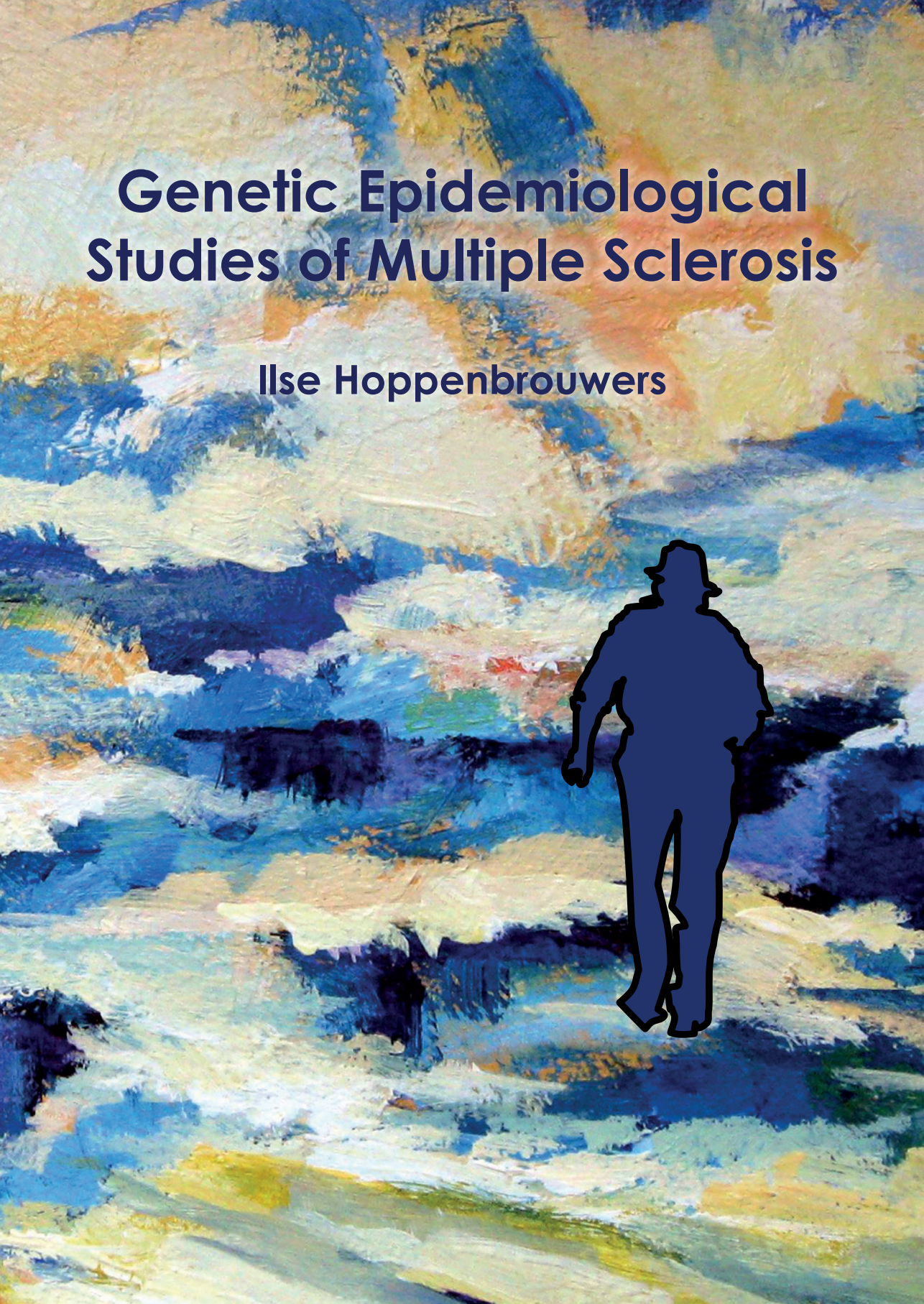
Genetic Epidemiological Studies of Multiple Sclerosis

Ilse Hoppenbrouwers



Genetic Epidemiological Studies of Multiple Sclerosis

ISBN: 978-90-5335-354-7

Cover design by: Theo Bemelmans, Nuenen

Thesis lay-out and printing: Simone Vinke, Ridderprint Offsetdrukkerij B.V., Ridderkerk

Financial support for the printing of this thesis was kindly provided by Biogen Idec, Sanofi-Aventis, Boehringer Ingelheim, Bayer Schering Pharma, TEVA Pharma, Novartis Pharma BV, Merck Serono and the Erasmus MC

Genetic Epidemiological Studies of Multiple Sclerosis

Genetisch epidemiologische studies van multiple sclerose

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de rector magnificus
Prof.dr. H.G. Schmidt
en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op
woensdag 12 januari 2011 om 13.30 uur

door

Ilse Anne Hoppenbrouwers

Promotiecommissie

Promotoren:	Prof.dr.ir. C.M. van Duijn
	Prof.dr. R.Q. Hintzen
	Prof.dr. B.A. Oostra
Overige leden:	Prof.dr. P.A. van Doorn
	Prof.dr. J.D. Laman
	Prof.dr. C. Wijmenga

Table of contents

Manuscripts based on the studies described in this thesis

Chapter 1
Hoppenbrouwers IA, Hintzen RQ. Genetics of multiple sclerosis. Accepted for publication in Biochim Biophys Acta.

Chapter 2
Hoppenbrouwers IA, Cortes LM, Aulchenko YS, Sintnicolaas K, Njajou O, Snijders PJ, Oostra BA, van Duijn CM, Hintzen RQ. Familial clustering of multiple sclerosis in a Dutch genetic isolate. Mult Scler 2007;13:17-24.

Chapter 3
Hoppenbrouwers IA, Liu F, Aulchenko YS, Ebers GC, Oostra BA, van Duijn CM, Hintzen RQ. Maternal transmission of multiple sclerosis in a dutch population. Arch Neurol 2008;65:345-348.

Chapter 4.1
Hoppenbrouwers IA, Aulchenko YS, Ebers GC, Ramagopalan SV, Oostra BA, van Duijn CM, Hintzen RQ. EVI5 is a risk gene for multiple sclerosis. Genes Immun 2008;9:334-337.

Chapter 4.2
Hoppenbrouwers IA, Aulchenko YS, Janssens AC, Ramagopalan SV, Broer L, Kayser M, Ebers GC, Oostra BA, van duijn CM, Hintzen RQ. Replication of CD58 and CLEC16A as genome-wide significant risk genes for multiple sclerosis. J Hum Genet 2009;54:676-680.

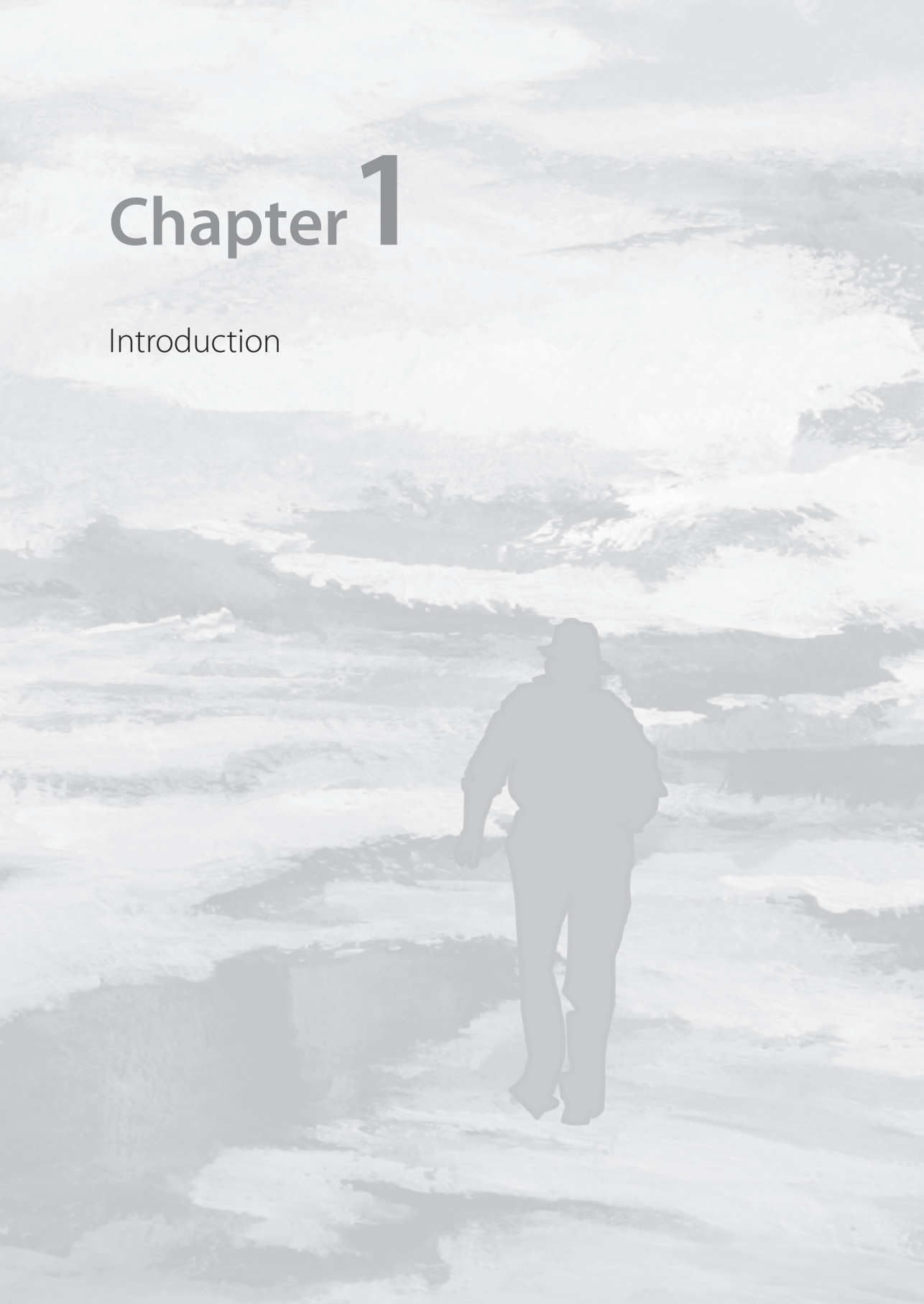
Chapter 5
Hoppenbrouwers IA,* Aulchenko YS,* Ramagopalan SV, Broer L, Jafari N, Hillert J, Link J, Lundström W, Greiner E, Dessa Sadovnick A, Goossens D, Van Broeckhoven C, Del-Favero J, Ebers GC, Oostra BA, van Duijn CM, Hintzen RQ. Genetic variation in the KIF1B locus influences susceptibility to multiple sclerosis. Nat Genet 2008;40:1402-1403.

Chapter 6
Hoppenbrouwers IA, Hintzen RQ. Genetics of multiple sclerosis. Accepted for publication in Biochim Biophys Acta.

* shared first authors

Chapter 1

Introduction



Multiple sclerosis (MS) is a chronic neurological disorder characterized by inflammation, myelin loss, axonal pathology in the central nervous system and progressive neurological dysfunction. It is presumed to be an autoimmune disorder and is believed to arise from complex interactions of both environmental and genetic factors. For the majority of affected individuals (80-90%), the disease begins as episodic, with full recovery after a relapse, and then over a period of time becomes progressive. In around 10-20% of patients the illness is progressive from onset, without clinical remissions. The relapsing phase of the disease is mediated by focal bursts of inflammation in white matter in the brain and spinal cord, whereas axonal and neuronal loss predominate during the progressive phase. Current therapies serve mainly to moderate the initial relapsing-remitting phase, but they have much less effect on disease progression and long term disability.[1]

MS is a common disease, affecting over 2 million people worldwide. The population prevalence in North America and Northern Europe is ~0.1 %.[2] The disease is characterized by a female excess, the female-to-male ratio now exceeds 3.2:1.[3] Incidence tends to be low in childhood (2 % of patients with MS present before age of 10 years and 5 % before age of 16 years), increases after the age of 18, reaching a peak between 20 and 35 years and then declines becoming rare at age older than 50.[2]

Are environmental factors involved in MS?

There is substantial evidence for environmental factors to be important in MS susceptibility. The disease incidence and prevalence variates according to geography, and is higher with increasing distance from the equator.[4-6] Migration from a low to high risk region for MS or from a high to low risk region for MS can alter the risk to develop MS,[7] indicating that endemic exposures occur and are the primary determinant of the disease. However, migration studies are typically unable to establish timing of environmental exposures owing to small numbers.

Further evidence for environmental factors being involved in susceptibility for MS comes from twin studies. In monozygotic twins the concordance rate is ~30 % instead of the 100 % that is expected when only genetic factors are important in susceptibility for MS.[8] In Canada, the risk of MS for a dizygotic twin of a patient with MS is almost twice the risk for a full non-twin sibling, which may implicate environmental factor(s), such as shared timing of gestation, birth, or both.[8]

Studies of month of birth in several countries in the northern hemisphere showed latitude-correlated increased MS risks for May births and decreased risks for November births, compared to population controls. In addition to population controls a Canadian study used unaffected siblings as a second control group. The unaffected sibling controls confirmed and extended the findings as they were, like the population control group, significantly more often born in November compared to their affected brothers and sisters. Thus month of birth and MS risk are associated which implies interactions between genes and environment. The abrupt change in risk by month suggests a threshold effect for both increased and decreased risk, which cannot easily be explained. The risk factors responsible for the effect of timing of birth must vary seasonally and are probably related to climate.[9] They may act during gestation or shortly after birth in individuals born in the northern countries.

Finally, the female-to-male MS gender ratio has shown to be increasing for at least 50 years in studies performed in Europe and North America. Since there is no indication that MS in men has decreased, it seems that the sex ratio change is determined by a preferential increase in affected women. These changes in sex ratio occur too fast to be accredited to genetic mutations. The factors must be environmental, perhaps resulting from gene-environment interactions.[3,10,11] In a study of avuncular pairs (aunts/uncles and nieces/nephews), this change in sex ratio has also been shown. The overall sex ratio for aunts/uncles was significantly lower than for nieces/nephews. Only in the maternal families this difference in sex ratio between two generations was shown. In the paternal families the aunt/uncle ratio was not different from the niece/nephew ratio.[12]

Is there a genetic component in MS?

Although familial aggregation of MS has long been accepted, systematic age-adjusted recurrence risks for relatives of persons with MS were first published in 1988.[13] Subsequent studies showed that first-, second- and third-degree relatives of patients with MS were more likely to have the disease than the general population.[14,15] The risk for a first-degree relative of a MS proband (2-5%) is ~20 to 50 times greater than that for the general population (0.1%).[14,16] Studies in twins showed a significant excess of concordance in monozygotic (~30 % concordance) as compared to dizygotic twins (~5% concordance).[8,17-20] In a study of individuals with MS who were adopted, a higher risk in the genetically related family was found, whereas no higher risk in the adopting family was found.[21] Affected husband and wife couples do not occur more frequently than expected by chance. However, the risk for their offspring is greater than if just one parent is affected.[22] For half-siblings the risk is approximately half the risk for full siblings, regardless whether they were raised together or apart.[23,24] Together, these data suggest that living with someone who has MS only increases your MS risk if you are a relative of that person and this risk increases with relatedness.

Mode of inheritance in MS

Although, data confirm that genetic factors are unequivocally relevant in MS, most MS families contain no more than two or three affected individuals and no clear mode of inheritance can be inferred from segregation analysis.

If common variants were to be responsible for MS susceptibility, between 20 and 100 common variants, each increasing risk by only a modest factor of 1.2-1.5, would be sufficient to account for the prevalence and heritability of MS.[25,26] If rare variants were to be responsible for MS susceptibility, many hundreds if not thousands of rare variants would be required, even if these individually increase risk by as much as 10-20 times.[26] In fact we can not know for sure where the balance between common and rare variants lies in determining susceptibility for MS, but in line with other complex diseases a complex interaction of both common variants with low penetrance and rare variants with high penetrance is expected.[27]

Although a single-gene aetiology can not be ruled out for a subset of pedigrees, no example

of a Mendelian variant of MS has yet been found. Larger more extended pedigrees, with a more than usual concordant disease expression would be expected, when rare but highly penetrant variants explain MS susceptibility.

Parent-of-origin effect in MS

Genetic-epidemiological studies showed that maternal half-siblings, connected through an unaffected mother, were at significantly higher risk of developing MS when compared with paternal half-siblings, connected through an unaffected father (2.35% versus 1.31%; $P=0.048$).[24] Avuncular pairs are significantly more often connected through unaffected mothers than through unaffected fathers.[12]

However, when patients with affected parents were studied, the fathers with MS transmitted the disease to their offspring significantly more often than the mothers with MS.[28] In a large Canadian population-based cohort equal transmission of MS was seen from affected fathers and from affected mothers.[29] Thus, whenever affected parent-child pairs are left out of the model, there is maternal transmission of MS and whenever affected parent-child pairs are studied at the nuclear family level, paternal transmission catches up or even exceeds maternal transmission. These observations could be the result of technical differences between studies.[30]

In the avuncular study a significantly increased niece/nephew sex ratio, compared to the aunt/uncle sex ratio was only seen in the maternal families. These data suggest that there is also a strong maternal influence on the determination of the sex ratio of MS offspring.[29]

A parent-of-origin effect in MS could be the result of several mechanisms. These may operate differently in alternative inheritance models for MS: genetic versus environmental, or both. One of the possible genetic mechanisms is a threshold effect due to an increased number of penetrant susceptibility genes in a given parental lineage. Another possibility is the involvement of epigenetic mechanisms in disease transmission by the affected parent. Epigenesis refers to a mechanism by which environmental factors affect gene expression. This can occur through a wide variety of mechanisms, including the selective methylation of DNA bases and modification of histones. The underlying DNA sequence is not changed. If epigenesis occurs in utero, which may explain the maternal parent of origin effect in MS, it can affect the availability of the critical gene product or modify the risk associated with a given gene polymorphism. In this way, it can directly contribute to the risk of a disease. Epigenetic effects do not have to occur in all cells from one person, they can also operate only in the cells of specific tissues.[30] In single-gene disorders, imprinting is a mechanism behind the differential parent-of-origin effects observed.[31] A third genetic possibility for the parent of origin effect in MS is that the multifactorial threshold and epigenetic mechanisms of inheritance occur in conjunction. This is especially likely in complex disorders as MS.[30]

Susceptibility genes for MS

The HLA region

The first association of MS with genes in the MHC region was identified in 1972.[32,33] Until very recently, despite many linkage and association studies, only the HLA-class II region of the *HLA-DR2* containing haplotype on chromosome 6p21, in both primary progressive and relapsing-remitting patients, was significantly associated with MS. Three candidate risk genes of this haplotype, *HLA-DRB1*1501* (encoding HLA-DR2b), *HLA-DRB5*0101* (encoding HLA-DR2a) and *HLA-DQB1*0602* (encoding HLA-DQ6), are so tightly linked that they are almost invariably inherited together.[34] Therefore, classic genetic studies failed to discriminate between them. Recently, powerful genetic studies have implicated *HLA-DRB1*1501* as the main susceptibility allele in MS.[35,36] The association is strongest in northern Europeans, but is seen in virtually all populations with a notable exception being some Mediterranean populations where MS is associated with *DR3* and *DR4*.[37,38]

Complex interactions within the HLA region

It has become clear that *HLA-DRB1*15* is not the sole risk increasing allele in the MHC region.[39-41] *HLA-DRB1*17* increases the risk of MS, but to a lesser extent than *HLA-DRB1*15*.[39,41,42]

Dominant negative epistasis (the phenomenon that the effects of a gene are modified by one or several other genes) is also seen in the HLA region, as *HLA-DRB1*14* completely abrogates any risk associated with *HLA-DRB1*15*, when they are inherited together.[39-41,43] The relative risk of MS for an individual who carries the *HLA-DRB1*15* allele is about 3. *HLA-DRB1*14* together with *HLA-DRB1*15* reduces the relative risk to approximately 1.[41] *HLA-DRB1*11* bearing haplotypes are protective as well. The protective effect of this haplotype over *HLA-DRB1*15* is weaker than that of *HLA-DRB1*14*.[41]

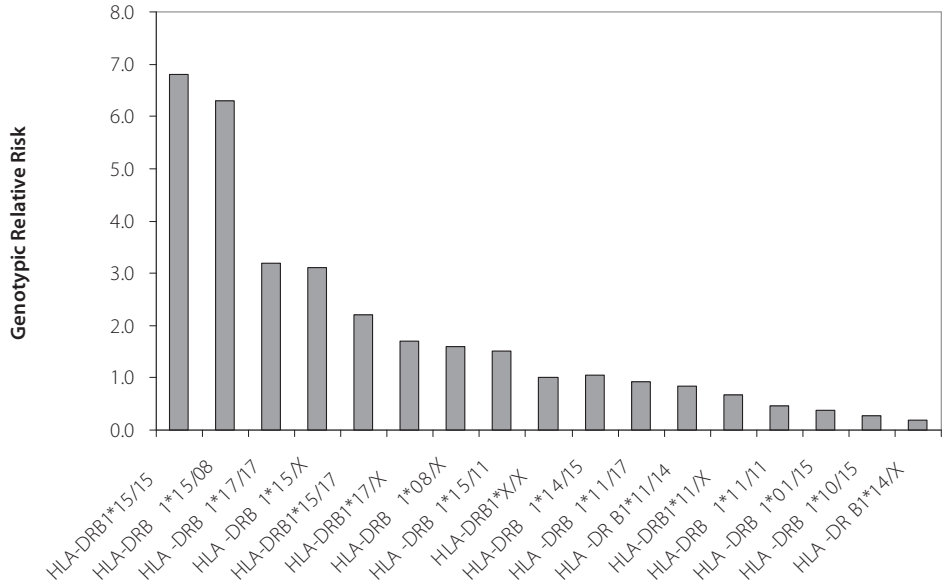
While *HLA-DRB1*11* and *HLA-DRB1*14* have an independent protective effect, *HLA-DRB1*01* and *HLA-DRB1**10 are only protective in the presence of *HLA-DRB1*15*,[39,41] although reports from Sweden suggest that *HLA-DRB1*01* may be protective on its own.[44]

*HLA-DRB1*08* shows the opposite trend as it increases MS risk only in the presence of *HLA-DRB1*15*.[39,41,43] In the figure, the relative risks of MS for combinations of alleles at the *HLA-DRB1* locus is shown.[39,41,43]

Furthermore, there is some evidence for epigenetic effects in the *HLA* region in MS. Using the large Canadian database, it has been possible to analyse affected aunt/uncle-niece/nephew (AUNN) pairs. The *HLA-DRB1*15* frequency in affected males remained the same over the two generations, whereas affected aunts had significantly lower *HLA-DRB1*15* frequency compared with their affected nieces. The risk carried by *HLA-DRB1*15* was greater in families with affected second-degree relatives (AUNN: odds ratio (OR) 4) when compared to those consisting of only affected first-degree relatives (affected sibling pairs (ASP): OR 2). Thus, *HLA-DRB1*15* contributed to risk in both family types, however the degree of contribution is not the same and is dependent on family structure and of MS status of the relatives of the transmitting individual. The exact reason for this is yet to be determined; however, gene-

environment interactions are implied.[45] Moreover, the observations demonstrate that the human MHC is directly involved in the increasing MS incidence in females, suggesting that the MHC is the site of a gene-environment interaction central to disease susceptibility.

Figure 1 Genotyping relative risks for MS for combinations of alleles at the *HLA-DRB1* locus



Data derived from *HLA-DRB1* genotyping performed by the Canadian Collaborative Study Group, comprising 2,535 individuals with definite MS and 4,799 of their unaffected first- degree relatives. x/x=individual with no disease associated alleles with baseline risk of 1, x= any non-disease associated allele; 15, 17, 11, 08, 14, 10, 01-*HLA-DRB1* alleleles. All genotypes are statistically significantly different from baseline, except for *HLA-DRB1*14/15* and *HLA-DRB1*11/17*. No *HLA-DRB1*14* homozygotes were observed, and it should be noted that *HLA-DRB1*10* and *01* provide greater protection against *HLA-DRB1*15* than *HLA-DRB1*14*.
(figure provided by S. Ramagopalan and G.E. Ebers)[43]

Environmental factors and the HLA region

Little is known about gene-environment interaction. There is some evidence that vitamin D and infectious mononucleosis (IM) influence the risk associated with *HLA-DRB1*15*. Epidemiological studies provided strong evidence that the geographical distribution of MS risk is the result of environmental factors operating at the population level. Sunlight, specifically through its role in generating active vitamin D, is a likely important environmental factor for the disease.[4,46,47] It has been shown that dietary vitamin D intake reduces disease risk[46] and that MS patients are deficient in vitamin D.[47] Vitamin D has its actions on immune and central nervous system development and function. Thus vitamin D could influence MS risk in this way.

