93	0.26	0.87

(continued)

Table 2. *Continued*

Gene	rs-Number	p-value logrank test on time to reach EDDS6	Corrected p-value (FDR) logrank test on time to reach EDDS6	p-value (Kruskal Wallis) on linear MSSS	Corrected p-value (FDR) on linear MSSS	p-value (ChiSquare) on dichotomous MSSS	Corrected p-value on ChiSquare on dichotomous MSSS
NDUFS5	rs6981	0.47	0.97	0.20	0.68	0.26	0.87
NDUFS5	rs2889683	0.24	0.76	0.38	0.84	0.61	0.97
NDUFS7	rs2074897	0.07	0.55	0.19	0.68	0.81	0.97
NOS2	rs1137933	0.47	0.97	0.42	0.84	<**0.01**	0.12
NOS2	rs2779248	0.85	0.97	0.60	0.91	0.38	0.97
NOTCH4	rs367398	0.70	0.97	0.87	0.93	0.96	0.97
PDCD1	rs11568821	0.85	0.97	0.61	0.91	0.75	0.97
PITPNC1	rs1318	0.19	0.66	0.10	0.65	**0.01**	0.12
PITPNC1	rs2365403	0.84	0.97	0.83	0.93	0.24	0.87
PNMT	rs876493	0.38	0.87	0.30	0.80	**0.01**	0.19
PRKCA	rs887797	0.84	0.97	0.25	0.69	0.35	0.93
PRKCA	rs3890137	0.13	0.64	0.69	0.93	0.80	0.97
PRKCA	rs2078153	0.69	0.97	0.83	0.93	0.90	0.97
PRKCA	rs7220007	0.90	0.98	0.70	0.93	0.85	0.97
PTPN22	rs2476601	0.51	0.97	0.81	0.93	0.64	0.97
SPP1	rs2853744	0.09	0.64	0.10	0.65	0.59	0.97
SPP1	rs4754	0.76	0.97	0.51	0.88	0.65	0.97
SPP1	rs9138	0.74	0.97	0.56	0.90	0.67	0.97
SPP1	rs1126616	0.77	0.97	0.55	0.90	0.68	0.97
SPP1	rs1126772	0.99	1.00	0.78	0.93	0.94	0.97
TNF	rs1800629	1.00	1.00	0.54	0.90	0.27	0.87
TNFSF10	rs9880164	0.82	0.97	0.24	0.69	0.18	0.87
UCP2	rs659366	0.95	1.00	0.84	0.93	0.26	0.87
VDR	rs731236	0.65	0.97	0.87	0.93	0.84	0.97
VDR	rs1544410	0.85	0.97	0.79	0.93	0.95	0.97

Abbreviations:
ADAMTS14 = a disintegrin and metalloproteinase with thrombospondin motif = type 1 motif 14; BTNL2 = butyrophilin-like 2; CACNG4 = calcium channel = voltage-dependent = gamma subunit 4; CCDC46 = coiled coil domain containing 46; CCL5 = chemokine (C-C motif) ligand 5; CCR5 = chemokine (C-C motif) receptor 5; CIITA = class II, major histocompatibility complex = transactivator; CNTF = ciliary neurotrophic factor; CRYAB = Alpha B crystallin; CTLA4 = cytotoxic T-lymphocyte-associated protein 4; EBF1 = early B-cell factor 1; FAS = TNF receptor superfamily, member 6; GABBR1 = gamma-aminobutyric acid (GABA) B receptor, 1; HELZ = helicase with zinc finger; HLA = human leucocyte antigen; HLA-DRA = human leucocyte antigen DR alpha; HSPB2 = heat shock 27kDa protein 2; IFNAR1 = interferon (alpha, beta and omega) receptor 1; IFNGR2 = interferon gamma receptor 2 (interferon gamma transducer 1); IL1B = interleukin 1, beta; IL1RN = interleukin 1 receptor antagonist; IL2 = interleukin 2; IL4R = interleukin 4 receptor; IL7R = interleukin 7 receptor; IL10 = interleukin 10; MC1R = melanocortin 1 receptor; MEFV = mediterranean fever; MOG = myelin oligodendrocyte glycoprotein; NDUFA7 = NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 7; NDUFS5 = NADH dehydrogenase (ubiquinone) Fe-S protein 5; NDUFS7 = NADH dehydrogenase (ubiquinone) Fe-S protein 7; NFKIBL1 = nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-like 1; NOS2 = nitric oxide synthase 2; NOTCH4 = Notch homolog 4; PDCD1 = programmed cell death 1; PITPNC1 = phosphatidylinositol transfer protein, cytoplasmic 1; PNMT = phenylethanolamine N-methyltransferase; PRKCA = protein kinase C, alpha; PTPN22 = protein tyrosine phosphatase = non-receptor type 22; PTPRC = protein tyrosine phosphatase, receptor type, C; SPP1 = secreted phosphoprotein 1/osteopontin; TNF = tumor necrosis factor; TNFSF10 = tumor necrosis factor (ligand) superfamily = member 10; UCP2 = uncoupling protein 2; VDR = vitamin D (1,25- dihydroxyvitamin D3) receptor.