Most biological effects of vitamin D are regulated by the vitamin D receptor (VDR). This

receptor influences the rate of transcription of vitamin D responsive genes by acting as a ligand activated transcription factor that binds to vitamin D response elements (VDREs) in gene promotors. Using sequence analysis, Ramagopalan and colleagues localized a single VDRE to the promotor region of *HLA-DRB1*.[48] Sequencing of this promotor in more than 1,000 chromosomes from *HLA-DRB1* homozygotes showed conservation of this VDRE on *HLA-DRB1*15*. In contrast, among non-MS associated haplotypes, there was striking variation.[48]

In addition, a functional role for this VDRE was demonstrated.[48] B cells transiently transfected with the *HLA-DRB1*15* gene promotor showed increased expression upon stimulation with 1,25-dihydroxyvitamin D3 (P =0.002), that was lost either upon deletion of the VDRE or with homologous VDRE sequence found in non-MS-associated *HLA-DRB1* haplotypes.[48]

Additionally to vitamin D, interactions may exist with Epstein-Barr virus (EBV) induced IM. Epidemiological and serological studies already showed that infection with EBV is associated with higher risk of MS.[49,50] In particular, primary EBV infection, manifestating as IM seems to be associated with increased MS risk.

In a Danish study in IM-naive individuals, *DRB1*15* carried a 2.4-fold (95% confidence interval [CI], 2.0-3.0) increased MS risk, whereas in persons with IM history *DRB1*15* was associated with a 7.0-fold (95% CI, 3.3-15.4) increased MS risk. Thus, the MS risk conferred by *HLA-DRB1*15* was 2.9 (95% CI, 1.3-6.5)-fold stronger in the presence than in the absence of a history with IM.[51]

Non-HLA MS risk single nucleotide polymorphisms

There are different reasons why until very recently no other risk genes than the *HLA* risk gene were identified. The attributable risk of other risk alleles is very small. Another reason is the heterogeneity of the populations studied. Undoubtedly the most important factor is the lack of power.

To overcome the problem of statistical power, the International MS Genetics Consortium (IMSGC) performed a genome-wide association study (GWAS). In the screening phase of this study more than 300,000 single nucleotide polymorphisms (SNPs) were tested in 931 families. For the replication phase 110 SNPs were selected and genotyped in another set of more than 2,500 MS patients and controls. Finally, a combined analysis was performed.

For 17 SNPs located in 14 regions an association with susceptibility for MS was identified. Only two regions, *HLA-DR* and *IL2RA*, achieved genome-wide significance ($P < 5 \times 10^{-08}$).[52] For a third gene, *IL7R*, already singled out as a strong candidate gene,[53-56] convincing functional support was obtained.[58] Later genome-wide significance was established in a joint analysis of 11,019 patients and 13,616 controls.[59] It is of note that the three risk sequence-allelic variants are all common variants.

Follow-up studies and further genome-wide analyses have now provided genome-wide significant support for several non-HLA MS risk genes: *IL7R, IL2RA* and *KIF1B,*[60] *CLEC16A* (*KIAA0350*)[61-63], *CD226* (DNAX accessory molecule 1[DNAM-1]),[61] *CD58,*[63] *CD6,*[64] *IRF8 (ICSBP1),* [64] *TNFRSF1A,*[64] a locus on chromosome 12 most probably relating to the gene *CYP27B1,*[65, 66] *TYK2,*[67-69] *STAT3,*[70] *KIF21B,*[71, 72] and *TMEM39A.*[71]

From genes to pathogenesis

The identification of risk genes for MS can help us in understanding the biological pathways involved in MS. It is important to recognize that their individual contribution is at present of no clinical value, because each identified risk variant is a common variant in the general population and their contribution in MS risk is very small.

It should be stressed that all associations discovered are at the SNP level. Sometimes they are situated in intronic areas or inside the gene desserts. For most of the SNPs a causative pathway remains to be determined. Most of the above immune related SNPs (see also the discussion of this thesis) have overlapping associations with autoimmune conditions such as rheumatoid arthritis, inflammatory bowel disease, SLE, Type 1 diabetes and thyroid autoimmunity (Table 1). This suggests that in MS and other autoimmune diseases, common pathways are involved in pathogenesis.

Table 1 Overlapping associations of risk genes for MS with other autoimmune diseases

Gene	Chr	Function	Disease	Reference
IL7R	5	Homeostasis of the memory T-cell pool,	T1D, chronic inflammatory arthropathies	73,74
IL2R	10	Regulation of T-cells	T1D, GD RA, JIA	75-78
CLEC16A	16	Provides signals for decisions between tolerance and immunity	Addison's disease, T1D, CD, JIA, RA, AITD	60,79-82
CD226	18	Adhesion and co-stimulation T-cells	T1D, RA, AITD CED, Wegener's granulomatosis	73,83-89
CYP27B1	12	Hydroxylates 25-hydroxyvitamin D into the bioactive form	T1D, Addison's disease	90-92
TNFRSF1A	12	Influences the TNFα pathway	Tumor necrosis factor associated periodic syndrome (TRAPS), IBD, persistent palindromic rheumatism	93-95
CD58	1	Influences T-cell proliferation and differentiation	RA, chronic inflammatory polyneuropathies?	96, 97
TYK2	19	Signalling by type I interferons and induction of Th1 cell differentiation upon antigen stimulation of dendritic cells	SLE	98
STAT3	17	Involved in multiple pathways and functions, including the Jak-STAT pathway, neuron axonal guidance, apoptosis, activation of immune responses and Th17 cell differentiation	CD, UC, hyperimmunoglobulin E recurrent infection syndrome (HIES)	99-102
KIF21B	1	Involved in axonal transport, also expressed in a variety of autoimmune cells	IBD	99

Abbreviations: AITD, autoimmune thyroid disease; CED, celiac disease; CD, Crohn's disease; GD, Graves'disease; IBD, inflammatory bowel disease; JIA, juvenile idiopathic arthritis; RA, rheumatoid arthritis; TID, Type 1 diabetes; UC, ulcerative colitis.

Scope of this thesis

The objective of this thesis was to find new risk alleles for MS. This may finally result in a better understanding of the pathogenesis of MS. Knowledge of MS disease pathways can direct strategies for prevention, diagnosis and therapy. In our study, we included MS patients from a genetically isolated population in the southwest of the Netherlands. We followed this strategy because of the relative genetic homogeneity of an isolated population and because relationships are known between patients. In **chapter 2** we assessed whether MS patients from this population were more often related to each other compared to controls from the same population. We investigated the parental relationship of MS patients using extensive genealogical information available from the Genetic Research in Isolated Populations (GRIP) program in **chapter 3**. In **chapter 4**, the results of two replication studies of the 17 by the IMSGC identified risk SNPs are described. In **chapter 4.1**, we verified the association of the 17 risk SNPs in MS patients and controls from the genetically isolated population. The second replication study (**chapter 4.2**) was performed in three independent cohorts: from the Dutch genetically isolated population, from the Dutch general population, and from the Canadian Collaborative Project on the Genetic Susceptibility to MS. The results were pooled with those of a recently published Australian replication study and with those of the original IMSGC study. Finally, we conducted a GWAS in the isolate and replicated the results in four independent cohorts, as reported in **chapter 5**. In the last chapter (**chapter 6**), we reflect our main findings, discuss methodological issues, speculate on the implications of our results and propose future studies.

References

1. Lopez-Diego RS, Weiner HL. Novel therapeutic strategies for multiple sclerosis--a multifaceted adversary. Nat Rev Drug Discov 2008;7(11):909-25.
2. Compston A, Coles A. Multiple sclerosis. Lancet 2002;359(9313):1221-31.
3. Orton SM, Herrera BM, Yee IM, et al. Sex ratio of multiple sclerosis in Canada: a longitudinal study. Lancet Neurol 2006;5(11):932-6.
4. Acheson ED, Bachrach CA, Wright FM. Some comments on the relationship of the distribution of multiple sclerosis to latitude, solar radiation, and other variables. Acta Psychiatr Scand 1960;35(147):132-47.
5. Miller DH, Hammond SR, McLeod JG, Purdie G, Skegg DC. Multiple sclerosis in Australia and New Zealand: are the determinants genetic or environmental? J Neurol Neurosurg Psychiatry 1990;53(10):903-5.
6. Kurtzke JF, Beebe GW, Norman JE, Jr. Epidemiology of multiple sclerosis in U.S. veterans: 1. Race, sex, and geographic distribution. Neurology 1979;29(9 Pt 1):1228-35.
7. Hammond SR, English DR, McLeod JG. The age-range of risk of developing multiple sclerosis: evidence from a migrant population in Australia. Brain 2000;123 (Pt 5):968-74.
8. Willer CJ, Dyment DA, Risch NJ, Sadovnick AD, Ebers GC. Twin concordance and sibling recurrence rates in multiple sclerosis. Proc Nat Acad Sci U S A 2003;100(22):12877-82.
9. Willer CJ, Dyment DA, Sadovnick AD, Rothwell PM, Murray TJ, Ebers GC. Timing of birth and risk of multiple sclerosis: population based study. BMJ 2005;330(7483):120.
10. Grytten N, Glad SB, Aarseth JH, Nyland H, Midgard R, Myhr KM. A 50-year follow-up of the incidence of multiple sclerosis in Hordaland County, Norway. Neurology 2006;66(2):182-6.
11. Alonso A, Hernan MA. Temporal trends in the incidence of multiple sclerosis: a systematic review. Neurology 2008;71(2):129-35.
12. Herrera BM, Ramagopalan SV, Lincoln MR, et al. Parent-of-origin effects in MS: observations from avuncular pairs. Neurology 2008;71(11):799-803.
13. Sadovnick AD, Baird PA, Ward RH. Multiple sclerosis: updated risks for relatives. Am J Med Genet 1988;29(3):533-41.
14. Sadovnick AD. Familial recurrence risks and inheritance of multiple sclerosis. Curr Opin Neurol Neurosurg 1993;6(2):189-94.
15. Carton H, Vlietinck R, Debruyne J, et al. Risks of multiple sclerosis in relatives of patients in Flanders, Belgium. J Neurol Neurosurg Psychiatry 1997;62(4):329-33.
16. Robertson NP, Fraser M, Deans J, Clayton D, Walker N, Compston DA. Age-adjusted recurrence risks for relatives of patients with multiple sclerosis. Brain 1996;119 (Pt 2):449-55.
17. Ebers GC, Bulman DE, Sadovnick AD, et al. A population-based study of multiple sclerosis in twins. N Engl J Med 1986;315(26):1638-42.
18. Mumford CJ, Wood NW, Kellar-Wood H, Thorpe JW, Miller DH, Compston DA. The British Isles survey of multiple sclerosis in twins. Neurology 1994;44(1):11-5.
19. Hansen T, Skytthe A, Stenager E, Petersen HC, Bronnum-Hansen H, Kyvik KO. Concordance for multiple sclerosis in Danish twins: an update of a nationwide study. Mult Scler 2005;11(5):504-10.
20. Sadovnick AD, Armstrong H, Rice GP, et al. A population-based study of multiple sclerosis in twins: update. Ann Neurol 1993;33(3):281-5.
21. Ebers GC, Sadovnick AD, Risch NJ. A genetic basis for familial aggregation in multiple sclerosis. Canadian Collaborative Study Group. Nature 1995;377(6545):150-1.
22. Ebers GC, Yee IM, Sadovnick AD, Duquette P. Conjugal multiple sclerosis: population-based prevalence and recurrence risks in offspring. Canadian Collaborative Study Group. Ann Neurol 2000;48(6):927-31.
23. Sadovnick AD, Ebers GC, Dyment DA, Risch NJ. Evidence for genetic basis of multiple sclerosis. The Canadian Collaborative Study Group. Lancet 1996;347(9017):1728-30.
24. Ebers GC, Sadovnick AD, Dyment DA, Yee IM, Willer CJ, Risch N. Parent-of-origin effect in multiple sclerosis: observations in half-siblings. Lancet 2004;363(9423):1773-4.
25. Lindsey JW. Familial recurrence rates and genetic models of multiple sclerosis. Am J Med Genet A 2005;135(1):53-8.

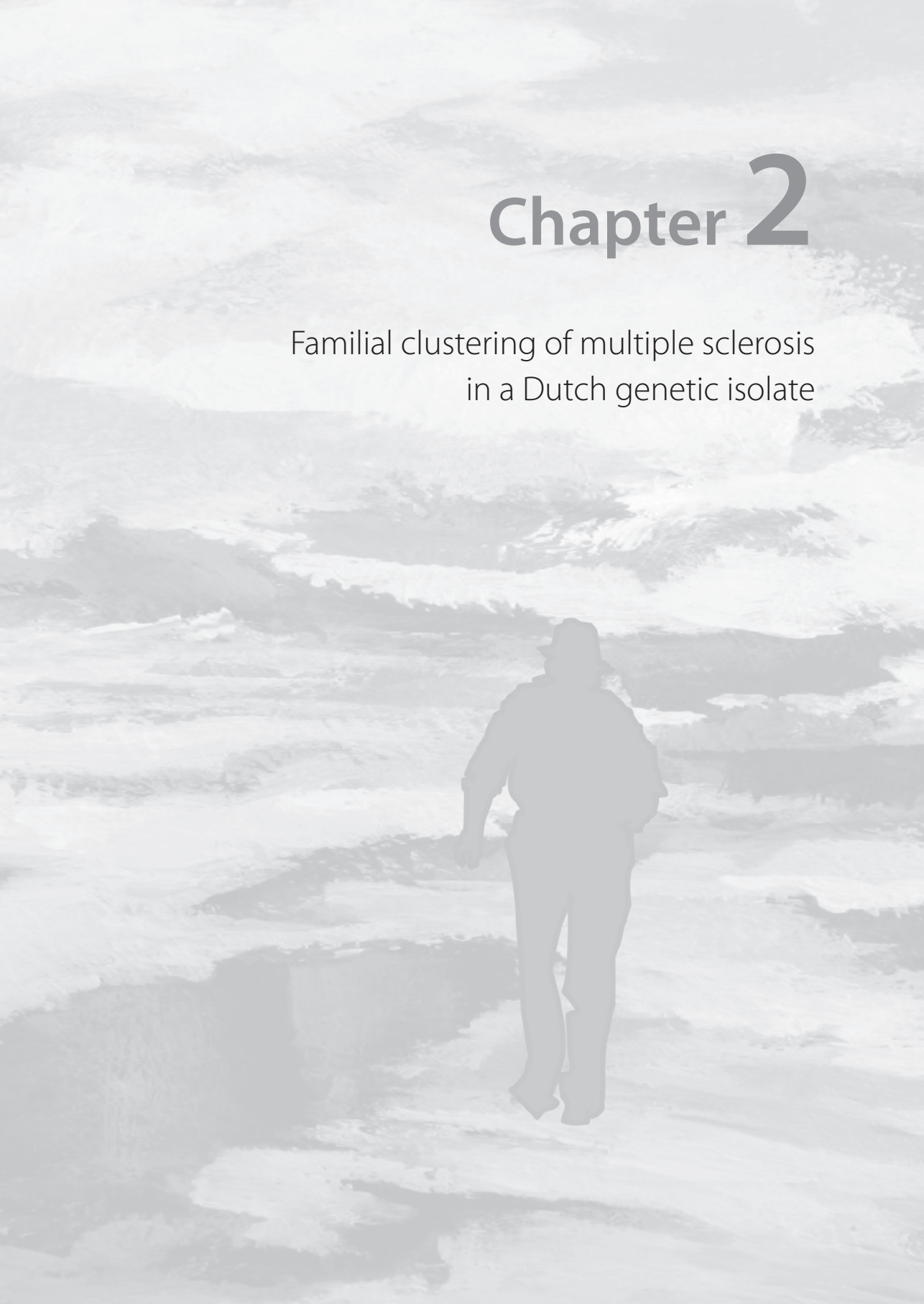
26. Yang Q, Khoury MJ, Friedman J, Little J, Flanders WD. How many genes underlie the occurrence of common complex diseases in the population? Int J Epidemiol 2005;34(5):1129-37.
27. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461(7265):747-53.
28. Kantarci OH, Barcellos LF, Atkinson EJ, et al. Men transmit MS more often to their children vs women: the Carter effect. Neurology 2006;67(2):305-10.
29. Herrera BM, Ramagopalan SV, Orton S, et al. Parental transmission of MS in a population-based Canadian cohort. Neurology 2007;69(12):1208-12.
30. Kantarci OH, Spurkland A. Parent of origin in multiple sclerosis: understanding inheritance in complex neurologic diseases. Neurology 2008;71(11):786-7.
31. Soejima H, Wagstaff J. Imprinting centers, chromatin structure, and disease. J Cell Biochem 2005;95(2):226-33.
32. Jersild C, Svejgaard A, Fog T. HL-A antigens and multiple sclerosis. Lancet 1972;1(7762):1240-1.
33. Naito S, Namerow N, Mickey MR, Terasaki PI. Multiple sclerosis: association with HL-A3. Tissue antigens 1972;2(1):1-4.
34. Miretti MM, Walsh EC, Ke X, et al. A high-resolution linkage-disequilibrium map of the human major histocompatibility complex and first generation of tag single-nucleotide polymorphisms. Am J Hum Genet 2005;76(4):634-46.
35. Lincoln MR, Montpetit A, Cader MZ, et al. A predominant role for the HLA class II region in the association of the MHC region with multiple sclerosis. Nat Genet 2005;37(10):1108-12.
36. Oksenberg JR, Barcellos LF, Cree BA, et al. Mapping multiple sclerosis susceptibility to the HLA-DR locus in African Americans. Am J Hum Genet 2004;74(1):160-7.
37. Marrosu MG, Murru MR, Costa G, Murru R, Muntoni F, Cucca F. DRB1-DQA1-DQB1 loci and multiple sclerosis predisposition in the Sardinian population. Hum Mol Genet 1998;7(8):1235-7.
38. Marrosu MG, Murru MR, Costa G, et al. Multiple sclerosis in Sardinia is associated and in linkage disequilibrium with HLA-DR3 and -DR4 alleles. Am J Hum Genet 1997;61(2):454-7.
39. Dyment DA, Herrera BM, Cader MZ, et al. Complex interactions among MHC haplotypes in multiple sclerosis: susceptibility and resistance. Hum Mol Genet 2005;14(14):2019-26.
40. Barcellos LF, Sawcer S, Ramsay PP, et al. Heterogeneity at the HLA-DRB1 locus and risk for multiple sclerosis. Hum Mol Genet 2006;15(18):2813-24.
41. Ramagopalan SV, Morris AP, Dyment DA, et al. The inheritance of resistance alleles in multiple sclerosis. PLoS Genet 2007;3(9):1607-13.
42. Modin H, Olsson W, Hillert J, Masterman T. Modes of action of HLA-DR susceptibility specificities in multiple sclerosis. Am J Hum Genet 2004;74(6):1321-2.
43. Ramagopalan SV, Ebers GC. Epistasis: multiple sclerosis and the major histocompatibility complex. Neurology 2009;72(6):566-7.
44. Brynedal B, Duvefelt K, Jonasdottir G, et al. HLA-A confers an HLA-DRB1 independent influence on the risk of multiple sclerosis. PLoS ONE 2007;2(7):e664.
45. Chao MJ, Ramagopalan SV, Herrera BM, et al. Epigenetics in multiple sclerosis susceptibility: difference in transgenerational risk localizes to the major histocompatibility complex. Hum Mol Genet 2009;18(2):261-6.
46. Munger KL, Levin LI, Hollis BW, Howard NS, Ascherio A. Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. JAMA 2006;296(23):2832-8.
47. Nieves J, Cosman F, Herbert J, Shen V, Lindsay R. High prevalence of vitamin D deficiency and reduced bone mass in multiple sclerosis. Neurology 1994;44(9):1687-92.
48. Ramagopalan SV, Maugeri NJ, Handunnetthi L, et al. Expression of the multiple sclerosis-associated MHC class II Allele HLA-DRB1*1501 is regulated by vitamin D. PLoS Genet 2009;5(2):e1000369.
49. Nielsen TR, Rostgaard K, Nielsen NM, et al. Multiple sclerosis after infectious mononucleosis. Arch Neurol 2007;64(1):72-5.
50. Levin LI, Munger KL, Rubertone MV, et al. Temporal relationship between elevation of epstein-barr virus antibody titers and initial onset of neurological symptoms in multiple sclerosis. JAMA 2005;293(20):2496-500.
51. Nielsen TR, Rostgaard K, Askling J, et al. Effects of infectious mononucleosis and HLA-DRB1*15 in multiple sclerosis. Mult Scler 2009;15(4):431-6.

52. Hafler DA, Compston A, Sawcer S, et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 2007;357(9):851-62.
53. Sundvall M, Jirholt J, Yang HT, et al. Identification of murine loci associated with susceptibility to chronic experimental autoimmune encephalomyelitis. Nat Genet 1995;10(3):313-7.
54. Ramanathan M, Weinstock-Guttman B, Nguyen LT, et al. In vivo gene expression revealed by cDNA arrays: the pattern in relapsing-remitting multiple sclerosis patients compared with normal subjects. J Neuroimmunol 2001;116(2):213-9.
55. Bomprezzi R, Ringner M, Kim S, et al. Gene expression profile in multiple sclerosis patients and healthy controls: identifying pathways relevant to disease. Hum Mol Genet 2003;12(17):2191-9.
56. Zhang Z, Duvefelt K, Svensson F, et al. Two genes encoding immune-regulatory molecules (LAG3 and IL7R) confer susceptibility to multiple sclerosis. Genes Immun 2005;6(2):145-52.
57. 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(9):1083-91.
58. 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(9):1108-13.
59. Refining genetic associations in multiple sclerosis. Lancet Neurol 2008;7(7):567-9.
60. Aulchenko YS, Hoppenbrouwers IA, Ramagopalan SV, et al. Genetic variation in the KIF1B locus influences susceptibility to multiple sclerosis. Nat Genet 2008;40(12):1402-3.
61. The expanding genetic overlap between multiple sclerosis and type 1 diabetes. Genes Immun 2009;10(1):11-4.
62. Rubio JP, Stankovich J, Field J, et al. Replication of KIAA0350, IL2RA, RPL5 and CD58 as multiple sclerosis susceptibility genes in Australians. Genes Immun 2008;9(7):624-30.
63. Hoppenbrouwers IA, Aulchenko YS, Janssens AC, et al. Replication of CD58 and CLEC16A as genome-wide significant risk genes for multiple sclerosis. J Hum Genet 2009;54(11):676-80.
64. 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(7):776-82.
65. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. Nat Genet 2009;41(7):824-8.
66. Sundqvist E, Baarnhielm M, Alfredsson L, Hillert J, Olsson T, Kockum I. Confirmation of association between multiple sclerosis and CYP27B1. Eur J Hum Genet.
67. Burton PR, Clayton DG, Cardon LR, et al. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. Nat Genet 2007;39(11):1329-37.
68. Ban M, Goris A, Lorentzen AR, et al. Replication analysis identifies TYK2 as a multiple sclerosis susceptibility factor. Eur J Hum Genet 2009;17(10):1309-13.
69. Mero IL, Lorentzen AR, Ban M, et al. A rare variant of the TYK2 gene is confirmed to be associated with multiple sclerosis. Eur J Hum Genet;18(4):502-4.
70. Jakkula E, Leppa V, Sulonen AM, et al. Genome-wide association study in a high-risk isolate for multiple sclerosis reveals associated variants in STAT3 gene. Am J Hum Genet;86(2):285-91.
71. Comprehensive follow-up of the first genome-wide association study of multiple sclerosis identifies KIF21B and TMEM39A as susceptibility loci. Hum Mol Genet;19(5):953-62.
72. Goris A, Boonen S, D'Hooghe M B, Dubois B. Replication of KIF21B as a susceptibility locus for multiple sclerosis. J Med Genet 2010, in press.
73. Todd JA, Walker NM, Cooper JD, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet 2007;39(7):857-64.
74. O'Doherty C, Alloza I, Rooney M, Vandenbroeck K. IL7RA polymorphisms and chronic inflammatory arthropathies. Tissue antigens 2009;74(5):429-31.
75. Vella A, Cooper JD, Lowe CE, et al. Localization of a type 1 diabetes locus in the IL2RA/CD25 region by use of tag single-nucleotide polymorphisms. Am J Hum Genet 2005;76(5):773-9.
76. Brand OJ, Lowe CE, Heward JM, et al. Association of the interleukin-2 receptor alpha (IL-2Ralpha)/CD25 gene region with Graves' disease using a multilocus test and tag SNPs. Clin Endocrinol 2007;66(4):508-12.
77. Kurreeman FA, Daha NA, Chang M, et al. Association of IL2RA and IL2RB with rheumatoid arthritis: a replication study in a Dutch population. Ann Rheum Dis 2009;68(11):1789-90.

78. Hinks A, Ke X, Barton A, et al. Association of the IL2RA/CD25 gene with juvenile idiopathic arthritis. Arthritis Rheum 2009;60(1):251-7.
79. Skinningsrud B, Husebye ES, Pearce SH, et al. Polymorphisms in CLEC16A and CIITA at 16p13 are associated with primary adrenal insufficiency. J Clin Endocrinol Metab 2008;93(9):3310-7.
80. Marquez A, Varade J, Robledo G, et al. Specific association of a CLEC16A/KIAA0350 polymorphism with NOD2/CARD15(-) Crohn's disease patients. Eur J Hum Genet 2009;17(10):1304-8.
81. Skinningsrud B, Lie BA, Husebye ES, et al. A CLEC16A variant confers risk for juvenile idiopathic arthritis and anti-CCP negative rheumatoid arthritis. Ann Rheum Dis 2009.
82. Awata T, Kawasaki E, Tanaka S, et al. Association of type 1 diabetes with two Loci on 12q13 and 16p13 and the influence coexisting thyroid autoimmunity in Japanese. J Clin Endocrinol Metab 2009;94(1):231-5.
83. Douroudis K, Nemvalts V, Rajasalu T, Kisand K, Uibo R. The CD226 gene in susceptibility of type 1 diabetes. Tissue antigens 2009;74(5):417-9.
84. Smyth DJ, Plagnol V, Walker NM, et al. Shared and distinct genetic variants in type 1 diabetes and celiac disease. N Engl J Med 2008;359(26):2767-77.
85. Hafler JP, Maier LM, Cooper JD, et al. CD226 Gly307Ser association with multiple autoimmune diseases. Genes Immun 2009;10(1):5-10.
86. Seldin MF, Amos CI. Shared susceptibility variations in autoimmune diseases: a brief perspective on common issues. Genes Immun 2009;10(1):1-4.
87. Maiti AK, Kim-Howard X, Viswanathan P, et al. Non-synonymous variant (Gly307Ser) in CD226 is associated with susceptibility to multiple autoimmune diseases. Rheumatology 2010;49(7):1239-44.
88. Tan RJ, Gibbons LJ, Potter C, et al. Investigation of rheumatoid arthritis susceptibility genes identifies association of AFF3 and CD226 variants with response to anti-tumour necrosis factor treatment. Ann Rheum Dis 2010;69(6):1029-35.
89. Wieczorek S, Hoffjan S, Chan A, et al. Novel association of the CD226 (DNAM-1) Gly307Ser polymorphism in Wegener's granulomatosis and confirmation for multiple sclerosis in German patients. Genes Immun 2009;10(6):591-5.
90. Bailey R, Cooper JD, Zeitels L, et al. Association of the vitamin D metabolism gene CYP27B1 with type 1 diabetes. Diabetes 2007;56(10):2616-21.
91. Fichna M, Zurawek M, Januszkiewicz-Lewandowska D, et al. Association of the CYP27B1 C(-1260)A Polymorphism with Autoimmune Addison's Disease. Exp Clin Endocrinol Diabetes 2010;118(8):544-9.
92. Jennings CE, Owen CJ, Wilson V, Pearce SH. A haplotype of the CYP27B1 promoter is associated with autoimmune Addison's disease but not with Graves' disease in a UK population. J Mol Endocrinol 2005;34(3):859-63.
93. Hoffmann LA, Lohse P, Konig FB, Feneberg W, Hohlfeld R, Kumpfel T. TNFRSF1A R92Q mutation in association with a multiple sclerosis-like demyelinating syndrome. Neurology 2008;70(13 Pt 2):1155-6.
94. Lappalainen M, Halme L, Turunen U, et al. Association of IL23R, TNFRSF1A, and HLA-DRB1*0103 allele variants with inflammatory bowel disease phenotypes in the Finnish population. Inflamm Bowel Dis 2008;14(8):1118-24.
95. Lu LY, Keng HM, Chu JJ, Lin XT, Hsu CM, Sung PK. TNF receptor I polymorphism is associated with persistent palindromic rheumatism. Scand J Rheumatol 2007;36(4):278-84.
96. Van Rhijn I, Van den Berg LH, Bosboom WM, Otten HG, Logtenberg T. Expression of accessory molecules for T-cell activation in peripheral nerve of patients with CIDP and vasculitic neuropathy. Brain 2000;123 (Pt 10):2020-9.
97. Raychaudhuri S, Thomson BP, Remmers EF, et al. Genetic variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk. Nat Genet 2009;41(12):1313-8.
98. Suarez-Gestal M, Calaza M, Endreffy E, et al. Replication of recently identified systemic lupus erythematosus genetic associations: a case-control study. Arthritis Res Ther 2009;11(3):R69.
99. Barrett JC, Hansoul S, Nicolae DL, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet 2008;40(8):955-62.
100. Holland SM, DeLeo FR, Elloumi HZ, et al. STAT3 mutations in the hyper-IgE syndrome. New Engl J Med 2007;357(16):1608-19.
101. Minegishi Y, Saito M, Tsuchiya S, et al. Dominant-negative mutations in the DNA-binding domain of STAT3 cause hyper-IgE syndrome. Nature 2007;448(7157):1058-62.
102. Cenit MC, Alcina A, Marquez A, et al. STAT3 locus in inflammatory bowel disease and multiple sclerosis susceptibility. Genes Immun;11(3):264-8.

Chapter 2

Familial clustering of multiple sclerosis in a Dutch genetic isolate



Abstract

Multiple sclerosis (MS) is a complex disease with a substantial, yet poorly identified genetic influence. We estimated the pattern of familial aggregation of MS in a recent genetically isolated population in the Netherlands. Forty-eight MS patients were identified. Their relationship was evaluated by tracing extended pedigrees, making use of municipal and church records. Of the 48 MS patients, 24 could be linked to a common ancestor in 14 generations. However, multiple relationships exist between patients and to take these into account we calculated inbreeding and kinship coefficients. We found that MS patients from the isolate were significantly more often related to each other and significantly more often inbred than a non-MS control group, drawn from the same isolate. There was no clustering of Type 1 diabetes and autoimmune thyroid diseases in families of MS patients from this isolate. Finally, HLA typing was performed. Although there was a trend towards a higher prevalence of the *HLA DRB1*15* allele in patients compared to controls, differences did not reach significance. This study suggests familial aggregation in the genetically isolated population. The high level of inbreeding makes this population valuable for finding novel genes involved in MS.

Introduction

Although familial aggregation of multiple sclerosis (MS) has long been recognized,[1,2] thus far, the mode of inheritance remains unclear.

Some studies suggested that a single genetic abnormality might be sufficiently severe to result in a form of inheritance following a Mendelian pattern. An autosomal dominant mode of inheritance of MS with reduced penetrance has been proposed for one family.[3] Autosomal recessive inheritance has been suggested in other studies,[4,5] but either mode of inheritance remains unproven. In the majority of patients, MS is probably the outcome of the additive effects of multiple genes with complex gene-gene and gene-environment interactions.

Data in most studies on family aggregation of MS have been based on interview-based information on family history. For distantly-related persons, data may be difficult to reproduce in this way. Availability of genealogical information over multiple generations is needed to unravel distant relationships. Such information is usually available in isolated populations, eg, in Italy,[6] Iceland,[7] and Northern Sweden.[8] Studying familial aggregation and constructing extended and multigenerational pedigrees becomes possible when well-documented genealogical records are available.

The presence of inbreeding, which may reveal recessive forms of disease, can be more easily assessed in the extended pedigrees than by patient interview only, especially when patients are related more than a few generations ago.

Of interest is the recent finding of clustering of MS with other autoimmune disorders, including Type 1 diabetes and thyroid disease.[9-11] Another striking feature in the Sardinian population is that the risk of MS is not associated with the *HLA-DRB1*15* allele,[12] but rather with the *HLA-DRB1*03* and-*DRB1*04* alleles.[13,14] The *HLA-DRB1*03* allele is also associated with an increased risk of Type 1 diabetes.[10,15] This association may partly explain the familial clustering with Type 1 diabetes, although other genetic factors also appear to play a role.[16]

The purpose of this study was to investigate the pattern of familial aggregation of MS in a genetically isolated population in the Netherlands. We studied both clustering of MS patients in families and clustering of MS, Type 1 diabetes and autoimmune thyroid disease. We tested if the clinical characteristics of the MS patients from the isolate were different from patients from outside the isolate. To determine if the *HLA* distribution in this population is comparable with the *HLA* distribution in a control group from the same isolate we performed *HLA* typing.

Material and methods

Study population

MS patients were identified in a genetically isolated community in the southwest of the Netherlands. This population is studied as part of the 'Genetic Research in Isolated Populations' (GRIP) program. About 150 individuals founded the GRIP population around 1750. Until recently, this population was characterized

by minimal inward and outward migration. During the last two centuries, this population underwent an exponential growth. Currently, the population consists of >20,000 individuals.[17]

Patients

Patients were traced through local general practitioners. The Medical Ethics Committee of the Erasmus Medical Center Rotterdam, the Netherlands, approved the study protocol.

Fifty-four individuals, diagnosed with multiple sclerosis, were identified in this genetically isolated population. Forty-five patients were traced by their general practitioner. Six patients were presented by their relative with MS. Two patients were traced by the genealogist and one patient, living in the area, by a routine visit at the outpatient clinic of our MS Centre.

Of the 54 identified individuals with MS, 50 subjects (15 male, 35 female) agreed to participate in the study. After written informed consent was given, all medical records, as well as MRI scans of brain and spinal cord, if available, were reviewed by at least three independent neurologists to confirm diagnosis. In case of discrepancy, the final diagnosis was established at a consensus meeting.

Of the 50 participating patients, 41 (85%) subjects fulfilled the criteria of Poser[18] for definite MS. Six patients fulfilled the criteria for clinically probable and one for laboratory supported MS. Two male patients were excluded from this study, one because he was diagnosed with CADASIL instead of MS, and the other because he did not fulfill MS criteria after thorough revision of the clinical data. Therefore, 48 MS patients (13 male, 35 female) were analysed in this study (Figure 1). To test if age at onset and gender distribution of these 48 MS patients were comparable to those of a standard MS population, we compared the characteristics of the 48 MS patients with those of a group of 73 sporadic Dutch MS patients, from outside the isolate.[19]

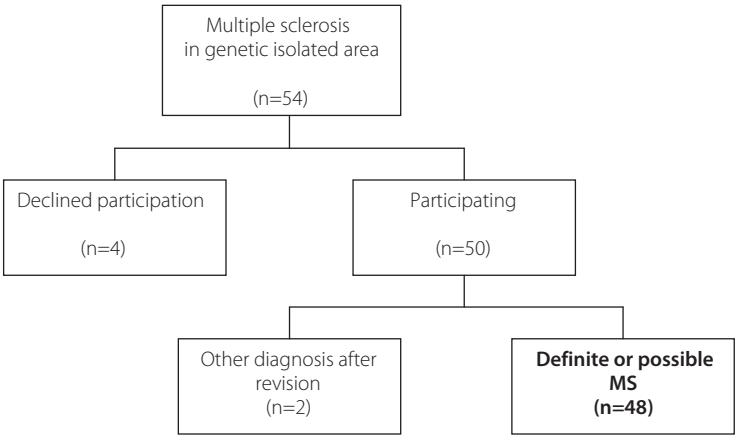


Figure 1 Ascertainment of 48 MS patients in the genetically isolated population.

Genealogical information

Genealogical data, comprising name, date and place of birth and death of relatives were obtained from the participants at home interviews. By means of local municipal and church registers, and a large genealogical database, containing information for almost 70,000 individuals, this information was extended up to 14 generations. This information was used to study familial clustering of MS in the genetic isolate.

Family history

In all patients, we carried out a questionnaire on medical history, family history and co-occurrence of MS and other diseases in relatives of MS patients. Special emphasis was paid on the co-occurrence of Type 1 diabetes or autoimmune thyroid disease in MS patients and their relatives. We defined Type 1 diabetes in accordance with criteria reported by the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus.[20] Diagnosis of autoimmune thyroid disease in MS patients and first-degree relatives of MS patients was verified by medication intake, cause of thyroid dysfunction, and/ or by information of the general practitioner. Evident non-autoimmune causes of hypothyroidism and hyperthyroidism were excluded.

HLA typing

Venous blood (30 mL) was obtained for DNA isolation from each patient. Venous blood was also obtained from a control set, consisting of 39 spouses or close friends of 39 MS patients. Genomic DNA was extracted from peripheral blood leucocytes, according to a standard protocol.[21]

HLA-DRB1 typing was performed at the two-digit level using a commercially available typing system, in which exon 2 of the *HLA-DRB1* gene is amplified and the amplified product is analysed with allele-specific probes in a line probe assay (INNO-LiPA, Innogenetics, Ghent, Belgium). This resulted in the determination of the allele groups DRB1*1, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18.[22]

Statistical analysis

We performed general descriptive statistics using χ^2 statistics for dichotomous variables and a Student's *t*-test for continuous variables. Kinship and inbreeding coefficients were calculated, using PEDIG software,[23] based on the genealogical information of the total population, consisting of 74,461 subjects.

Genealogical relationships were expressed as the pairwise kinship coefficient. This is the probability that a randomly drawn allele from one person is identical by descent with a randomly drawn allele at the same locus from the other person. If individuals are related to each other, the kinship coefficient is higher than zero, if not, the kinship coefficient is equal to zero.

Inbreeding coefficients were also calculated. The inbreeding coefficient is the probability that two alleles at the same locus in an individual are identical by descent and represents the degree of consanguinity between parents of an individual. The presence of consanguinity in affected people may point to the presence of a recessive mutation underlying the disease. Inbred individuals have an

inbreeding coefficient higher than zero, while the inbreeding coefficient is zero for those who are not inbred.

To evaluate familial clustering of MS patients, we considered only the patients who were probands, ascertained through the general practitioner. We excluded the patients that were identified via their relatives or by the genealogist. We also excluded MS patients who were born outside the isolate. Genealogical information for this population is only available for people born in the isolate. The inclusion of patients who were born outside the isolate could result in false low kinship coefficients, because of the lack of genealogical information for these subjects. Thus, only 17 MS patients were included for calculation of kinship and inbreeding coefficients. Of these 17 probands, 13 probands were members of the large pedigree.

Further, we studied whether kinship and inbreeding values in the 17 MS patients deviate from what might be expected in a random sample of individuals in the isolated population. The random samples of individuals, drawn from the genealogical database, were matched by age, sex and place of birth with the 17 MS patients. We performed 1,000 replicas of this sampling in order to obtain a null distribution for the kinship and inbreeding coefficients. *P*-values were computed using bootstrapping.

Results

Patients

In total, 48 MS patients participated in the study. General descriptives of the participants are presented in Table 1. We found no significant differences between age of symptom onset and gender of the 48 MS patients and age of symptom onset and gender of a group of 73 sporadic Dutch MS patients, from outside the isolate.[19]

Table 1 Description of the total study group, MS patients that could be and could not be linked to a common ancestor

Variables	MS patients (n=48)	MS patients with common ancestor (n=24)	MS patients, no common ancestor (n=24)
Age			
Of onset symptoms	32 ± 11	29 ± 9	35 ± 12
Of diagnosis	38 ± 12	36 ± 11	41 ± 12
Sex			
Female (%)	73	79	67
Clinical phenotype at onset			
Relapse-onset (%)	85	83	88
Primary progressive (%)	15	17	13

Dichotomous variables are expressed as percentage. Values of continuous variables are expressed as mean ± standard deviation.

Genealogy

Eleven (23%) of the 48 MS patients reported that they did have a first- or second-degree relative with MS. After an extensive genealogical analysis, we found that 24 of 48 MS patients (50%) could be linked to one common ancestor in 14 generations. Figure 2 shows the genealogical lineages of the patients

of the pedigree. Numerous connections exist between these patients via multiple ancestors. General descriptions for the 24 MS patients, who are in the pedigree, are presented in Table 1 as well as the general characteristics for the 24 MS patients who are not in the pedigree.

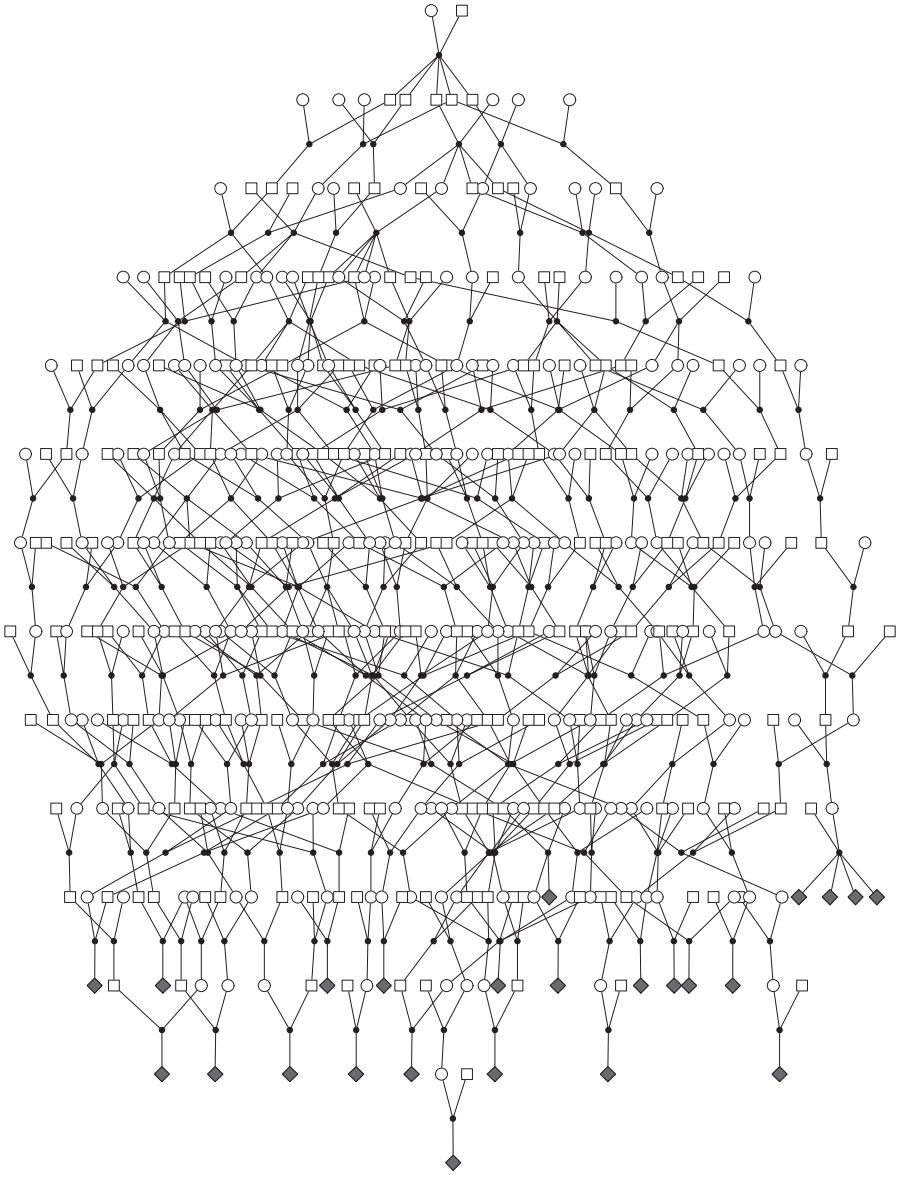


Figure 2 Pedigree of 24 MS patients and their relationships to a common ancestor. White solid squares indicate males, white solid circles indicate females. Black dots indicate meiosis. Diamonds indicate the MS patients.

The average age of onset of symptoms for the 24 related patients was lower than for the 24 non-related patients, with borderline significance (P =0.05). Clinical phenotype, regarding age of diagnosis, gender and disease course,[24] did not show a significant difference between patients of the pedigree and patients that could not be linked to a common ancestor. When we compared the 24 related MS patients with the 73 sporadic Dutch MS patients,[19] from outside the isolate, we could not find a significant difference between age of symptom onset (P =0.57). In addition, gender was equally distributed amongst these two groups.

We tested, by calculating kinship coefficients, whether the 17 MS probands born in the isolate were more closely related than controls, who were matched for sex, age and place of birth and who were randomly drawn from the genealogical database. This analysis showed that 95% of MS patients were related to each other, compared to 34% expected in the random group (P <0.0001) (Table 2). In addition, significantly more MS patients could be linked to a common ancestor (88%) than expected by chance (49%, P <0.0001).

Table 2 Distribution of kinship coefficients for patients and controls

Kinship Coefficient	MS probands	Controls
> 0	129 (95)*	45,803 (34)
0	7 (5)*	90,197 (66)
Total number of pairs	136 (100)	136,000 (100)

All values are absolute numbers with percentages of the total number of pairs in brackets.
* Significantly different from controls (P <0.0001).

We further calculated the inbreeding coefficients for patients and randomly selected controls. The percentage of MS patients with inbreeding was 82%, compared to 36% for the control group, which was significantly higher (P <0.0001) (Table 3).

Table 3 Distribution of inbreeding coefficients for patients and controls

Inbreeding Coefficient	MS probands	Controls
> 0	14 (82)*	6,172 (36)
0	3 (18)*	10,828 (64)
Total numbers	17(100)	17,000 (100)

All values are absolute numbers with percentages in brackets.
* Significantly different from controls (P <0.0001).

Type 1 diabetes and thyroid disease

None of the 48 MS patients was also diagnosed with Type 1 diabetes. The 48 MS patients had a total of 361 first-degree relatives. Type 1 diabetes was reported in two of the 361 (0.6%) first-degree relatives, which is very similar to the prevalence of about 0.4% in the general Dutch population (www.diabetesfonds.nl).

Only one of the 48 MS patients was diagnosed with an autoimmune thyroid disease. Seven out of the total of 361 (1.9%) first-degree relatives had a thyroid disease, reportedly due to an autoimmune disease. Prevalence of thyroid disease seems to be similar to that observed in the English population (about 25%).[25] There is no information about prevalence of autoimmune thyroid disease in the Netherlands. Families with no co-occurrence of autoimmune diseases did not differ in clinical characteristics from families with co-occurrence of such ailments.

HLA typing

Figure 3 shows the HLA allele distribution in cases and controls. Allele and genotype proportions in MS patients and controls did not deviate from Hardy-Weinberg equilibrium (P =0.988 and 0.997 respectively). The *HLA-DRB1*15* allele was not found significantly more often in MS patients (28% of 96 alleles) than in controls from the isolate (18% of 78 alleles) (OR: 1.79; 95%: 0.86-3.71; P =0.12). Additionally, we did not find significant differences between cases and controls for the other *HLA-DRB1* alleles.

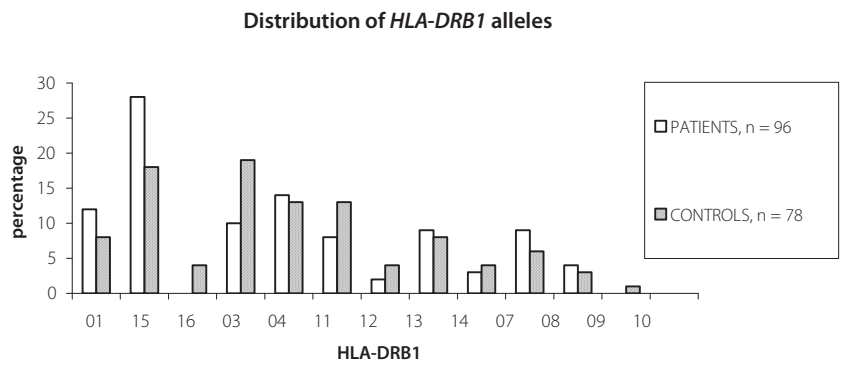


Figure 3 Distribution of *HLA-DRB1* alleles among MS patients and controls.

Discussion

In this study, we identified 48 MS patients in a genetically isolated area in the Netherlands. Twenty-four (50%) of the 48 patients could be linked to one common ancestor, thus providing the largest MS pedigree reported so far. There were no special clinical features seen in these patients, compared to the general population, including age of onset. Previous studies have shown that patients with evidence for strong genetic loading for MS tend to have earlier onset of the disease.[26] The age of symptom onset of MS patients who could be linked to a common ancestor was lower than the age of symptom onset of MS patients who could not be linked to a common ancestor, although just reaching borderline significance. Age of symptom onset did not reach significant differences between patients of the pedigree and a set of sporadic MS patients from our outpatient clinic. A slight skew towards females was found for patients who could be linked to a common ancestor (79%), not reaching significance.

An increased relation to a common ancestor, as well as kinship between patients indicates familial clustering of the disease. Familial clustering of MS was also found in a northern Swedish rural district[8] and in the isolated population of Sardinia.[27] Within the pedigree we observed a substantial amount of inbreeding, and we calculated that there were significantly more MS patients with inbreeding in this population, compared with controls from the same population. This could be in line with an autosomal recessive mode of inheritance, at least in a subset of patients. However, at this stage in our study, we cannot exclude the presence of susceptibility genes for which homozygosity shows the highest increase in risk. Both models may be detected by homozygosity mapping. Identification of the gene involved will ultimately reveal the true mode of inheritance.

We examined the co-occurrence of autoimmune disease in MS families, since this was an important feature in the Sardinian isolate,[10] and, to a lesser extend, also observed in the UK.[9] In total, 0.6% of all 361 first-degree relatives were reported to have Type 1 diabetes. In the Netherlands, a prevalence of about 0.4% for Type 1 diabetes mellitus is reported by the national diabetes fund (www.diabetesfonds.nl). Therefore, in contrast to the isolated population in Sardinia, the prevalence of Type 1 diabetes in the first degree relatives of the MS patients does not appear to be enhanced in this isolated population (P =0.6).

In total seven out of 361 (1.9%) first degree relatives were reported to have a thyroid autoimmune disease. A population-based study for the prevalence of thyroid disease in the north of England was carried out by Tunbridge *et al.*[25] The prevalence of thyroid disease was about 2.5% . Only patients with an abnormal function of the thyroid gland were included for calculation of the prevalence of thyroid disease and not only autoimmune thyroid diseases. By far the highest percentage of thyroid diseases with an abnormal function of the thyroid gland is due to autoimmune thyroid diseases in the UK and the Netherlands. Therefore, the percentage we observed in the isolated population does not appear to differ much from what is found in the general population. Thus, in contrast to MS families from the UK, the prevalence of autoimmune thyroid disease in the first degree relatives of the MS patients also does not seem to be high in this isolated population.

Several studies have shown association of MS with the *HLA-DRB1*15* allele. Also, in a large North American pedigree, linkage of this allele with another susceptibility locus on 12p12 has been suggested to be important for development of MS.[28] In our isolated population, although a trend towards a slightly higher prevalence of the *HLA-DRB1*15* allele was found, differences did not reach significance. This can probably be explained by the low numbers of patients and controls in whom *HLA* has been genotyped. The prevalence of *HLA-DRB1*15* is more similar to that in local controls. This finding is in accordance with findings in a genetically isolated population in Northern Sweden, where no significant sharing of the *HLA* region was found.[29] In the Sardinian genetic isolate, there was no association of the *HLA-DRB1*15* allele with the MS population. In that area *HLA-DRB1*03* and *HLA-DRB1*04* are over-represented along with an increased incidence of Type 1 diabetes.[10,14,15] The fact that we also did not find any indication for over-representation of HLA Class II loci in our population supports the view that other loci outside the MHC region are shared by individuals with MS in this population,[16] such as observed in the northern

Sweden isolate.[29] Here, a novel susceptibility gene for MS was suggested in chromosome 17p11. We conclude that extensive genealogical search in this genetic isolate revealed a large proportion of MS cases, which were linked in a genetic way. The extended pedigree may prove to be helpful for identifying the genes involved, using an autosomal recessive inheritance model. Given the lack of several special clinical characteristics in this MS population, the genetic pathways involved could have relevance for MS in the general population.

References

1. Dyment DA, Ebers GC, Sadovnick AD. Genetics of multiple sclerosis. Lancet Neurol 2004;3(2):104-10.
2. Compston A, Coles A. Multiple sclerosis. Lancet 2002;359(9313):1221-31.
3. Dyment DA, Cader MZ, Willer CJ, Risch N, Sadovnick AD, Ebers GC. A multigenerational family with multiple sclerosis. Brain 2002;125(Pt 7):1474-82.
4. Modin H, Masterman T, Thorlacius T, et al. Genome-wide linkage screen of a consanguineous multiple sclerosis kinship. Mult Scler 2003;9(2):128-34.
5. Sadovnick A, Yee I, Ebers G. Recurrence risks to sibs of MS index cases: impact of consanguineous matings. Neurology 2001;56:784-5.
6. Angius A, Melis PM, Morelli L, et al. Archival, demographic and genetic studies define a Sardinian sub-isolate as a suitable model for mapping complex traits. Hum Genet 2001;109(2):198-209.
7. Gudmundsson H, Gudbjartsson DF, Frigge M, Gulcher JR, Stefansson K. Inheritance of human longevity in Iceland. Eur J Hum Genet 2000;8(10):743-9.
8. Binzer M, Forsgren L, Holmgren G, Drugge U, Fredrikson S. Familial clustering of multiple sclerosis in a northern Swedish rural district. J Neurol Neurosurg Psychiatry 1994;57(4):497-9.
9. Broadley SA, Deans J, Sawcer SJ, Clayton D, Compston DA. Autoimmune disease in first-degree relatives of patients with multiple sclerosis. A UK survey. Brain 2000;123 (Pt 6):1102-11.
10. Marrosu MG, Cocco E, Lai M, Spinicci G, Pischedda MP, Contu P. Patients with multiple sclerosis and risk of type 1 diabetes mellitus in Sardinia, Italy: a cohort study. Lancet 2002;359(9316):1461-5.
11. Dorman JS, Steenkiste AR, Burke JP, Songini M. Type 1 diabetes and multiple sclerosis: together at last. Diabetes Care 2003;26(11):3192-3.
12. Olerup O, Hillert J. HLA class II-associated genetic susceptibility in multiple sclerosis: a critical evaluation. Tissue Antigens 1991;38(1):1-15.
13. Marrosu MG, Murru MR, Costa G, Murru R, Muntoni F, Cucca F. DRB1-DQA1-DQB1 loci and multiple sclerosis predisposition in the Sardinian population. Hum Mol Genet 1998;7(8):1235-7.
14. Marrosu MG, Murru MR, Costa G, et al. Multiple sclerosis in Sardinia is associated and in linkage disequilibrium with HLA-DR3 and -DR4 alleles. Am J Hum Genet 1997;61(2):454-7.
15. Lernmark A. Multiple sclerosis and type 1 diabetes: an unlikely alliance. Lancet 2002;359(9316):1450-1.
16. Marrosu MG, Motzo C, Murru R, et al. The co-inheritance of Type 1 Diabetes and Multiple Sclerosis in Sardinia cannot be explained by genotype variation in the HLA region alone. Hum Mol Genet 2004;13(23):2919-24.
17. Aulchenko YS, Heutink P, Mackay I, et al. Linkage disequilibrium in young genetically isolated Dutch population. Eur J Hum Genet 2004;12(7):527-34.
18. Poser CM, Paty DW, Scheinberg L, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol 1983;13(3):227-31.
19. Buljevac D, Flach HZ, Hop WC, et al. Prospective study on the relationship between infections and multiple sclerosis exacerbations. Brain 2002;125(Pt 5):952-60.
20. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care 1997;20(7):1183-97.
21. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988;16(3):1215.
22. Thonnard J, Deldime F, Heusterspreute M, et al. HLA class II genotyping: two assay systems compared. Clin Chem 1995;41(4):553-6.
23. Boichard D. PEDIG: A FORTRAN package for pedigree analysis studied for large populations. In. Montpellier, France: CD-ROM Communication no. 28-13; 2002:Proceeding of the 7th World Congr. Genet. Appl. Livest. Prod. 2002.
24. Lublin FD. The diagnosis of multiple sclerosis. Curr Opin Neurol 2002;15(3):253-6.
25. Tunbridge WM, Evered DC, Hall R, et al. The spectrum of thyroid disease in a community: the Whickham survey. Clin Endocrinol (Oxf) 1977;7(6):481-93.
26. Sadovnick AD, Yee IM, Ebers GC, Risch NJ. Effect of age at onset and parental disease status on sibling risks for MS. Neurology 1998;50(3):719-23.

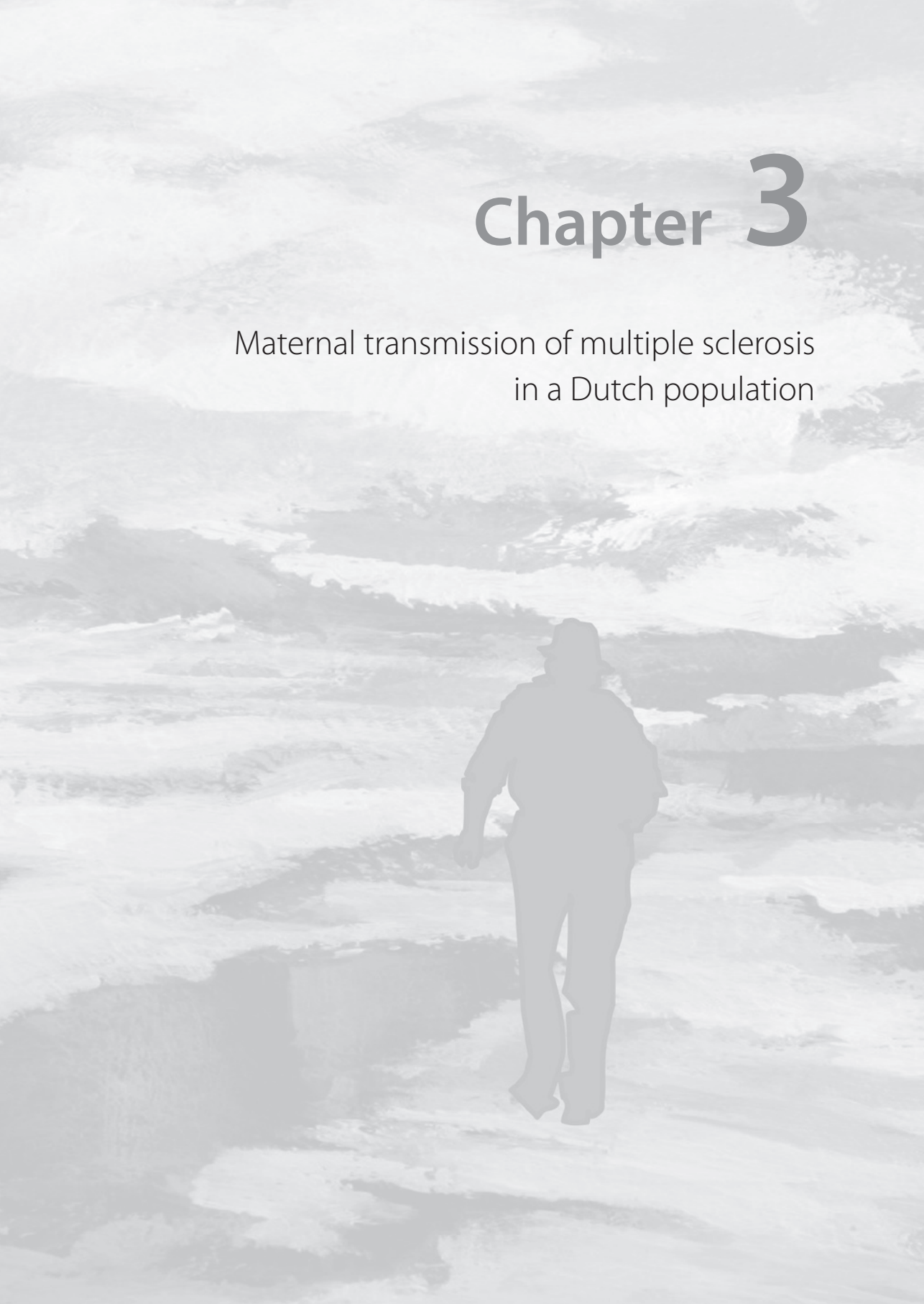
27. Marrosu MG, Lai M, Cocco E, et al. Genetic factors and the founder effect explain familial MS in Sardinia. Neurology 2002;58(2):283-8.

28. Vitale E, Cook S, Sun R, et al. Linkage analysis conditional on HLA status in a large North American pedigree supports the presence of a multiple sclerosis susceptibility locus on chromosome 12p12. Hum Mol Genet 2002;11(3): 295-300.

29. He B, Giedraitis V, Ligers A, et al. Sharing of a conserved haplotype suggests a susceptibility gene for multiple sclerosis at chromosome 17p11. Eur J Hum Genet 2002;10(4):271-5.

Chapter 3

Maternal transmission of multiple sclerosis in a Dutch population



Abstract

Objective: To investigate the parental relationship of patients with multiple sclerosis (MS) from an extended pedigree with extensive genealogical information up to the middle of the 18th century.

Design: Multiple sclerosis is a complex disease resulting from genetic and environmental factors. Parent-of-origin effect, a phenomenon when the same allele may express differently depending on the sex of the transmitting parent, may influence the risk for MS. We investigated parental relationships between patients with MS using extensive genealogical information available from the Genetic Research In Isolated Populations (GRIP) program. We compared the average kinship of the parents of MS patients. We further explored the distribution of shortest genealogical links between parents of MS patients.

Subjects: Twenty-four MS patients from the isolated population who could be linked within a large complex pedigree, including 2,471 people in total.

Results: The results consistently indicate a higher prevalence of maternal transmission of MS. The kinship between mothers of patients was 3.8 times higher than that between fathers (bootstrap $P=0.01$). Among the 814 shortest connections between parents, 333 were maternal (41% versus 25% expected), 98 were paternal (12% versus 25% expected), and 383 were maternal-paternal (47% versus 50% expected) ($P<0.001$).

Conclusions: Mothers of MS patients were more closely related than their fathers. This skewed relationship shows evidence for a maternal effect in MS. The most likely explanation is a gene-environment effect that takes place in utero.

Introduction

Multiple sclerosis (MS) is a complex disease resulting from the interplay between both genetic and environmental factors. Genetic factors are implicated to play a role in determining MS risk by adoption studies,[1] studies with half-siblings,[2] twin studies[3] and studies of conjugal MS.[4] Migration studies, have especially pointed to environmental factors as playing a role in MS.[5,6]

Other influences, as parent-of-origin effects, are also described in determining MS risk. Parent of origin is a phenomenon when the same allele is expressed differently depending on the sex of the transmitting parent. A study of half-siblings showed that maternal half-siblings have significantly higher risk of developing MS compared with paternal ones, suggesting a maternal parent-of-origin effect.[2] In contrast, observations in offspring of affected parents demonstrated excess of paternal vs maternal transmission.[7]

Both studies in which a parent-of-origin effect in MS was suggested were performed in differentially selected patient groups: with nonaffected parents in the study by Ebers et al[2] and affected parents in the study by Kantarci et al.[7] They had in common that only recent generations of patients with MS were studied.

We studied parent-of-origin effects of MS transmission in a genetically isolated population, with extensive genealogical information over centuries. Most cases were not closely related but were initially diagnosed as sporadic MS cases. The clinical MS phenotype was similar to MS in the general Dutch population.[8]

Methods

Study population and patient ascertainment

This study was performed within the framework of the previously described Genetic Research in Isolated Populations (GRIP) program.[9,10] The Medical Ethics Committee of the Erasmus MC approved GRIP protocols. The GRIP population is a genetically isolated community in the southwest of the Netherlands. Fewer than 400 individuals founded the population in the middle of the 18th century. Considerable population growth subsequently occurred. Until recently, there was minimal immigration. An estimated 20,000 descendants of this population are scattered over eight adjacent communities. The genealogical database contains information on more than 90,000 people spanning 23 generations. Residents in the GRIP area are generally related via multiple lines of descent.

Ascertainment and clinical characteristics of MS patients in the GRIP population have been described in detail previously.[8] In brief, 24 MS patients (5 men and 19 women) could be linked to the most recent common ancestor in 14 generations. These 24 MS patients had given written informed consent. Numerous connections exist between these patients via multiple common ancestors.

Parent-of-origin effect

We compared average kinship of parents of the 24 MS patients, testing whether patients were more often related through paternal or maternal lineage. Kinship coefficients were computed using a computer software package (PEDIG).[11] Under the null hypothesis of no parent-of origin effect, the average degree of relationship of fathers should not be different from the one of mothers, assuming random mating.

When there is deviation from random mating (in particular, preferential outbreeding for one of the sexes), the test based on comparison of average kinships between mothers and fathers is not correct. For example, if there is systematic outbreeding of males, one would expect to observe higher kinship between mothers of any randomly selected group of people from the population.

To assess the empirical null distribution of differences in parental relationship specific for our population, we performed analysis using 10,000 replicas. For each time, 24 individuals, adjusted for sex and age with the patients, were randomly sampled from the GRIP genealogy database. This procedure is also known as bootstrap. Average kinship was computed for fathers and mothers of these randomly sampled individuals, and the ratio between maternal and paternal kinships was computed. The proportion of realizations in which the ratio was the observed ratio or more gives the empirical *P* value.

We further explored genealogical links between MS patients via all of their common ancestors using a software package (FCN; available at http://mga.bionet.nsc.ru/soft/index.html). We next focused on the shortest connection between patients. Under the hypothesis that a genetic factor plays a role in MS, one assumes that patients are more closely related to each other than control subjects. In terms of genealogical links, this implies that, on average, links between MS patients are shorter than that those between controls. For the shortest genealogical link between two MS patients there are three possible patterns: two patients are related through their mothers, through the father of one patient and the mother of the other patient, or through their fathers. Under the null hypothesis of no parent-of-origin effect, and assumption of random mating, the expected distribution of these three patterns is 25% through their mothers, 25% through their fathers and 50% through a father and a mother. A χ^2 test with 2 *df* was used as the test statistic.

In all analyses, when siblings were present as patients, the parents were included in paternal/maternal samples only once to avoid possible bias.

We had the unique opportunity to test parent-of-origin effect in the same way for several other diseases in the GRIP area, including Parkinson disease, late-onset Alzheimer disease, and Type 1 diabetes.

Results

The general characteristics of the 24 MS patients are as follows. Their mean (SD) age at onset of symptoms was 29 (9) years; at diagnosis, 36 (11) years. There was no significant (P =0.57) difference in age of symptom onset between the 24 MS patients and 73 sporadic Dutch MS patients from outside the isolate.[8,12] There were 19 women (79%) in the sample. The distribution of relapse-onset MS (20 patients

[83%]) and primary progressive onset MS (4 patients [17%]) was not different from that in the general MS population. Of the 24 patients, 15 were initially diagnosed as isolated MS cases and nine were from families in which patients were themselves aware of having two or more MS cases in their families. Of these nine patients, four were affected sisters from one family, all participating in this study. The other five patients were from families in which two people were diagnosed as having MS: two participating patients were third-degree relatives, two participating patients were second-degree relatives, and one patient had a first-degree relative with MS who did not want to participate in this study. None of the parents or grandparents of the 24 patients was diagnosed as having MS.

The characteristics of the genealogy of the MS patients are as follows. The mean (SD) number of consanguineous loops per patient was 139.9 (278.9), with a range from 0 to 1,205; the mean (SD) number of meioses per loop was 12.4 (1.3), with a range from 0 to 29; and the mean (SD) inbreeding was 1.4×10^{-3} (1.9×10^{-3}), with a range from 0 to 7.7×10^{-3}. Patients with MS could be linked to multiple common ancestors in different generations. Each pair of patients shares, on average, 264.2 (SD, 449.4) common ancestors (range 2-2,936 common ancestors). The mean number of meioses separating a pair of patients was 22.2 (SD, 2.1), with a range from 2 to 36, which means that their common ancestry is, on average, 11 generations ago. The most recent common ancestor for all 24 patients could be identified 14 generations ago. Finally, the mean (SD) kinship for the patients was 6.7×10^{-3} (3.7×10^{-2}), with a range from 2×10^{-7} to 0.3.

To assess the parent-of-origin effect, we compared the mean kinship of mothers and the one of fathers. The mean kinship of mothers (0.0031) was 3.9 times higher than the one of fathers (0.0008). A 2-independent-samples *t* test comparing these kinships gives a significant *P* value (<0.001). The empirical null distribution of the ratio between maternal and paternal kinships using 10,000 samples of 24 age- and sex-matched controls randomly selected from our genealogical database was derived. The probability of observing the same or a more extreme kinship ratio in this population was $P = 0.01$, indicating that closer kinship between mothers of MS patients is not a statistical artifact or a consequence of violation from the random mating assumption.

We further explored genealogical links between MS patients. In total 36,466 pairwise connections between each possible pair of 24 MS patients were identified via 796 ancestors. Among these connections, we counted shortest connections between each pair of MS patients in respect to the three possible patterns described in the : "Methods" section. Among 814 shortest connections, 333 were between mothers (40.9%; 25.0% expected), 98 were between fathers (12.0%; 25.0% expected), and 383 were between a father and a mother (47.1%; 50.0% expected) ($P < 0.001$) (Figure 1).

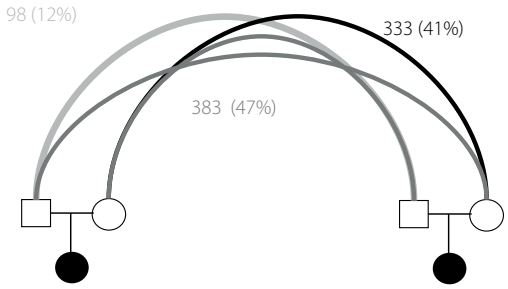


Figure 1 Observed distribution of connections between patients with multiple sclerosis (MS). Solid circles indicate female patients with MS; open circles, unaffected mothers of patients with MS; and open squares, unaffected fathers of patients with MS.

We also investigated parent-of-origin effect for several other diseases in the GRIP area, using the same database and the same method. For 67 patients with Parkinson disease, 103 with late-onset Alzheimer disease (P =0.052), and 39 with Type 1 diabetes (P =0.38) who could be linked to a common ancestor, no significant parent-of-origin effect could be found.

Discussion

A pedigree of 24 patients with common clinical phenotypes of MS was recently reconstructed.[8]

Herein, we show that the shortest connection to a common ancestor between two individuals with MS was significantly more often through their nonaffected mother than through their nonaffected father, suggesting a maternal parent-of-origin effect. No significant parent-of-origin effect was observed in several other diseases in the same area. Thus, the effect seems to be specific for MS.

Mothers of the 24 MS patients were also more closely related to each other than their fathers.

Multiple sclerosis has a female-male ratio of 3:1. Potentially, this carries the risk of bias when studying disease transmission over two generations when parents are affected and women are at increased risk of MS. Our approach had no bias into this skew because none of the parents and grandparents of the 24 MS patients were affected with MS.[8] A few other studies analyzed parent-of-origin effects, all using the information of two generations.[2,7,13] The significant difference with our study is that by using extensive genealogical information we were able to study the shortest links by a multigenerational approach.

Our findings are in line with results of the Canadian study in which the difference in MS risk was investigated for half-siblings of MS patients with a shared non-affected parent. A shared mother was associated with higher MS risk.[2]

In contrast, a recent study in the United States observed that fathers transmit the disease more frequently.[7] This was explained by the Carter effect. According to this theory, men are more resistant to MS because they require a higher genetic load and thus are more likely to transmit the genetic risk of the disease to their children. This phenomenon was not observed by others.[13] Our approach differed

from that of Kantarci et al[7] in that we did not study multiplex families and parents were not affected.

Taking together all available data on parental transmission of MS, one can conclude that the MS-affected status of the parent may influence transmission of disease.

Maternal transmission could be a result of several factors: genetic, environmental, or both. There are at least three single genetic explanations. None of them are supported heavily by current data. First, maternal effects could be exerted by direct transmission via mitochondrial genes or indirectly by autosomal genes involved in mitochondrial pathways, such as *UCP2* (uncoupling protein 2 gene) (GenBank MIM 601,693). This gene has a neuroprotective function and may contribute to MS susceptibility.[14,15] Thus far, attempts to link mitochondrial mutations with MS susceptibility have been disappointing.[16-18]

A second possible genetic explanation is genomic imprinting.[19] This has been insufficiently explored.

A third explanation could be the interaction between genes and female-specific environmental factors, such as hormonal, intrauterine, or perinatal factors.

Recent findings have demonstrated an increasing female to male sex ratio, strongly suggesting an environmental origin. It seems that the relative role of environmental factors in determining MS susceptibility has been changed during the past century, in possible interaction with genetic factors.[20]

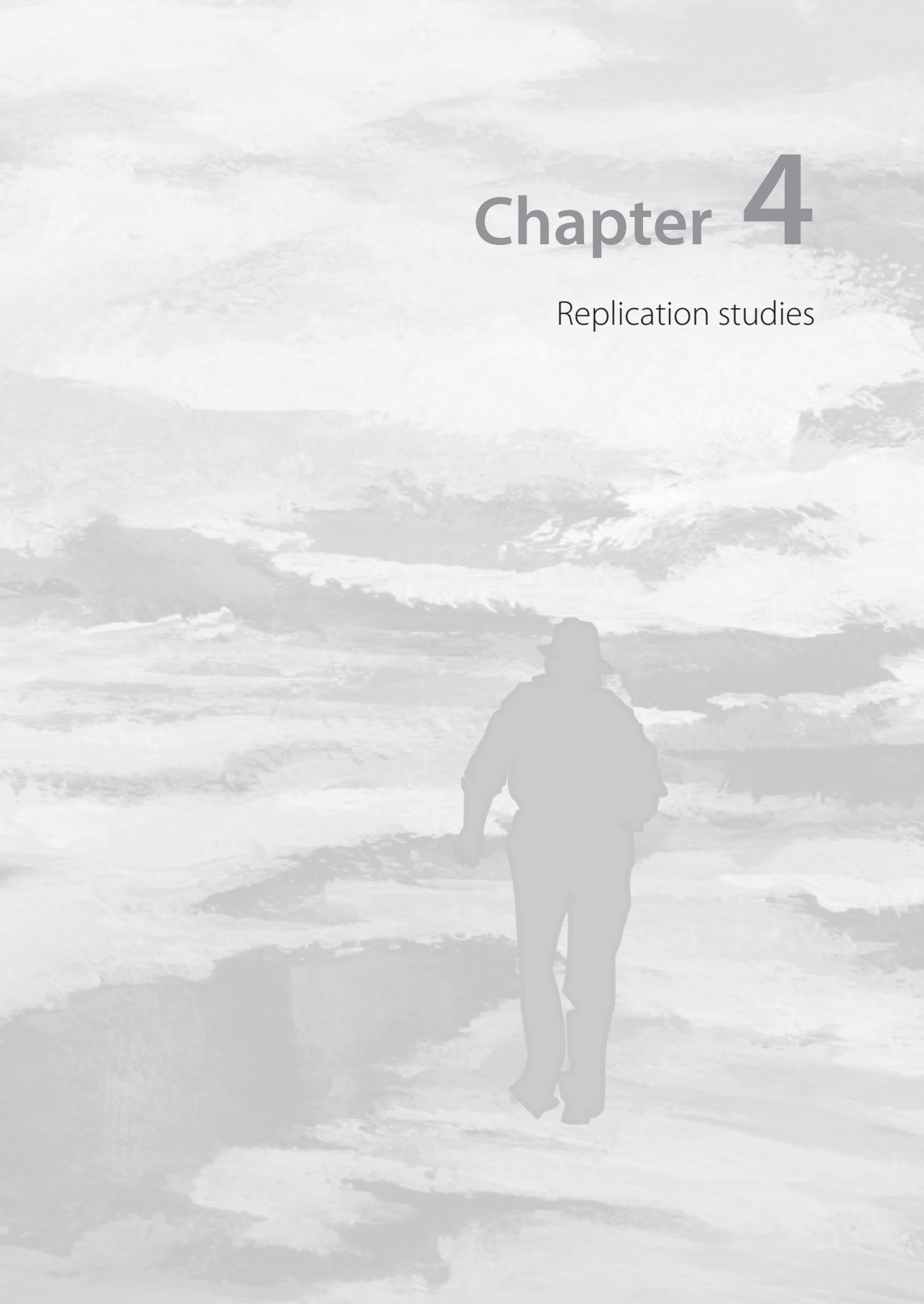
Our data show that MS patients are more often related through their mother than through their father, at least in the pedigree studied herein. Because MS in this pedigree has a common clinical phenotype and most cases were initially diagnosed as sporadic MS cases, our findings are likely to be representative for the total MS population. These findings support a maternal parent-of-origin effect in MS. There is reason to assume that this effect is caused by an interaction of genetic and environmental factors.[21] Dense genotyping in this pedigree can help to unravel the genetic contribution, thus aiding in resolving the nature-nurture dilemma in MS.

References

1. Ebers GC, Sadovnick AD, Risch NJ. A genetic basis for familial aggregation in multiple sclerosis. Canadian Collaborative Study Group. Nature 1995;377:150-1.
2. Ebers GC, Sadovnick AD, Dyment DA, Yee IM, Willer CJ, Risch N. Parent-of-origin effect in multiple sclerosis: observations in half-siblings. Lancet 2004;363:1773-4.
3. Willer CJ, Dyment DA, Risch NJ, Sadovnick AD, Ebers GC. Twin concordance and sibling recurrence rates in multiple sclerosis. Proc Natl Acad Sci U S A 2003;100:12877-82.
4. Robertson NP, O'Riordan JI, Chataway J, et al. Offspring recurrence rates and clinical characteristics of conjugal multiple sclerosis. Lancet 1997;349:1587-90.
5. Alter M, Kahana E, Loewenson R. Migration and risk of multiple sclerosis. Neurology 1978;28:1089-93.
6. Hammond SR, English DR, McLeod JG. The age-range of risk of developing multiple sclerosis: evidence from a migrant population in Australia. Brain 2000;123(Pt 5):968-74.
7. Kantarci OH, Barcellos LF, Atkinson EJ, et al. Men transmit MS more often to their children vs women: the Carter effect. Neurology 2006;67:305-10.
8. Hoppenbrouwers IA, Cortes LM, Aulchenko YS, et al. Familial clustering of multiple sclerosis in a Dutch genetic isolate. Mult Scler 2007;13:17-24.
9. Aulchenko YS, Heutink P, Mackay I, et al. Linkage in young genetically isolated Dutch population. Eur J Hum Genet 2004;12:527-34.
10. Pardo LM, MacKay I, Oostra B, van Duijn CM, Aulchenko YS. The effect of genetic drift in a young genetically isolated population. Ann Hum Genet 2005;69:288-95.
11. Boichard D. PEDIG: A FORTRAN package for pedigree analysis studied for large populations. In. Montpellier, France 2002:CD-ROM Communication.
12. Buljevac D, Flach HZ, Hop WC, et al. Prospective study on the relationship between infections and multiple sclerosis exacerbations. Brain 2002;125:952-60.
13. Herrera BM, Ramagopalan SV, Orton S, et al. Parental transmission of MS in a population-based Canadian cohort. Neurology 2007;69(12):1208-12.
14 Vogler S, Goedde R, Miterski B, et al. Association of a common polymorphism in the promoter of UCP2 with susceptibility to multiple sclerosis. J Mol Med 2005;83:806-11.
15 Vogler S, Pahnke J, Rousset S, et al. Uncoupling protein 2 has protective function during experimental autoimmune encephalomyelitis. Am J Pathol 2006;168:1570-5.
16. Chalmers RM, Robertson N, Compston DAS, Harding AE. Sequence of mitochondrial DNA in patients with multiple sclerosis. Ann Neurol 1996;40:239-43.
17. Kalman B, Alder H. Is the mitochondrial DNA involved in determining susceptibility to multiple sclerosis? Acta Neurol Scand 1998;98:232-7.
18. Harding AE, Sweeney MG, Miller DH, et al. Occurrence of a multiple sclerosis-like illness in women who have a Leber's hereditary optic neuropathy mitochondrial DNA mutation. Brain 1992;115(Pt 4):979-89.
19. Morison IM, Paton CJ, Cleverley SD. The imprinted gene and parent-of-origin effect database. Nucleic Acids Res 2001;29:275-77.
20. Orton SM, Herrera BM, Yee IM, et al. Sex ratio of multiple sclerosis in Canada: a longitudinal study. Lancet Neurol 2006;5:932-36.
21. Willer CJ, Dyment DA, Sadovnick AD, et al. Timing of birth and risk of multiple sclerosis: population based study. BMJ 2005;330:120-4.

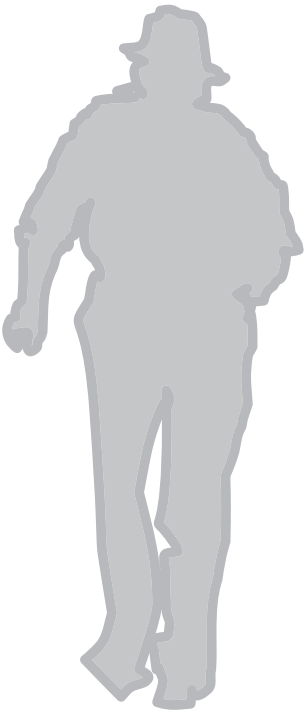
Chapter 4

Replication studies



Chapter 4.1

EVI5 is a risk gene for multiple sclerosis



Abstract

HLA-DRB1 is the major locus associated with risk for multiple sclerosis (MS). A recent genome-wide study showed three additional single-nucleotide polymorphisms (SNPs), within the *IL2RA* and *IL7RA* genes respectively, also to be associated with MS. Consistent association but lower significance was found for 13 other SNPs.

In this study we aimed to verify association of these SNPs with MS in 46 MS patients and 194 controls from a Dutch genetically isolated population. Apart form the human leukocyte antigen (HLA) locus, the *EVI5* gene on chromosome 1 was confirmed as a novel risk gene, with odds ratios (ORs) even higher than those from the MS Consortium (OR 2.01 and 1.9; P =0.01). The risk effect of *EVI5* was further validated for the general MS population in an independent set of 1,318 MS patients from the Canadian Collaborative Project on the Genetic Susceptibility to MS. On the basis of the transmission disequilibrium testing, a weak but significant risk effect was observed (OR 1.15; P =0.03 and OR 1.15; P =0.04). This study confirms *EVI5* as another risk locus for MS; however, much of the genetic basis of MS remains unidentified.

Introduction

Multiple sclerosis (MS) is a complex disease, resulting from genetic as well as environmental factors. For long, *HLA-DRB1* has been the only locus consistently involved with higher risk for MS. Next to the human leukocyte antigen (HLA) region, a recent genome-wide study also showed two single-nucleotide polymorphisms (SNPs) within the *IL2RA* gene and one SNP within the *IL7RA* gene to be strongly associated with MS susceptibility. Thirteen other SNPs, although less significant, also showed evidence for association with MS. In total, 17 SNPs were found to be associated with MS both in the screening phase and the replication phase of the study, of which the SNP in the *HLA* region again showed the strongest association.[1]

In this study, we assessed the risk contribution of these 17 SNPs in MS patients from a Dutch genetically isolated population. Apart form the *HLA* locus, a novel risk gene was confirmed. This finding is further validated for the general MS population in an independent large set of Canadian MS patients.

Patients and methods

Participants from the Dutch genetically isolated population

This study was performed within the framework of the previously described Genetic Research in Isolated Populations (GRIP) program.[2,3] The Medical Ethics Committee of the Erasmus MC approved GRIP protocols. The GRIP population is a genetically isolated community in the southwest of the Netherlands. The isolate was founded by less than 400 individuals around the middle of the 18th century. Minimal inward migration and considerable population growth subsequently occurred. An estimated 20,000 descendants of this population are now scattered over eight adjacent communities. The genealogical database currently contains information on more than 90,000 people spanning 23 generations. Residents in the GRIP area are generally related via multiple lines of descent.

The ascertainment and clinical characteristics of MS patients in the GRIP population have been described in detail previously.[4] Originally, 48 MS patients (13 male, 35 female) were included. Later on, another female MS patient was identified and included.

All 49 MS patients (13 male, 36 female) were diagnosed according to standard diagnostic criteria, with regular clinical phenotypes. Although most were originally diagnosed as sporadic MS patients, a total of 25 (51%) could be linked to an extended pedigree.[4]

Three of the 49 MS patients were excluded from analysis because they were sisters of one MS patient. Clinical characteristics from the 46 MS patients who were included for analysis are described in Table 1a.

In this population, there was a trend towards a higher prevalence of the *HLA-DRB1*15* allele in patients compared to controls, but no significant differences could be found (OR 1.79, P =0.12).[4]

For the control group, we included 194 healthy controls from the same area who were all distantly related (≥ 5 meioses).

All patients and controls gave written informed consent to participate in this study.

Participants from the Canadian Collaborative Project on the Genetic Susceptibility to MS

A total of 2,825 individuals from 756 families were typed as part of the Canadian Collaborative Project on the Genetic Susceptibility to MS (CCPGSMS) for which the methodology has been described.[5,6] This includes 1,318 individuals with definite MS and 1,507 of their unaffected first-degree relatives. Clinical characteristics from the 1,318 patients are described in Table 1b.

The Canadian families consisted of 456 multicase families (that is, parents with two or more affected offspring) and 300 parent-child trios.

Table 1a Description of the 46 Dutch MS patients

Variables	MS patients (n=46)
Age(years)	
Of onset symptoms	33 ± 11
Of diagnosis	39 ± 12
Sex	
Female (%)	72 (33)
Clinical phenotype at onset	
Relapse (%)	85 (39)
Primary progressive (%)	15 (7)

Abbreviation: MS, multiple sclerosis.

Table 1b Description of the 1,318 Canadian MS patients

Variables	MS patients (n=1,318)
Age	
Of onset symptoms	31 ± 9
Of diagnosis	34 ± 10
Sex	
Female (%)	60 (791)
Clinical phenotype at onset	
Relapse (%)	75 (989)
Primary progressive (%)	25 (329)

Abbreviation: MS, multiple sclerosis.
Dichotomous variables are expressed as percentage, absolute numbers in between brackets. Values of continuous variables are expressed as mean ± s.d.

Genotyping of individuals from the Dutch genetically isolated population

We genotyped all participants using the Affymetrix GeneChip Mapping 250K according to the protocol. This array set contains the Nsp array and includes approximately 262,000 SNPs.

Genotyping of participants from the Canadian Collaborative Project on the Genetic Susceptibility to MS

All genotypes were generated blind to pedigree structure and disease status of the individual.

Genotyping of SNPs was performed using the Sequenom MassEXTEND protocol (www.sequenom.com). Only conservative and moderate genotyping calls were accepted in this study. Samples having aggressive or low-probability quality genotypes were re-analysed. The concordance of genotyping between Affymetrix and Sequenom platforms has previously been validated.[1]

Statistical analyses of data from the Dutch genetically isolated population

For analysis, we used the R library GenABEL version 1.1-8.[7] We used the Armitage's test to estimate *P*-values. We used the genomic control method[8] to adjust for the relationship between GRIP participants.[9] The inflation factor was estimated to be 1.28. We imputed the SNPs that were not available from the Affymetrix GeneChipMapping 250K Array using HAPMAP CEU haplotypes as a reference. Mach software was used for imputations.[10,11]

Statistical analysis of data from the Canadian Collaborative Project on the Genetic Susceptibility to MS

Transmission disequilibrium test was performed using the PLINK analysis package.[12] PLINK implements sib-transmission disequilibrium test, which calculates empirical probabilities for χ^2 statistics, accurately reflecting association independent of linkage within families. This calculation is done by permuting parent alleles while fixing the identical by descent status of sibs within a family.

The transmission disequilibrium test counts the number of times an allele is transmitted to affected offspring from heterozygous parents. For transmission disequilibrium tests, the χ^2 distribution was used to assess significance. The OR was calculated as described in Kazeem and Farrall.[13]

Results

We tested the 17 MS SNPs that were reported to be associated in the collaborative genome-wide MS study (Table 2).[1] The *HLA-DRB1* surrogate SNP (rs3135388) was significantly associated with MS in this study (*P* =0.001, odds ratio (OR) 2.99, 95% confidence interval (CI) 1.56-5.74). Furthermore, two SNPs, both on chromosome 1, located in the *EVI5* (ecotropic viral integration site 5) gene gave significant *P*-values in our replication study: rs10735781 (*P* =0.01, OR =2.01, 95% CI 1.19-3.39) and rs6680578 (*P* =0.01, OR =1.9, 95% CI 1.16-3.11). These ORs are considerably higher than those reported by the International MS Consortium; the confidence intervals were not overlapping. The two SNPs in the *EVI5* gene were in nearly complete linkage disequilibrium (D' =0.99). The *IL7RA* rs689732 and for *IL2RA* rs12722489 and rs2104286, SNPs were not significantly associated with MS (Table 2).

The two SNPs located in the *EVI5* gene were subsequently tested in an independent set of 756 Canadian families containing 1,318 MS patients. Both SNPs had a weak but significant contribution in this population (rs10735781: *P* =0.03, OR =1.15, 95% CI 1.01-1.30; rs6680578: *P* =0.04, OR =1.15, 95% CI 1.01-1.30) (Table 3) and were in nearly complete linkage disequilibrium (D' =0.98).

Table 2 Results of replication analysis of 17 SNPs reported to be associated with MS in a Dutch genetically isolated population

Rs	Chr	Gene	Risk allele	NEJM, GWA				GRIP replication				
				RAF	OR	95% CI	*P*	Imputed	RAF	OR	95% CI	*P*
rs3135388	6	HLA-DRA	A	0.23	1.99	(1.84-2.15)	$8.94x10^{-81}$	Yes	0.14	2.99	(1.56-5.74)	0.001
rs12722489	10	IL2RA	C	0.85	1.19	(1.08-1.31)	$4.56x10^{-4}$	No	0.89	1.45	(0.70-3.03)	0.32
rs2104286	10	IL2RA	T	0.75	1.16	(1.08-1.25)	$1.49x10^{-4}$	Yes	0.79	1.23	(0.47-3.25)	0.67
rs6897932	5	IL7R	C	0.75	1.18	(1.09-1.27)	$2.75x10^{-5}$	Yes	0.7	1.34	(0.79-2.29)	0.28
rs6498169	16	KIAA0350	G	0.37	1.16	(1.09-1.24)	$1.89x10^{-5}$	No	0.31	1.11	(0.62-1.96)	0.73
rs6604026	1	RPL5	C	0.29	1.13	(1.05-1.22)	$9.58x10^{-4}$	No	0.28	0.87	(0.51-1.49)	0.61
rs10984447	9	DBC1	A	0.77	1.14	(1.05-1.24)	$1.27x10^{-3}$	No	0.78	0.83	(0.47-1.49)	0.54
rs12044852	1	CD58	C	0.92	1.2	(1.07-1.35)	$2.06x10^{-3}$	Yes	0.89	0.97		1
rs7577363	2	ALK	A	0.03	1.34	(1.11-1.62)	$3.15x10^{-3}$	Yes	0.05	0.67	(0.14-3.26)	0.62
rs7536563	1	FAM69A	A	0.38	1.08	(1.01-1.16)	$2.17x10^{-3}$	No	0.40	1.18	(0.72-1.93)	0.51
rs11164838	1	FAM69A	C	0.57	1.09	(1.02-1.16)	$1.30x10^{-2}$	No	0.52	1.38	(0.83-2.30)	0.22
rs10975200	9	ANKRD15	G	0.18	1.11	(1.02-1.21)	$2.12x10^{-2}$	No	0.16	1.77	(0.94-3.33)	0.08
rs10735781	1	EVI5	G	0.38	1.08	(1.01-1.16)	$2.01x10^{-2}$	No	0.33	2.01	(1.19-3.39)	0.01
rs6680578	1	EVI5	T	0.38	1.09	(1.01-1.16)	$1.86x10^{-2}$	No	0.31	1.9	(1.16-3.11)	0.01
rs4763655	12	KLRB1	A	0.38	1.09	(1.01-1.16)	$1.83x10^{-2}$	No	0.34	1.11	(0.62-1.99)	0.72
rs12487066	3	CBLB	T	0.73	1.08	(1.00-1.16)	$3.53x10^{-2}$	Yes	0.66	0.9		1
rs1321172	1	PDE4B	C	0.49	1.07	(1.01-1.14)	$3.95x10^{-2}$	Yes	0.51	0.58	(0.31-1.08)	0.08

Abbreviations: Chr, chromosome; CI, confidence interval; GRIP, Genetic Research in Isolated Populations; MS, multiple sclerosis; OR, odds ratio; SNP, single nucleotide polymorphism.

Table 3 Results of replication of the two SNPs located in gene *EVI5* in a Canadian cohort

SNP	Risk allele	TR:NT	Odds Ratio (95% CI)	χ^2	*P* asymptotic
rs10737581	G	510:443	1.151 (1.014-1.308)	4.71	0.02998
rs6680578	T	510:445	1.146 (1.009-1.301)	4.424	0.03544

Abbreviations: CI, confidence interval; NT, not transmitted; SNP, single nucleotide polymorphism; TR, transmitted.

Discussion

This study replicates in a Dutch genetically isolated population, the recent indication that *EVI5* is a risk gene for MS. Another interesting observation in the genetic isolate is that we found a very significant association of the *HLA-DRB1*15* tagging SNP rs3135388 with MS, whereas previous extensive *HLA-DRB1* typing in this same population showed no significant association with the known *HLA-DRB1*15* allele or other alleles at the two-digit level. A possible explanation for this may have been the lack of power in the initial study on this small population.[4]

The study on the genetic isolate had limited power to verify the low ORs of the recent genome-wide study.[1] The fact that we did not find statistical evidence for 14 of 17 SNPs that were earlier reported to be associated with MS does not exclude an effect of these SNPs/genes, as the probability of false-negative findings is high in this study. It is striking that despite this relatively low power, our study does show convincing evidence for *EVI5* at chromosome 1p22 as a risk allele. This underlines the strength of studying genetically isolated populations in complex diseases.[2,14] The ORs observed for *EVI5* are higher than the ORs reported by the International MS Consortium.[1]

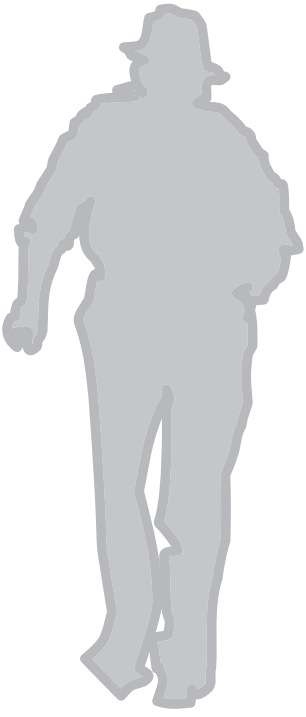
EVI5 was also found to be significantly associated with MS in a separate group ascertained as part of the CCPGSMS. This strengthens the evidence that *EVI5* is associated with MS risk. It is still not clear whether the causal allele acts through *EVI5* itself. *EVI5* is a common site of retroviral integration and has been linked to lymphomagenesis.[15] It remains to be seen whether and to what extent it could influence T-cell function and also if it could be related to retroviral elements associated with MS.[16] Identification of the exact causal allele will require the sequencing and genotyping of additional samples. In addition, functional immunological studies stratified according to *EVI5* genotype are planned.

References

1. The International Multiple Sclerosis Genetics Consortium. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 2007;357(9):851-62.
2. Aulchenko YS, Heutink P, Mackay I, et al. Linkage disequilibrium in young genetically isolated Dutch population. Eur J Hum Genet 2004;12(7):527-34.
3. Pardo LM, MacKay I, Oostra B, van Duijn CM, Aulchenko YS. The effect of genetic drift in a young genetically isolated population. Ann Hum Genet 2005;69:288-95.
4. Hoppenbrouwers IA, Cortes LM, Aulchenko YS, et al. Familial clustering of multiple sclerosis in a Dutch genetic isolate. Mult Scler 2007;13:17-24.
5. Sadovnick AD, Risch NJ, Ebers GC. Canadian collaborative project on genetic susceptibility to MS, phase 2: rationale and method. Canadian Collaborative Study Group. Can J Neurol Sci 1998;25:216-21.
6. Ramagopalan SV, Dyment DA, Valdar W, et al. Autoimmune disease in families with multiple sclerosis: a population-based study. Lancet Neurol 2007;6:604-10.
7. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics 2007;23:1294-6.
8. Devlin B and Roeder K. Genomic control for association studies. Biometrics 1999;55:997-1004.
9. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. Nat Genet 2007;39:770-5.
10. Mach: http://www.sph.umich.edu/csg/abecasis/MACH/index.html
11. Li Y and Abecasis GR. Mach 1.0: Rapid Haplotype Reconstruction and Missing Genotype Inference. Am J Hum Genet 2006;S79:2290.
12. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;1(3):559-75.
13. Kazeem GR, Farrall M. Integrating case-control and TDT studies. Ann Hum Genet 2005;69(Pt3):329-35.
14. Heutink P, Oostra BA. Gene finding in genetically isolated populations. Hum Mol Genet 2002;11(20):2507-15.
15. Liao X, Buchberg AM, Jenkins MA and Copeland NG. Evi-5, a common site of retroviral integration in AKXD T-cell lymphomas, maps near Gfi-1 on mouse chromosome 5. J Virol 1995;69(11):7132-7.
16. Perron H, Garson JA, Bedin F, et al. Molecular identification of a novel retrovirus repeatedly isolated from patients with multiple sclerosis. The Collaborative Research Group on Multiple Sclerosis. Proc Natl Acad Sci U S A 1997;94(14):7583-8.

Chapter 4.2

Replication of CD58 and CLEC16A as genome-wide significant risk genes for multiple sclerosis



Abstract

A recent genome-wide association study by the International Multiple Sclerosis Genetics Consortium (IMSGC) reported association of 17 single-nucleotide polymorphisms (SNPs) in 14 loci with multiple sclerosis (MS). Only two loci, *HLA-DRA* and *IL2RA*, reached genome-wide significance ($P < 5 \times 10^{-08}$). In our study, we determined whether we could replicate the results of the IMSGC and whether more SNPs are genome-wide significantly associated with MS. We assessed the association between the 17 IMSGC SNPs and MS in three cohorts (total number of subjects 3,981, among these 1,853 cases). We performed a meta-analysis of the results of our study, the original IMSGC results and the results of a recent replication study performed in the Australian population. Of the 17 IMSGC SNPs, five SNPs showed genome-wide significant association with MS: *HLA-DRA* ($P = 8 \times 10^{-124}$), *IL7*R ($P = 6 \times 10^{-09}$), *IL2RA* ($P = 1 \times 10^{-11}$), *CD58* ($P = 4 \times 10^{-09}$) and *CLEC16A* ($P = 3 \times 10^{-12}$). Therefore, genome-wide significance has now been demonstrated for SNPs in different non-HLA MS risk genes. Several of these risk genes, including *CD58* and *CLEC16A,* are shared by different autoimmune diseases. Fine mapping studies will be needed to determine the functional contributions to distinct autoimmune phenotypes.

Introduction

Multiple sclerosis (MS) is a complex disease resulting from genetic and environmental factors. The genetic influence on MS susceptibility is substantial, as evidenced by the 20-fold increase in risk for siblings of patients. Much of the high recurrence risk is explained by the MHC Class II region.[1] A recent genome-wide association study (GWAS)[2] conducted by the International Multiple Sclerosis Genetics Consortium (IMSGC) reported the association of MS with 17 single-nucleotide polymorphisms (SNPs) located in 14 regions. Only two regions, *HLA-DRA* and *IL2RA*, achieved genome-wide significance ($P < 5 \times 10^{-08}$). For a third gene, *IL7R*, convincing functional support was obtained[2-4] and genome-wide significance was established in a joint analysis of 11,019 cases and 13,616 controls.[5] In a recent Australian replication study of the 17 IMSGC risk SNPs, besides *IL2RA*, *CLEC16A*, *RPL5* and *CD58* were found to be associated with susceptibility for MS, although not genome-wide significant.[6]

We here assessed MS association of the 17 IMSGC reported SNPs in MS patients from a Dutch genetically isolated population (45 cases and 195 controls), in MS patients from the Dutch general population (490 MS cases and 426 controls), and in MS patients from the Canadian Collaborative Project on the Genetic Susceptibility to MS (CCPGSMS; 1,318 affected MS patients with their parents). In total, we studied 3,981 subjects, including 1,853 MS affected individuals. Results obtained in this study were also pooled with those obtained in the original IMSGC study[2] and the recent Australian replication study.[6]

Material and methods

Patients and genotyping

All patients fulfilled either Poser's criteria for definite MS or McDonald's criteria for MS.

The Dutch outbred cohort consisted of MS patients who are part of an ongoing nationwide study on genetic susceptibility in MS. A total of 490 MS patients were included, 370 sporadic MS patients and 120 cases from 120 multiplex MS families (that is, parents with two or more affected offspring). Overall, ten percent of the patients (n =51) had a clinically isolated syndrome at the time of enrollment. The 426 healthy controls consisted of 26 unrelated spouses, together with 400 healthy blood donors. Further, we have sampled 45 MS patients within the framework of Genetic Research in Isolated Populations (GRIP) program.[7] As controls, we included 195 healthy individuals from the same area who were all distantly related. Details on ascertainment are given elsewhere.[8] A total of 1,318 individuals with definite MS and 1,507 of their unaffected first-degree relatives were typed as part of the CCPGSMS.[9] The research protocol was approved by the respective Medical Ethics Committees and written informed consent had been obtained.

Genotyping was carried out using the MassARRAY system/Homogeneous MassExtend assay, following the protocol provided by Sequenom. PCR and extension primers were designed using the Assay Design 3.0 program (Sequenom, San Diego, CA, USA). Briefly, 20 ng genomic DNA is PCR amplified using Titanium Taq DNA Polymerase (Clontech, Mountain View, CA, USA). PCR primers were used at 100

nM final concentrations for a PCR volume of 10 μl. The PCR condition was 95°C for 15 min, followed by denaturing at 94°C for 20 sec, annealing at 56°C for 30 sec, extension at 72°C for 1 min for 45 cycles and finally incubation at 72°C for 3 min. PCR products were first treated with shrimp alkaline phosphatase (Sequenom) for 20 min at 37°C to remove excess dNTPs. ThermoSequenase (Sequenom) was used for the base extension reactions. Analysis and scoring were performed using the program Typer 3.3 (Sequenom).

Statistical analysis

All analyses were performed using R software (http://www.r-project.org/). Estimates of odds ratios (ORs) and significance were tested using logistic regression as implemented in "glm" function. Analysis was performed without including covariates, therefore, effectively the analysis is equivalent to the Armitage trend test for proportions in genotypic 2x3 table. Thus, our analysis was similar to and compatible with these performed in external cohorts included in this meta-analysis, in that no adjustment was done for covariates and allelic ORs were estimated.

In the genetically isolated population, over-dispersion of the standard errors was estimated and corrected using genomic control approach.[10] Genomic control lambda was estimated as 1.37.[11] Meta-analysis of log (OR) was performed using a fixed model approach with inverse of the square of the estimates of standard error used as weights. Test for the heterogeneity of effects between studies was performed using standard Cochrane's *Q*-test; random effect model was estimated using 'rmeta' library for R (by T Lumley: http://cran.r-project.org/web/packages/rmeta/rmeta.pdf). We used a *P*-value of 5 x 10^{-08} as the threshold for genome-wide significance.

The combined predictive value of the multiple genetic variants was investigated in a simulation study. The methods have been described elsewhere,[12] but briefly, the simulation strategy creates a dataset that includes genotypes and disease status for 100,000 individuals in such way that all genotype frequencies and odds ratios are the same as reported in this paper and the disease prevalence is one in 1000. Predicted risks for all individuals are obtained using Bayes' theorem in which the earlier risk of disease (1 in 1000) is multiplied by the likelihood ratios of all single variants under the assumption of independent genetic effects. For this reason, we included one polymorphism per gene, selecting the polymorphism with the strongest OR. Therefore, in total we included 14 of the 17 SNPs. We examined the discriminative accuracy, which is the extent to which test results can discriminate between individuals who will develop MS and those who will not, and is commonly assessed by the area under the receiver operating characteristic curve (AUC). The AUC is the probability that the test correctly identifies the diseased individual from a pair, of whom one is affected and one is unaffected, and ranges from 0.5 (total lack of discrimination) to 1.0 (perfect discrimination). To obtain more robust AUC estimates, the simulation study was repeated 100 times and 95% confidence intervals were calculated. The AUC was obtained as the c-statistic by the function somers,[2] which is available in the Hmisc library of R software (Harrell FE. Design and Hmisc R function library. Available at: http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/RS).

Results

Results of the meta-analysis are summarized in Table 1, which provides ORs and *P*-values for each study and for the combined analysis of all studies, The Supplementary Table 1 additionally provides allelic frequencies in cases and controls (where available) and random effect model meta-analysis.

The SNP rs3135388^{A} located in the *HLA-DRA* region was most strongly associated with MS in all cohorts, as expected (Table 1). Meta-analysis, including the original IMSGC study, resulted in an OR estimate of 2.1 and a pooled *P*-value was 8×10^{-124}.

The SNP located in *IL7R* (rs6897932^{C}) was consistently confirmed in the three populations and the meta-analysis reached genome-wide significance ($P = 6 \times 10^{-09}$). In addition, two SNPs within the *IL2RA* gene rs2104286^{T} and rs12722489^{C} reached genome-wide significance. Moreover, our analysis established two other SNPs with genome-wide significance. For rs12044852^{C} located in *CD58* the OR was 1.23 ($P = 4 \times 10^{-09}$) and for rs6498169^{G} located in the *CLEC16A* locus (also named *KIAA0350)*, the OR was 1.17 ($P = 3 \times 10^{-12}$). Forest plots for these SNPs are shown in Figure 1a and b. There is some support in our analysis that another locus, rs10735781^{G} in *EVI5* is associated with MS (OR=1.12, $P = 2 \times 10^{-06}$) as independent replication was observed (Figure 1c) and the *P*-value of the meta-analysis improved compared with that reported in the initial IMSGC report. However, the overall analysis did not reach genome-wide significance (Table 1). Meta-analysis of SNPs in two genes that are in close proximity to *EVI5*, *RPL5* and *FAM69A*, as well as SNPs in three other genes, *ALK*, *CBLB* and *KLRB1*, all did not change the *P*-values observed by the IMSGC substantially. For three loci (*PDE4B, DBC1* and *ANKRD15*), our results suggest that the previously reported association in the initial IMSGC screen was a false positive association.

Among genome-wide significantly implicated variants, two (rs6897932 at *IL7R* and rs3135388 at *HLA-DRA*) showed nominally significant heterogeneity (Table 1). The source of heterogeneity for rs6897932 may be the data of the Australian study, which shows risk allele OR which is opposite (though not significant) to other studies. For the rs3135388, the heterogeneity may be introduced by the original study,[2] which shows relatively low OR (1.99) and the Dutch isolated population, showing a relatively high OR (3.08).

We calculated the AUC of the 14 genes (HLA included), found to be associated with MS by the IMSGC, for MS with the ORs we obtained in our meta-analysis. The AUC was 0.68 (95% confidence interval 0.67-0.68). The AUC for HLA alone was 0.63 (95% confidence interval 0.63-0.64), the AUC of the other 13 genes was 0.60 (95 % confidence interval 0.59-0.61).

Table 1 Association of 17 SNPs with MS in five cohorts

					IMSGC		*DO*		*DI*		*Can*		*Aus*		*All*		
Chromosome	**SNP**	**Gene**	**RA**	**EA**	**OR**	***P***	**OR**	***P***	**OR**	***P***	**OR**	***P***	**OR**	***P***	**OR**	***P***	P_{het}
1p31	*rs1321172*	*PDE4B*	*G*	*C*	*1.08*	*0.006*	*0.9*	*0.262*	*0.61*	*0.039*	*1.01*	*0.89*	*0.97*	*0.59*	*1.04*	*0.110*	*0.03*
1p22	rs10735781	EVI5	C	G	1.11	3E-04	1.1	0.297	1.96	0.004	1.15	0.03	1.09	0.2	1.12	2E-06	0.19
1p22	rs6680578	EVI5	A	T	1.11	5E-04	1.08	0.423	1.78	0.013	1.15	0.04	1.05	0.41	1.11	2E-05	0.25
1p22	rs6604026	RPL5	T	C	1.15	8E-06	1.04	0.687	0.89	0.623	1.02	0.75	1.14	0.041	1.12	7E-06	0.34
1p22	rs7536563	FAM69A	G	A	1.12	9E-05	1.05	0.577	1.2	0.418	1.09	0.14	1.06	0.35	1.10	2E-05	0.88
1p22	rs11164838	FAM69A	T	C	1.11	2E-4	0.99	0.919	1.34	0.217	1.11	0.16	1.07	0.24	1.10	5E-05	0.70
1p13	**rs12044852**	**CD58**	A	C	1.24	2E-05	1.57	0.003	0.95	0.881	1.18	0.009	1.2	0.042	1.23	**4E-09**	0.45
2p23	rs7577363	ALK	G	A	1.37	7E-05	0.8	0.436	0.67	0.498	1.22	0.22	1.32	0.12	1.29	6E-05	0.32
3q13	rs12487066	CBLB	C	T	1.09	0.005	1.17	0.109	0.99	0.976	1.05	0.44	1.05	0.49	1.08	0.001	0.88
5p13	**rs6897932**	**IL7R**	T	C	1.18	3E-07	1.16	0.148	1.54	0.113	1.25	6E-04	0.96	0.58	1.16	**6E-09**	0.05
6p21	**rs3135388**	**HLA-DRA**	G	A	1.99	9E-81	2.32	7^{E}-12	3.08	2E-04	–	–	2.49	1E-33	2.10	**8E-124**	0.02
9p24	*rs10975200*	*ANKRD15*	*A*	*G*	*1.14*	*3E-04*	*1.08*	*0.526*	*1.53*	*0.133*	*1.02*	*0.75*	*1.001*	*0.95*	*1.02*	*0.095*	*0.01*
9q33	*rs10984447*	*DBC1*	*G*	*A*	*1.17*	*9E-06*	*0.92*	*0.46*	*0.83*	*0.511*	*1.05*	*0.4*	*0.98*	*0.79*	*1.10*	*5E-04*	*0.05*
10p15	**rs2104286**	**IL2RA**	C	T	1.19	2E-07	1.21	0.065	–	–	1.32	4E-04	1.16	0.033	1.20	**1E-11**	0.62
10p15	**rs12722489**	**IL2RA**	T	C	1.25	3E-08	1.19	0.182	0.68	0.26	1.21	0.001	1.13	0.156	1.22	**7E-11**	0.39
12p13	rs4763655	KLRB1	G	A	1.1	7E-04	1.11	0.262	1.15	0.568	1.07	0.29	1.07	0.27	1.09	1E-04	0.99
16p13	**rs6498169**	**CLEC16A**	A	G	1.14	4E-06	1.29	0.007	1.19	0.481	1.17	7E-04	1.23	0.001	1.17	**3E-12**	0.64

Abbreviations: All, meta-analysis of all studies; Aus, Australian study; Can, Canadian population; DI, Dutch genetically isolated population; DO, Dutch Outbred; EA, effective allele; IMSGC, International MS Genetics Consortium study; OR, odds ratio; P_{het}, *P*-value from the test for heterogeneity of effects; RA, reference allele; **Bold**: SNPs achieving genome-wide significance ($P < 5 \times 10^{-08}$) in meta-analysis *Italic*: SNPs for which contradicting results were obtained between the original report of IMSGC and our data, as indicated by at least x10 increase in *P*-value.

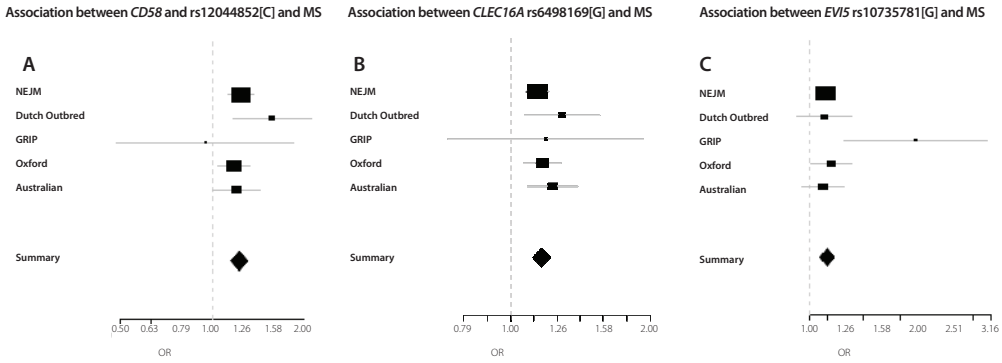


Figure 1 (a) Association between *CD58* rs12044852^{C} and MS. (b) Association between*CLEC16A* rs6498169^{G} and MS. (c) Association between *EVI5* rs10735781^{G} and MS. MS, multiple sclerosis. Effect estimates and 95% confidence intervals are shown. Study names: NEJM: IMSGC study; Dutch Outbred; GRIP: Dutch genetically isolated population; Oxford: Canadian population; Australian study.

Discussion

These data further establish the genome-wide significant association with MS of six out of the 17 SNPs that were found to be associated with MS in a previous GWAS.[2] The overall ORs of the non-*HLA* SNPs were modest, between 1.16 and 1.23. The predictive power of the 14 SNPs together was too low to have clinical significance, for example for diagnostic or prognostic purposes. The AUC is comparable to that found in recent studies on the prediction of Type 2 diabetes based on 18 susceptibility genes and on the genetic prediction of coronary heart disease.[13] The discriminative accuracy of *HLA* alone was better than the discriminative value of the other 13 genes together (0.63 vs 0.60)

Also in the separate populations, ORs were generally below 1.30, with an exception of an 1.58 OR for the *CD58* SNP in the Dutch outbred population and an OR of 1.96 for the replicated *EVI5* SNP in the very small population of the genetic isolate.[8] This study and others found that the risk effect from the *HLA* locus is independent from the risk signal coming from the other non-*HLA* risk loci.[2,6]

We found genome-wide significant association of rs12044852 in the *CD58* gene with MS (overall $P = 4 \times 10^{-09}$), which is in line with two very recent studies. The Australian and New Zealand Multiple Sclerosis Genetics Consortium showed genome-wide significance for the *CD58* SNP rs1335532.[14] This SNP is in strong linkage disequilibrium (LD) with rs12044852 (R^2=0,93), both being located in intron 10/11 of the *CD58* gene.[15] In addition, a study in MS patients from the United Kingdom and the United States demonstrated genome-wide significance, and further fine mapping indicated rs2300747 as the best susceptibility allele within the *CD58* locus.[16,17] Todd and colleagues recently screened the rs12044852 SNP also in Type 1 diabetes (T1D) patients in whom it does not seem to have a role.[18] *CD58* encodes a ligand for the T-cell specific CD2 membrane molecule, an adhesion molecule that transduces important signals for T-cell proliferation and differentiation. In addition, a role for the CD58 molecules has been

suggested in chronic inflammatory polyneuropathies.[19]

The *CLEC16A* rs6498169 SNP (intron 22) was also found to be genome-wide significantly associated to MS when combining the Australian study[6] and the IMSGC screen ($P = 3 \times 10^{-08}$).[2] Genome-wide significance was further only noted for another SNP in this gene, the T1D-associated SNP rs12708716 located in intron 19 ($P = 1.6 \times 10^{-16}$).[20] In the recently published meta-analysis and replication study in MS patients from the UK and the US, the SNP rs11865121 in the *CLEC16A* gene was found to be associated with susceptibility to MS ($P = 1.77 \times 10^{-07}$) in the joint analysis.[17] A study in Sardinia that explored the contribution of T1D genes to MS risk, found association, although not genome-wide significant, with another SNP in intron 19 rs725613 ($P = 4 \times 10^{-05}$)[21] which is in perfect LD with the genome-wide significantly associated SNP rs12708716 (R^2=1.0). The *CLEC16A* SNP rs6498169 that was identified here as genome-wide significant is in a different haplotype block (R^2=0.2), suggesting that in a single gene different SNPs are involved to different autoimmune disorders. A third autoimmune disorder that has been associated with *CLEC16A* is autoimmune Addison's disease.[22] It remains to be determined to what extend the different genetic variants within the *CLEC16A* area contribute to the susceptibility for certain autoimmune disorders.

Still little is known regarding the function of CLEC16A protein in humans. It is a member of the C-type lectin family, of which members have described to provide signals for a decision between tolerance and immunity. They can bind bacterial products as well as endogenous ligands, and their signal can counteract the signal of Toll-like receptors, therewith influencing T-helper cell function. We previously implied a role for C-type lectin receptors in MS pathogenesis, and discussed a link with infections.[23] However, further research has to be undertaken to understand the exact function of the *CLEC16A* gene and subsequently how it could influence the susceptibility to MS.

Not surprisingly, many of the risk genes identified thus far are directly linked to adaptive immune functions, further stressing the autoimmune pathogenesis of the disease. IL2R and IL7R are receptors for the regulation of lymphocyte expansion and differentiation. *CD58* and *CLEC16A* both share functional characteristics with the recently identified MS risk gene *CD226*[20,24,25] that reached genome-wide significance,[20] by their involvement in cell-cell interaction, adhesion and signaling.

So far *KIF1B* has been reported the only neuronally expressed gene with genome-wide significant association with MS.[11]

It is of note that many by now validated and strongly suggested MS susceptibility loci are also associated with other autoimmune diseases such as TID and Graves' disease. Although these genes may very well account for the clustering of MS and other autoimmune diseases in certain populations and within families, also here again the protective as well as risk alleles MHC class II area may exert the strongest effects.

Some immunodulatory treatments can trigger autoimmune diseases such as Graves' disease or idiopathic thrombocytopenic purpura (ITP) in MS patients.[26,27] Genotyping of the overlapping autoimmune risk alleles[24] may identify patients at risk for such autoimmune side effects.

In conclusion, in this study genome-wide significant association of non-HLA risk genes with

susceptibility to MS is confirmed for *CD58, CLEC16A, IL2R* and *IL7R*. Several of the by now validated and strongly suggested risk genes are shared by different autoimmune diseases, including the risk genes for which we found genome-wide significance. A question that remains to be answered is if the SNPs in these are causative variants. Fine mapping studies will be needed to determine the functional contributions to the distinct autoimmune phenotypes.

References

1. Lincoln MR, Montpetit A, Cader MZ, et al. A predominant role for the HLA class II region in the association of the MHC region with multiple sclerosis. Nat Genet 2005;37:1108-12.
2. Hafler DA, Compston A, Sawcer S, et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 2007;357:851-62.
3. Lundmark F, Duvefelt K, Iacobaeus E, et al. Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. Nat Genet 2007;39:1108-13.
4. 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;39:1083-91.
5. International Multiple Sclerosis Genetics Consortium (IMSGC). Refining genetic associations in multiple sclerosis. Lancet Neurol 2008;7:567-9.
6. Rubio JP, Stankovich J, Field J, et al. Replication of KIAA0350, IL2RA, RPL5 and CD58 as multiple sclerosis susceptibility genes in Australians. Genes Immun 2008;9:624-30.
7. Aulchenko YS, Heutink P, Mackay I, et al. Linkage disequilibrium in young genetically isolated Dutch population. Eur J Hum Genet 2004;12:527-34.
8. Hoppenbrouwers IA, Cortes LM, Aulchenko YS, et al. Familial clustering of multiple sclerosis in a Dutch genetic isolate. Mult Scler 2007;13:17-24.
9. Sadovnick AD, Risch NJ, Ebers GC. Canadian collaborative project on genetic susceptibility to MS, phase 2: rationale and method. Canadian Collaborative Study Group. Can J Neurol Sci 1998;25:216-21.
10. Bacanu SA, Devlin B, Roeder K. The power of genomic control. Am J Hum Genet 2000;66:1933-44.
11. Aulchenko YS, Hoppenbrouwers IA, Ramagopalan SV, et al. Genetic variation in the KIF1B locus influences susceptibility to multiple sclerosis. Nat Genet 2008;40:1402-1403.
12. Janssens AC, Aulchenko YS, Elefante S, Borsboom GJ, Steyerberg EW, van Duijn CM. Predictive testing for complex diseases using multiple genes: fact or fiction? Genet Med 2006;8:395-400.
13. Janssens AC, van Duijn CM. Genome-based prediction of common diseases: advances and prospects. Hum Mol Genet 2008;17(R2):R166-73.
14. Australia and New Zealand Multiple Sclerosis Genetics Consortium (ANZgene). Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosome 12 and 20. Nat Genet 2009;41(7):824-8.
15. Baranzini SE, Wang J, Gibson RA, et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. Hum Mol Genet 2009;18(4):767-78.
16. de Jager PL, Baecher-Allan C, Maier LM, et al. The role of the CD58 locus in multiple sclerosis. Proc Natl Acad Sci U S A 2009;106:5264-9.
17. de Jager PL, Xiaoming J, 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(7):776-82.
18. Fung E, Smyth DJ, Howson JM, et al. Analysis of 17 autoimmune disease-associated variants in type 1 diabetes identifies 6q23/TNFAIP3 as a susceptibility locus. Genes Immun 2009;10:188-91.
19. van Rhijn I, van den Berg LH, Bosboom WM, Otten HG, Logtenberg T. Expression of accessory molecules for T-cell activation in peripheral nerve of patients with CIDP and vasculitic neuropathy. Brain 2000;123:2020-9.
20. International Multiple Sclerosis Genetics Consortium (IMSGC). The expanding genetic overlap between multiple sclerosis and type I diabetes. Genes Immun 2009;10:11-4.
21. Zoledziewska M, Costa G, Pitzalis M, et al. Variation within the CLEC16A gene shows consistent disease association with both multiple sclerosis and type 1 diabetes in Sardinia. Genes Immun 2009;10:15-7.
22. Skinningsrud B, Husebye ES, Pearce SH, et al. Polymorphisms in CLEC16A and CIITA at 16p13 are associated with primary adrenal insufficiency. J Clin Endocrinol Metab 2008; 93:3310-7.
23. 't Hart BA, Laman JD, Bauer J, Blezer E, van Kooyk Y, Hintzen RQ. Modelling of multiple sclerosis: lessons learned in a non-human primate. Lancet Neurol 2004;3:588-97.
24. Hafler JP, Maier LM, Cooper JD, et al. CD226 Gly307Ser association with multiple autoimmune diseases. Genes Immun 2009;10:5-10.
25. Seldin MF, Amos CI. Shared susceptibility variations in autoimmune diseases: a brief perspective on common issues. Genes Immun 2009;10:1-4.
26. Coles AJ, Wing M, Smith S, et al. Pulsed monoclonal antibody treatment and autoimmune thyroid disease in multiple sclerosis. Lancet 1999;354:1691-5.
27. Samijn JP, te Boekhorst PA, Mondria T, et al. Intense T cell depletion followed by autologous bone marrow transplantation for severe multiple sclerosis. J Neurol Neurosurg Psychiatry 2006;77:46-50.

SUPPLEMENTARY TABLE

Supplementary Table 1 Extended results of association of 17 SNPs with MS in five cohorts

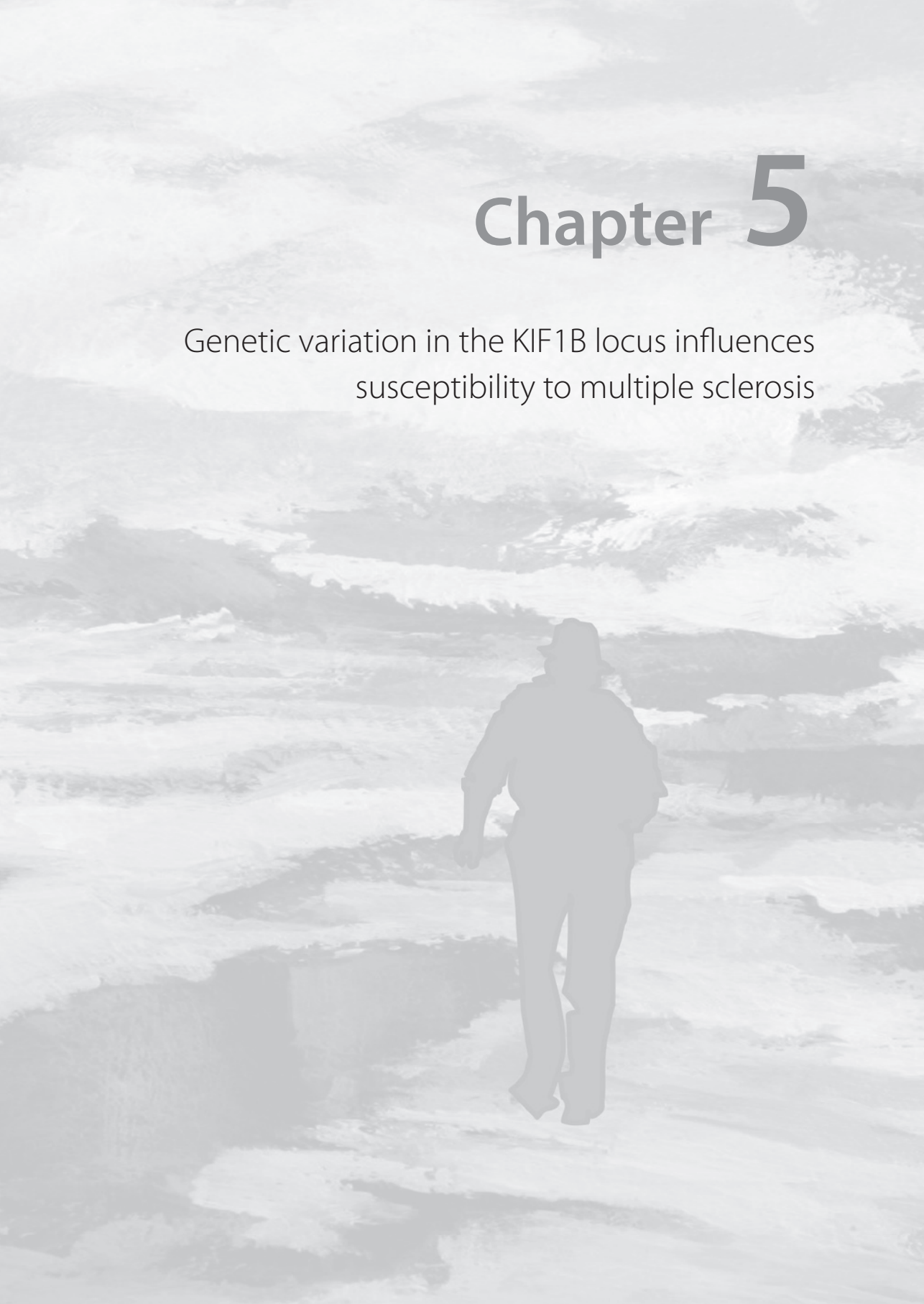
				NEJM		DO		DI						Can		Aus				J1			J1-random		J2		
SNP	Gene	RA	EA	OR	*P*	EAFa	EAFu	OR	*P*	RAFa	RAFu	OR	*P*	OR	*P*	RAFa	RAFu	OR	*P*	OR	*P*	P_{het}	OR	*P*	OR	*P*	P_{het}
rs1321172	PDE4B	G	C	1.08	0.006	0.47	0.5	0.9	0.262	0.41	0.53	0.61	0.039	1.01	0.89	0.519	0.526	0.97	0.59	1.04	0.110	0.03	0.98	0.74	0.96	0.26	0.21
rs10735781	EVI5	C	G	1.11	3E-04	0.37	0.35	1.1	0.297	0.48	0.31	1.96	0.004	1.15	0.03	0.388	0.37	1.09	0.2	1.12	2E-06	0.19	1.13	0.001	1.14	0.002	0.11
rs6680578	EVI5	A	T	1.11	5E-04	0.37	0.35	1.08	0.423	0.43	0.29	1.78	0.013	1.15	0.04	0.38	0.369	1.05	0.41	1.11	2E-05	0.25	1.11	0.002	1.11	0.012	0.15
rs6604026	RPL5	T	C	1.15	8E-06	0.27	0.27	1.04	0.687	0.26	0.28	0.89	0.623	1.02	0.75	0.3	0.273	1.14	0.041	1.12	7E-06	0.34	1.11	4E-4	1.06	0.12	0.52
rs7536563	FAM69A	G	A	1.12	9E-05	0.38	0.37	1.05	0.577	0.44	0.39	1.2	0.418	1.09	0.14	0.401	0.388	1.06	0.35	1.10	2E-05	0.88	1.10	2E-05	1.07	0.058	0.94
rs11164838	FAM69A	T	C	1.11	2E-4	0.57	0.57	0.99	0.919	0.58	0.51	1.34	0.217	1.11	0.16	0.572	0.555	1.07	0.24	1.10	5E-05	0.70	1.10	5E-05	1.07	0.077	0.63
rs12044852	CD58	A	C	1.24	2E-05	0.91	0.87	1.57	0.003	0.89	0.89	0.95	0.881	1.18	0.009	0.892	0.872	1.2	0.042	1.23	4E-09	0.45	1.23	5E-09	1.22	6E-05	0.31
rs7577363	ALK	G	A	1.37	7E-05	0.02	0.03	0.8	0.436	0.03	0.05	0.67	0.498	1.22	0.22	0.032	0.025	1.32	0.12	1.29	6E-05	0.32	1.26	0.004	1.16	0.18	0.37
rs12487066	CBLB	C	T	1.09	0.005	0.72	0.68	1.17	0.109	0.66	0.66	0.99	0.976	1.05	0.44	0.719	0.71	1.05	0.49	1.08	0.001	0.88	1.08	0.001	1.07	0.11	0.79
rs6897932	IL7R	T	C	1.18	3E-07	0.76	0.73	1.16	0.148	0.78	0.69	1.54	0.113	1.25	6E-04	0.751	0.758	0.96	0.58	1.16	6E-09	0.05	1.15	0.006	1.13	0.005	0.03
rs3135388	HLA-DRA	G	A	1.99	9E-81	0.28	0.15	2.32	7E-12	0.27	0.11	3.08	2E-04	–	–	0.312	0.166	2.49	1E-33	2.10	8E-124	0.02	2.27	5E-23	2.47	8E-47	0.67
rs10975200	ANKRD15	A	G	1.14	3E-04	0.17	0.16	1.08	0.526	0.22	0.15	1.53	0.133	1.02	0.75	0.161	0.161	1.001	0.95	1.02	0.095	0.01	1.06	0.16	1.00	0.76	0.44
rs10984447	DBC1	G	A	1.17	9E-06	0.76	0.77	0.92	0.46	0.76	0.79	0.83	0.511	1.05	0.4	0.758	0.761	0.98	0.79	1.10	5E-04	0.05	1.05	0.4	1.00	0.93	0.63
rs2104286	IL2RA	C	T	1.19	2E-07	0.76	0.72	1.21	0.065	–	–	–	–	1.32	4E-04	0.767	0.741	1.16	0.033	1.20	1E-11	0.62	1.20	1E-11	1.22	1E-05	0.46
rs12722489	IL2RA	T	C	1.25	3E-08	0.86	0.84	1.19	0.182	0.84	0.89	0.68	0.26	1.21	0.001	0.868	0.854	1.13	0.156	1.22	7E-11	0.39	1.21	4E-10	1.17	3E-04	0.39
rs4763655	KLRB1	G	A	1.1	7E-04	0.38	0.35	1.11	0.262	0.37	0.34	1.15	0.568	1.07	0.29	0.394	0.379	1.07	0.27	1.09	1E-04	0.99	1.09	1E-04	1.08	0.054	0.98
rs6498169	KIAA0350	A	G	1.14	4E-06	0.39	0.33	1.29	0.007	0.34	0.31	1.19	0.481	1.17	7E-04	0.376	0.331	1.23	0.001	1.17	3E-12	0.64	1.17	3E-12	1.20	7E-08	0.79

Abbreviations: EA, effective allele; EAFa, effective allele frequency in cases; EAFu, effective allele frequency in controls; OR, odds ratio; *P*, *P*-value; P_{het}, *P*-value from the test for heterogeneity of effects; RA, reference allele.

Study names: NEJM, original study of IMSGC; DO, Dutch Outbred; DI, Dutch genetically isolated; Aus, Australian study; J1, meta-analysis of all studies; J1-random, meta-analysis of all studies using random effects model; J2, meta-analysis excluding the NEJM study.

Chapter 5

Genetic variation in the KIF1B locus influences susceptibility to multiple sclerosis



Abstract

The few loci associated with multiple sclerosis (MS) are all related to immune function. We report a GWA study identifying a new locus replicated in 2,679 cases and 3,125 controls. An rs10492972[C] variant located in the *KIF1B* gene was associated with MS with an odds ratio of 1.35 ($P = 2.5 \times 10^{-10}$). *KIF1B* is a neuronally expressed gene plausibly implicated in the irreversible axonal loss characterizing MS in the long term.

MS is a complex disease resulting from genetic and environmental factors. The genetic influence on MS susceptibility is substantial, as evidenced by the 20-fold increase in risk for siblings of individuals with MS. Part of the high recurrence risk is explained by the MHC class II locus.[1] A recent GWA study[2] done by the International Multiple Sclerosis Genetics Consortium (IMSGC) reported association of MS with 17 SNPs in 14 loci. Although most loci were consistently associated, only variants in two genes, *HLA-DRA* and *IL2RA*, achieved genome-wide significance (P value $<5 \times 10^{-8}$). For a third gene, *IL7R*, convincing functional support was obtained. *IL2RA* and *IL7R* are both related to T-cell function and have also been linked to other autoimmune diseases.[2,3] However, these genes lack specificity for the central nervous system (CNS) pathology seen in individuals with MS.

We conducted a GWA study in 45 MS cases and 195 controls from a young genetically isolated Dutch population[4,5] using Affymetrix 250K Nsp array from GeneChip Human Mapping 500K Array Set (Supplementary Methods). We reasoned that the relative homogeneity of the sample would be advantageous. Association analysis did not reveal any SNP achieving genome-wide significance; the lowest P value observed was 10^{-6}. However, we did detect a signal in the *KIF1B* locus (chromosome 1p36.22) that was characterized by multiple associated SNPs (Figure 1). Although the significance for individual SNPs was modest for a GWA analysis (lowest P =0.0004 for rs10492972), the odds ratio (OR) for rs10492972[C] was 3.12 (95% CI =1.66–5.86) per allele, and 14 (52%) out of 27 SNPs in the 500-kb region surrounding rs10492972 showed association with P values < 0.05. There is extensive linkage disequilibrium (LD) in the *KIF1B* locus, and the boundaries of the major haplotype block, as estimated using HapMap CEU data (Figure 1), coincided well with the boundaries of association. When we included rs10492972 in a logistic regression analysis, no other SNPs remained significant (all $P > 0.05$), suggesting that association with the other 13 SNPs was most likely explained by LD with rs10492972. Strong LD is expected to occur around a causal variant, in particular when affected individuals are derived from a genetic isolate. As *KIF1B* is shown to encode a protein involved in axonal function, we followed up on rs10492972 in three replication sets of individuals with MS ascertained in outbred populations (Supplementary Methods): (i) a cohort from the Dutch general population (490 MS cases and 426 controls), (ii) a cohort from the Swedish general population (826 MS cases and 997 controls) and (iii) the Canadian Collaborative Project on the Genetic Susceptibility to MS (CCPGSMS; 1,318 individuals with MS and their parents). In total, we studied 2,634 cases and 2,930 controls in the replication phase.

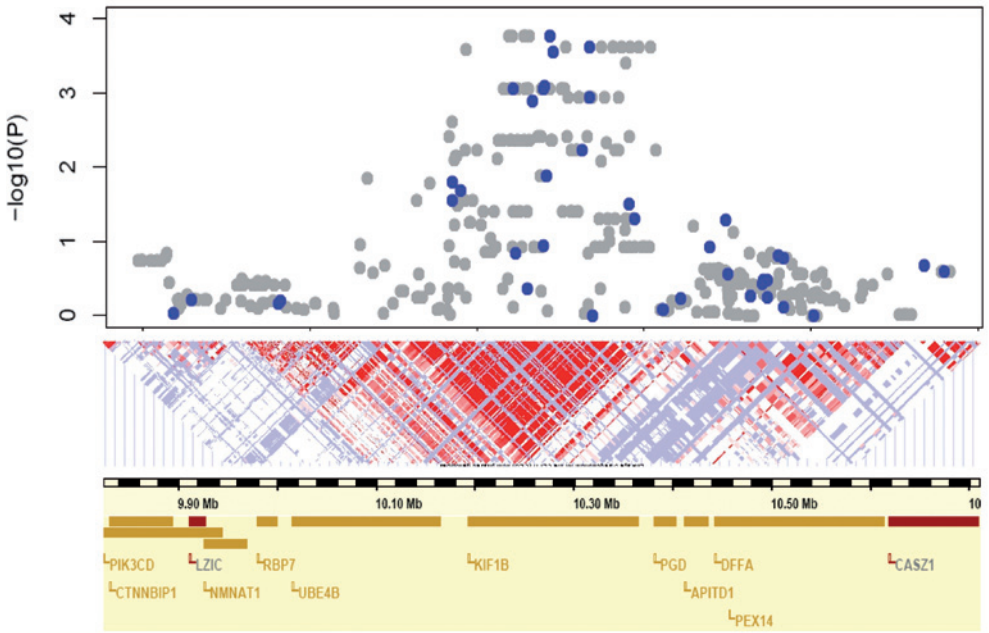


Figure 1 Association of MS with *KIF1B* region in a Dutch genetically isolated population. Blue, directly typed SNPs; gray, imputed SNPs. Linkage disequilibrium plot is constructed using the HapMap CEU data.

In each replication cohort, significant association of MS with the rs10492972[C] allele was demonstrated, with ORs consistently ranging from 1.3 (Swedish cohort) to 1.42 (Dutch outbred cohort). In each cohort, the frequency of the C allele was very similar and close to 0.34 in cases and to 0.27 in contrsols (Table 1). A meta-analysis of the data resulted in a joint OR of 1.35 (95% CI =1.23-1.48), with a significant *P* value of 2.5 x 10^{-10}. Analysis of genotypic model (Table 1) suggests that the effect of rs10492972[C] on the log(OR) scale is well approximated by the additive model.

Table 1 Results from the individual studies and meta-analysis of association between rs10492972[C] (*KIF1B* region) and MS

					Additive model		Genotypic model[a]			
Cohort	**No. cases**	**No. controls**	**RAF_a**	**RAF_u**	**OR_c**	***P***	**OR_{CT}**	**P_{CT}**	**OR_{CC}**	**P_{CC}**
Dutch isolated	45	195	0.43	0.21	3.12	0.0004[b]	2.73	0.007	11.23	0.0001
Dutch outbred	490	426	0.34	0.27	1.42	0.0009	1.50	0.004	1.85	0.010
Swedish	826	997	0.34	0.29	1.30	0.0003	1.25	0.025	1.75	0.001
Canada[c]	1,318	1,507	0.34	0.27	1.31	0.0012	1.21	0.002	1.45	0.004
Pooled	2,679	3,125	0.34	0.27	1.34	**2.5×10^{-10}**	1.27	1 x 10-6	1.67	4×10^{-8}

RAFa, risk allele frequency in affected individuals; RAFu, risk allele frequency in unaffected individuals.
[a]Ors and *P* values are given for the comparison with the TT reference group. [b]Corrected using genomic control. [c]Trio design, TdT analysis.

Rs10492972 is located in intron 5 of the *KIF1B* gene. Although analysis of imputed data (Figure 1) did not reveal any SNP more strongly associated with MS than the original rs10492972, we cannot claim that rs10492972 is the true causative variant. The question of whether rs10492972 itself, or an associated structural or regulatory polymorphism, is responsible for the association observed remains open. *KIF1B* has been linked to the inherited peripheral neuropathy Charcot-Marie-Tooth disease (CMT2A) in a single Japanese family.[6] However, later studies have shown that mutations in *MFN2* (mitofusin-2), which is located 2 Mb from *KIF1B*, are responsible for most cases of CMT2A linked to this locus.[7] In our study, SNPs in *MFN2* and its surrounding area (± 100 Kb) did not show evidence for association to MS (all $P > 0.01$).

Given the strong evidence for an autoimmune origin of MS, it was not unexpected that the first MS genes discovered by the IMSGC GWA[2] were related to T-cell function. There is, however, increasing evidence for neurodegenerative processes in MS pathology. Irreversible axonal loss is an important mechanism in the development of permanent neurological symptoms.[8] Although primary demyelination may underlie early axonal loss, further progression of neurodegeneration occurs when the compensatory capacity of the CNS is exceeded and the threshold of axonal loss is reached.[8] Mechanisms proposed for this loss of nerve fibers include mitochondrial dysfunction, reduced ATP production and altered axonal expression of sodium channels.[9-11] *KIF1B* encodes a kinesin superfamily member believed to be responsible for axonal transport of mitochondria and synaptic vesicle precursors.[12, 13] It has an ATPase binding domain and is enriched in motor neurons. Recently, dysregulation of ATPases and mitochondrial mislocalization have been shown to have a role in several neurodegenerative diseases.[12] *KIF1B* knockout mice clearly showed CNS abnormalities such as atrophy.[6]

The discovery of *KIF1B* is based on what may be the smallest number of cases studied by a GWA study. Although the a priori power was low, even in a genetically isolated population, the ability to locate this gene is most likely due to the demonstrably extended LD around rs10492972 in the isolate.[14] Genome-wide significant replication of the findings in the series of MS cases from the general Dutch, Canadian and Swedish populations confirmed the association of s10492972 to MS. We have previously confirmed the role of *EVI5*,[15] one of the putative 17 MS loci proposed by the IMSGC consortium.[2] Notably, both *EVI5* and the newly identified *KIF1B* showed higher ORs in our recently isolated population, as compared to the general population ($P = 0.03$ and 0.01, respectively). For *KIF1B*, this is difficult to interpret, as ORs are often inflated in the gene discovery sets. However, the finding for *EVI5*, which is based on a confirmation study, suggests that founding effects and genetic drift may have led to an enrichment of genetic variants and their effects in the genetic isolate.

Our study establishes a new genetic variant, rs10492972[C], that is associated with the risk of MS in the general population. This SNP is located in intron 5 of the *KIF1B* gene and may explain part of the progressive neurodegeneration seen in individuals with MS. Further functional studies are required to establish the exact role of the *KIF1B* gene region in the pathogenesis of MS.

References

1. Ebers GC. Environmental factors and multiple sclerosis. Lancet Neurol 2008;7(3):268-77.
2. Hafler DA, Compston A, Sawcer S, et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 2007;357(9):851-62.
3. 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(9):1108-13.
4. Aulchenko YS, Heutink P, Mackay I, et al. Linkage disequilibrium in young genetically isolated Dutch population. Eur J Hum Genet 2004;12(7):527-34.
5. Hoppenbrouwers IA, Cortes LM, Aulchenko YS, et al. Familial clustering of multiple sclerosis in a Dutch genetic isolate. Mult Scler 2007;13(1):17-24.
6. Zhao C, Takita J, Tanaka Y, et al. Charcot-Marie-Tooth disease type 2A caused by mutation in a microtubule motor KIF1Bbeta. Cell 2001;105(5):587-97.
7. Kijima K, Numakura C, Izumino H, et al. Mitochondrial GTPase mitofusin 2 mutation in Charcot-Marie-Tooth neuropathy type 2A. Hum Genet 2005;116(1-2):23-7.
8. Dutta R, Trapp BD. Pathogenesis of axonal and neuronal damage in multiple sclerosis. Neurology 2007;68(22 Suppl 3):S22-31; discussion S43-54.
9. Dutta R, McDonough J, Yin X, et al. Mitochondrial dysfunction as a cause of axonal degeneration in multiple sclerosis patients. Ann Neurol 2006;59(3):478-89.
10. Craner MJ, Newcombe J, Black JA, Hartle C, Cuzner ML, Waxman SG. Molecular changes in neurons in multiple sclerosis: altered axonal expression of Nav1.2 and Nav1.6 sodium channels and Na+/Ca2+ exchanger. Proc Natl Acad Sci U S A 2004;101(21):8168-73.
11. Waxman SG. Axonal dysfunction in chronic multiple sclerosis: meltdown in the membrane. Ann Neurol 2008;63(4):411-3.
12. Boldogh IR, Pon LA. Mitochondria on the move. Trends Cell Biol 2007;17(10):502-10.
13. Nangaku M, Sato-Yoshitake R, Okada Y, et al. KIF1B, a novel microtubule plus end-directed monomeric motor protein for transport of mitochondria. Cell 1994;79(7):1209-20.
14. Service S, DeYoung J, Karayiorgou M, et al. Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. Nat Genet 2006;38(5):556-60.
15. Hoppenbrouwers IA, Aulchenko YS, Ebers GC, et al. EVI5 is a risk gene for multiple sclerosis. Genes Immun 2008;9(4):334-7.

SUPPLEMENTARY METHODS AND RESULTS

Genetic variation in the *KIF1B* locus influences susceptibility to multiple sclerosis

Patients & genotyping

All patients fulfilled either Poser's criteria for definite multiple sclerosis or McDonald's criteria for multiple sclerosis. Each participating site has obtained ethical approval from the relevant institutional review board.

The Netherlands MS series. In the screening phase, we performed a GWA study on 45 MS patients from the Genetic Research in Isolated Populations (GRIP) program. As controls, we included 195 healthy individuals from the same area who were all distantly related. Details on ascertainment are given elsewhere.[1] The 250K Nsp array from the GeneChip Human Mapping 500K Array Set was applied to genotype with cohort genome-wide; BRLMM algorithm was used for genotype calling. Details are described elsewhere.[2]

For the replication cohort we recruited and ascertained MS patients as part of an ongoing nationwide study on genetic susceptibility in MS. At total of 494 MS patients were included, 374 sporadic MS patients and 120 cases from 120 multiplex MS families (i.e. families with two or more MS patients). Ten percent of the patients (n=51) had a clinically isolated syndrome at the time of enrollment. The 429 healthy controls consisted of 29 unrelated spouses, together with 400 healthy blood donors.

Genotyping in the replication sample was carried out using the MassARRAY system Homogeneous MassExtend assay, following the protocol provided by Sequenom. PCR and extension primers were designed using the Assay Design 3.0 program (Sequenom). Briefly, 20 ng genomic DNA is PCR amplified using Titanium Taq DNA Polymerase (Clontech). PCR primers were used at 200 nM final concentrations for a PCR volume of 10 l. The PCR condition was 95°C for 15 min, followed by denaturing at 94°C for 20 sec, annealing at 56°C for 30 sec, extension at 72°C for 1 min for 45 cycles, and finally incubation at 72°C for 3 min. PCR products were first treated with shrimp alkaline phosphatase (Sequenom) for 20 min at 37°C to remove excess dNTPs. ThermoSequenase (Sequenom) was used for the base extension reactions. Analysis and scoring were performed using the program Typer 3.3 (Sequenom).

Swedish case-control data set. In the Swedish study, 826 MS patients were collected at Karolinska University Hospital, Stockholm, Sweden. Further, 997 healthy blood donors from the Stockholm area served as controls. Genotypes for rs10492972 were assessed using a pre-designed TaqMan® SNP Genotyping Assay (Assay ID: C__30400488_20, Applied Biosystems). Primers were diluted four times compared to the manufacturer's manual. Allelic discrimination was performed using the Applied Biosystems 7900HT Fast Real-Time PCR equipment and the interpretation was done using the SDS 2.2.1 software (both from Applied Biosystems).

Canadian Collaborative Project on the Genetic Susceptibility to MS (CCPGSMS). A total of 2,825 individuals from 300 parent-child trios and 456 multi-case nuclear families comprising affected children and their parents were typed as part of the CCPGSMS for which the methodology has been described.[3,4] This includes 1,318 individuals with definite MS and 1,507 of their unaffected first-degree relatives. The Canadian families consisted of 456 multiplex MS families and 300 parent-child trios. All genotypes were generated blind to pedigree structure and disease status of the individual. Genotyping of SNPs was performed using the Sequenom MassEXTEND protocol (www.sequenom.com). Only conservative and moderate genotyping calls were accepted in this study. Samples having aggressive or low probability quality genotypes were reanalysed.

Statistical analysis

Genome-wide association analysis was performed using GenABEL package.[5] Only SNPs having minor allele frequency of at least 1% and demonstrating call rate of >93% were included in the study.

We have studied relations between people included into a GWA scan performed in the Dutch genetically isolated population using the genomic kinship (IBD) data estimated as described by Amin *et al*.[6] The 45 cases and 195 controls used in this study made a total of 28,680 pairs; of these 99.2% had genomic kinship of <0.0156 ($1/2^6$). Twenty-two pairs (0.08%) had genomic kinship between 0.13 and 0.0625, indicating possible second-degree relations.

Principal component analysis of the genomic kinship matrix[7] has not revealed genetic outliers and has also shown that cases and controls are well genetically matched. This is not surprising given the nature of population and careful selection of cases and controls on the region of origin and ethnicity.

It is known that even a small fraction of related people may cause inflation of the association test statistic when highly heritable traits are studied. We[6] and others have shown that genomic control (GC) can be effectively used to correct for such inflation. Therefore all results reported in the manuscript use genomic control correction, as estimated using qtscore function of GenABEL package.[5] GC λ was estimated as 1.37.

The Armitage trend test was used to access the significance of association between the disease and genotypes (coded as 0, 1, or 2; additive model). All results reported for the GRIP use genomic control[8] correction for possible inflation of the test statistics. Furthermore, EIGENSTRAT analysis[7] with up to ten principal components was performed to ensure robustness of results against false-positives.

In the analysis of GWA data, four SNPs reached genome-wide significant level with $P < 5 \times 10^{-8}$. These SNPs exhibited very low risk allele frequency in controls (MAF<1%) and high frequency in cases (MAF>12%). Since array genotyping is susceptible to bias,[9] we have re-genotyped these 4 SNPs together with an additional 18 "top" novel hits (score test corrected *P*-values ranging from 1.4×10^{-9} to 1.6×10^{-5}) using Sequenom MassARRAY system. We further searched for extended regions showing evidence for association to the disease, which characterizes regions including disease genes.[10] This method is particularly powerful in isolated populations.[11] For this reason, the *KIF1B* rs10492972 (rank = 89) was also included in the Sequenome retyping.

These 23 SNPs were also genotyped in the Dutch outbred cohort (494 cases and 429 controls). After re-genotyping the four genome-wide significant hits were found to be false positives generated by genotyping error. The lowest *P*-value in our GWAS was 10^{-6}. Despite that, the *KIF1B* rs10492972 showed a high consistency (99.1%) between Affymetrix and Sequenom genotypes and, additionally, showed evidence for association (*P* =0.0008) with MS in the Dutch outbred cohort. Therefore this SNP was followed up in two other outbred cohorts (Swedish and Canadian samples).

In the Dutch genetically isolated population, the corrected *P*-value for the SNP rs10492972 was estimated as 0.0002 when using the score test and 0.0004 when using Wald test based on the corrected estimates obtained in the logistic regression analysis. As an extra quality check, we performed EIGENSTRAT analysis[7] using up to 10 principal components. This analysis did not change our results concerning *KIF1B*; the *P*-value for rs10492972 ranged from 0.0001 (1 component) to 0.0003 (6 components). Imputations in the regions of interest were performed using MACH program[12] with HapMap release 21 CEU population as reference. Only imputed SNPs showing the ratio of observed to the expected variance (r2) over 0.8 were considered in analysis (Figure 1).

In the three outbred populations, no genomic correction was performed because of absence of genome-wide data. However, the fact that the ORs obtained in Dutch and Swedish outbred cohorts were very similar to those obtained from the Canadian sample transmission-disequilibrium test analysis, makes it unlikely that these two series of patients, both derived from very homogeneous populations, were biased by population stratification and leaves no opportunity that *KIF1B* finding is a false-positive due to population stratification.

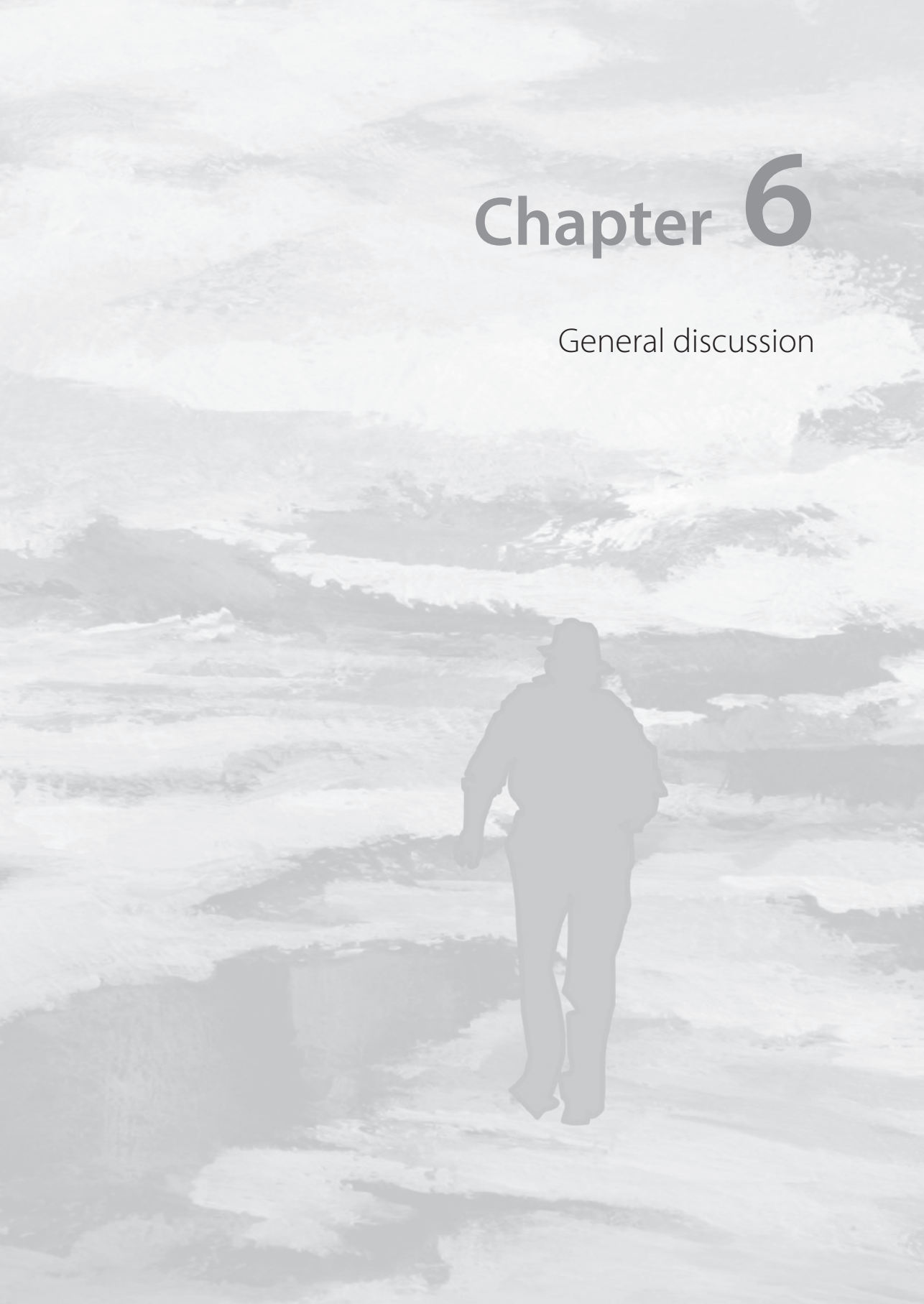
In the Dutch outbred population and Swedish cohorts, odds ratios and *P*-values were estimated using logistic regression as implemented in glm function of R v. 2.6.2. In the Canadian Collaborative Project on the Genetic Susceptibility to MS cohort, affected offspring analysis was performed with TDT as implemented in PLINK package.[13] The calculations are done by permuting parent alleles while fixing the IBD status of sibs within a family. The odds ratio was calculated as described in Kazeem and Farrall.[14] Meta-analysis was done using inverse variance pooling of log-odds ratios.

Supplementary references

1. Hoppenbrouwers IA, Cortes LM, Aulchenko YS, et al. Familial clustering of multiple sclerosis in a Dutch genetic isolate. Mult Scler 2007;13:17-24.
2. Kayser M, Liu F, Janssens AC, et al. Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. Am J Hum Genet 2008;82:411-23.
3. Sadovnick AD, Risch NJ, Ebers GC. Canadian collaborative project on genetic susceptibility to MS, phase 2: rationale and method. Canadian Collaborative Study Group. Can J Neurol Sci 1998;25:216-21.
4. Ramagopalan SV, Dyment, DA, Valdar W, et al. Autoimmune disease in families with multiple sclerosis: a population-based study. Lancet Neurol 2007;6:604-10.
5. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics 2007;23:1294-6.
6. Amin N, van Duijn CM, Aulchenko YS. A genomic background based method for association analysis in related individuals. PLoS ONE 2007;2:e1274.
7. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904-9.
8. Bacanu SA, Devlin B, Roeder K. The power of genomic control. Am J Hum Genet 2000;66:1933-44.
9. Clayton DG, Walker NM, Smyth DJ, et al. Population structure, differential bias and genomic control in a large-scale, case-control association study. Nat Genet 2005;37:1243-6.
10. Terwilliger JD. A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. Am J Hum Genet 1995;56:777-87.
11. Service S, DeYoung J, Karayiorgou M, et al. Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. Nat Genet 2006;38:556-60.
12. Li Y, Abecasis GR. Rapid Haplotype Reconstruction and Missing Genotype Inference. Am J Hum Genet 2006;2290.
13. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and populationbased linkage analyses. Am J Hum Genet 2007;81:559-75.
14. Kazeem GR, Farrall M. Integrating case-control and TDT studies. Ann Hum Genet 2005;69:329-35.

Chapter 6

General discussion



Multiple sclerosis (MS) is a complex disease and is thought to be the result of many environmental and genetic risk factors as well as a complex network of gene-gene and gene-environment interactions. For over 30 years the major histocompatibility complex (MHC) has been the only genetic association with susceptibility to MS.[1] The attributable risk of other risk alleles is probably very small. Detecting these is very difficult, requiring large populations. However, pooling studies might result in more heterogeneous populations, making it more difficult to detect risk genes.

In this chapter I will discuss our main findings and how our findings fit into current knowledge of MS genetics and pathogenesis. Finally, I will discuss the future prospectives in MS genetics and the consequences for the understanding of the pathogenesis of MS and the development of new therapies for MS.

Gene discovery

General principles

Family based versus population based genetic studies

There are two principal approaches to identify disease genes: linkage and association studies. In linkage studies families with a number of affected persons are required to investigate whether discrete chromosomal segments deviate from independent segregation and co-segregate (i.e. link) with the disease. Linkage studies will identify genes that are very rare but do have major effects.

In association studies the frequency of alleles at a polymorphic site in well-matched groups of affected and unaffected individuals are compared. Association studies have greater statistical power than linkage studies to detect common genetic variants that confer a modest risk of disease,[2] but are potentially important for defining new insights in disease pathogenesis. A disadvantage is that rare variants are generally not identified with association studies, whereas complex diseases are most likely to be caused by a complex interaction of both common variants with low penetrance and rare variants with high penetrance.[3]

Initially, most association studies have focussed on candidate genes. The main disadvantage of this approach is that it does require a hypothesis, based on the known function of the gene. Genome-wide association studies (GWAS), in which a sufficiently large number of well-spaced single nucleotide polymorphisms (SNPs) provide almost complete genomic coverage, do not require a previous hypothesis. GWAS are based on Chip technology, which allows the efficient genotyping of several hundred thousand SNPs throughout the genome simultaneously. However, given the many tests performed simultaneously, one has to make very stringent corrections for multiple testing. The level for genome-wide significance has been set at $P < 5 \times 10^{-8}$, representing Bonferroni correction for the number of blocks present in the human genome.[4]

Genetic association for complex diseases may be probed either with case-control studies of unrelated people or with family-based design. Both designs have advantages and disadvantages. The advantage of studies with unrelated cases and controls is that sufficiently large study populations

can readily be assembled without the need to enrol also family members of the recruited participants. However, a disadvantage of this approach is that confounding due to unaccounted population admixture remains a possible threat to the validity of obtained results.[5-7] On the other hand, family based study designs have the advantage that there is a common genetic background among the family members. Thus, the problem of population stratification is bypassed. Moreover, families tend to be more homogeneous regarding exposure to environmental factors possibly associated to the disease etiology.

Genetic studies in isolated populations

The selection of a population is crucial for genetic research. In order to reduce the problem of genetic heterogeneity, population isolates can be used in attempts to map genes underlying complex diseases.[8-12] Individuals from genetically isolated populations show significantly less genetic diversity compared with individuals from outbred populations.[13,14] The reduced genetic diversity is the result of a small number of founders, the occurrence of population bottlenecks and genetic drift. Population bottlenecks are the result of a sudden decrease in size, which is often seen in isolated populations in the form of famine, infectious disease epidemics and war. These population bottlenecks are followed by the survival and expansion of a small random sample of the original population. Predisposing genes for a disease are thought to be the same in patients from an isolate due to a restricted number of ancestor mutations.

In addition to the increased genetic homogeneity, isolates create a powerful setting to study the genetics of complex diseases through the usually available accurate genealogical records, which allows for the analysis of extended pedigrees.

In population isolates, particularly in those founded recently, fewer meiotic events will have occurred since a founder introduced a mutation in the gene pool. As a consequence it will be less likely that an ancestral haplotype has been disrupted by a recombination event, increasing the extent of linkage disequilibrium (LD) and thus the size of the shared haplotype. Hence, in comparison with outbred populations, such isolates may require fewer markers for GWAS or may achieve better genome-wide coverage with equivalent numbers of markers.[9,15] The disadvantage of GWAS carried out in genetic isolates is that the strong LD that initially helped to identify the disease locus may in the end hamper efforts to distinguish the biologically relevant variants from insignificant polymorphisms in complete LD with them.

Gene identification in multiple sclerosis

Although familial aggregation of MS has been recognized for long,[16,17] so far the mode of inheritance remains unclear.

Some studies suggested that a single genetic abnormality might be sufficiently severe to result in a form of inheritance following a Mendelian pattern. An autosomal dominant mode of inheritance of MS with reduced penetrance has been proposed for one family.[18] Autosomal recessive inheritance has been suggested in other studies[19,20] but either mode of inheritance remained unproven. In the majority

of patients, MS is probably the outcome of the additive effects of multiple genes with complex gene-gene and gene-environment interactions.

Family-based studies

Linkage screens with different levels of resolution and genome coverage have been completed in more than 30 data sets of familial MS cases.[21] Multiple chromosomal regions with potential involvement in MS susceptibility were suggested in each of these studies, consistent with the long-held view that MS is the result of many different genetic risk variants. However, no other region than the *HLA-class II* locus within the MHC region on chromosome 6p21.3 has exceeded the threshold for formal statistical significance. To assess the full potential of the linkage approach in MS, the International Multiple Sclerosis Genetics Consortium (IMSGC) reported in 2005 the results of a linkage screen in 730 multi-case MS families, using more than 4,500 SNPs. The higher peak LOD score of 11.7 in the *HLA* locus illustrates the substantial greater power of this high density screen compared with earlier efforts. Strikingly no other locus reached genome-wide significance in this screen.[22] However, there were non-MHC peaks of linkage observed, which failed to reach genome-wide statistical significance. These peaks do provide a useful guidance for the size of effects likely to be attributable to non-MHC susceptibility loci. It was calculated that common non-MHC risk alleles are highly unlikely to increase MS risk by more than a factor of 2.0.[2,22-24]

Genome-wide linkage screens were also done in a cohort of forty families in which four or more individuals were affected with MS[25] and in a single pedigree with a high prevalence of MS.[27] These screens also showed no evidence of linkage outside the MHC region. However, the size effect of the MHC appeared to be greater in families with many affected individuals.[26, 27] It was speculated that non-MHC susceptibility loci would have larger effects in multiplex families. Therefore 13 candidate non-MHC gene loci, identified as susceptibility loci for MS in an association screen performed by the IMSGC (described later in this discussion),[28] were tested for association in the multiplex families. The aggregate results demonstrate that the MS families have risk genes similar to that of the general MS population. However, the multiplex families showed higher risk allele frequencies than in sporadic cases, and they also exhibited a significant aggregation of susceptibility loci. Both factors seem to result in the increased susceptibility of MS within these families.[29]

Studies in isolated populations

Familial clustering of MS has been found in a northern Swedish rural district,[30] the county of Värmland in mid Sweden,[31] in an isolated population of Sardinia,[32] Southern Ostrobothnia in Finland [33] and in the Dutch genetic isolate we studied.[34]

As described in **chapter 2**, we identified MS patients in a genetically isolated population in the southwest of the Netherlands. We studied this population as part of the Genetic Research in Isolated Populations (GRIP) program. Extensive genealogic data were available for this population. We were able to construct a pedigree in which 24 of the 48 MS patients from the genetically isolated population could be linked to a common ancestor in 14 generations. Using the genealogical data we also found that the

MS patients were significantly more often related to each other and significantly more often inbred (higher probability that two alleles in one patient are identical by descent) than a non-MS control group derived from the same isolate.[34]

Studies in Southern Ostrobothnia in Finland demonstrate the power of population isolates in the identification of rare disease alleles. Southern Ostrobothnia is a genetic isolate in Finland with an exceptionally high prevalence and familial occurrence of MS.[33] In Finnish MS families enriched with cases from this isolate, linkage of MS to chromosome 5p had been detected.[8,35] The *IL7R* gene is located in the chromosome 5p region and this gene has been linked to MS in several population studies, as I will discuss further on. The 5p locus was studied further in the Finnish isolate and monitored for haplotype sharing. This analysis revealed only modest association at *IL7R*, whereas most significant association was found with one haplotype covering the *C7-FLJ40243* locus, 5.1 Mb centromeric from *IL7R*. The identified risk haplotype (~4% in the general European population) is relatively rare, but it has obviously become enriched in the southern Ostrobothnian high-risk MS region (12% of cases).[12] It contains complement component 7 (C7), an important factor for the innate immunity.

Interestingly, the risk genes for MS we confirmed (*EVI5*, **chapter 4**) and identified (*KIF1B*, **chapter 5**) in the GRIP population demonstrated higher odds ratios (OR) in our isolate compared to the general population, which is most probably due to the enrichment of genetic variants in the isolate as a consequence of founder effect and genetic drift.

Population studies

Candidate gene studies

Population-based association studies of candidate polymorphic genes, in which the frequencies of marker alleles in groups of patients and healthy controls are compared, have only been modestly successful in identifying disease-causing genes in MS, with the notable exception of the MHC locus. This is in part because of the difficulty in selecting from the many candidate genes without an unifying model of disease pathogenesis.

Genome-wide association and replication studies

Recently, the IMSGC performed a GWAS, using DNA microarray technology. In the screening phase >300,000 SNPs were tested in 931 trio families (half from the US and half from the UK). As would be predicted, the limited power provided by 931 trio families meant that no unequivocal associations were identified in the screening phase, outside the expected signals from the MHC region. For the replication phase 110 SNPs were selected on the basis of statistical significance and proximity to previously identified loci associated with susceptibility to autoimmune diseases. Finally, a joint analysis of data from 12,360 subjects was performed. The results revealed the existence of 16 non-MHC susceptibility SNPs in 13 gene loci of modest effect. However, only the non-MHC SNP in the *IL2RA* gene achieved genome-wide significant association with susceptibility to MS ($P = 3 \times 10^{-8}$).[28]

Other powerful GWAS and studies aiming to replicate genes identified in GWAS, including

ours as described in **chapter 4 and 5**, showed genome-wide significant association with susceptibility for MS for several variants: *HLA, IL7R,*[38] *IL2RA, KIF1B, CLEC16A* (*KIAA0350*),[36-38] *CD226* (DNAX accessory molecule 1, [*DNAM-1*]),[36] *CD58,*[38] *CD6,* [39] *IRF8* (*ICSBP1*),[39] *TNFRSF1A,*[39] a locus on chromosome 12 most probably relating to the gene *CYP27B,*[40] *TYK2,*[41-43] *STAT3,*[44] *KIF21B*[45,46] and *TMEM39A.*[45] These associated variants, except *TYK2* (mean allele frequency (MAF) <0,05) are common, have small odds ratios and explain only a fraction of the genetic risk.

In **chapter 4** we describe results of our studies, aiming to analyse the by the IMSGC identified MS risk genes.[28] In **chapter 4.1** we confirmed *EVI5* as a risk gene for MS in the Dutch genetically isolated population, despite the low numbers to verify the ORs of the risk genes identified by the IMSGC. Successively, we also confirmed *EVI5* as a risk gene for MS in an independent Canadian population, which strengthens the evidence that *EVI5* is associated with MS risk. In both populations the association was marginally significant (P =0.01 and 0.04 respectively). For the other 14 non-MHC SNPs described to be associated with MS risk by the IMSGC, we found no statistical evidence for an association. This does not exclude an effect of these SNPs as the probability of finding false-negative findings is high in our study, because of the low numbers. EVI5 is a common site of retroviral integration and has been linked to lymphomagenesis. Further studies will have to elucidate to what extend it could influence T-cell function and whether it could be related to retroviral elements associated with MS.[47]

The results of our study as described in **chapter 4.2** and our GWAS as described in **chapter 5** will be discussed later.

Do identified risk variants result in a better understanding of MS pathogenesis?

It remains important to recognize that identified disease associated genetic variants not necessarily cause the disease, but can also be markers of it. Most of the non-MHC loci are thought to confer disease susceptibility because they are the dominant candidate genes in a region of LD, in which a SNP or set of SNPs have been associated with disease. Many genes and associated variants are inherited together owing to LD in the genome. The current tools of human genetics cannot easily distinguish between a marker variant and a causal variant if the LD between them is either strong or absolute. The identification of the disease-causing variant will ultimately depend on functional studies.

Function of identified risk variants

The function of not all the currently identified risk variants is exactly known. Interestingly, most of the risk variants with known function have primarily immunologic functions.

I will describe what is known about the function of the different identified risk variants for which a genome-wide significant association with MS has been shown. Furthermore, I will give an overview of the level of evidence for association of each risk variant with MS (which can also be only a marker for the real disease causing genes as described above), dividing the risk variants in three groups: those with very high evidence, those with high evidence and those with modest evidence for an association with MS.

Very high evidence for an association

MHC class II

MHC class II molecules are highly polymorphic cell-surface glycoproteins that present antigen to CD4+ T helper cells and are integral to successful maintenance of self-tolerance by the immune system and the adaptive immune response to invading pathogens.[48] Each *HLA-DRB1* allele forms, by the presence of defined amino acid anchors, a number of specific pockets comprising a peptide-binding groove.[49] The binding affinities for disease-related peptides may thus be different for different HLA-DRB1 molecules as determined by their protein sequence. Subsequently this may influence composition of T-cell repertoires and ultimately result in *HLA-DRB1* alleles having varying effects on disease risk.[50] However, protein sequence analysis failed to provide unequivocal support for this hypothesis. Perhaps antigen presentation is not the sole mechanism of the MHC association in MS.[51]

Association of MS with the DR15 haplotype from the MHC on chromosome 6p21, has been established for more than 30 years[52] and is confirmed in nearly every population studied.[53]

IL7R

IL7R encodes for the receptor for interleukin-7, which is expressed on T and B-lymphocytes. The signal transduced by this receptor is crucial for lymphocyte survival and immune homeostasis. Initial functional data suggest a change in the rate of membrane-bound and soluble forms of the IL7RA for individuals carrying the risk allele.[54] It is still not known how this change could influence the function of the IL7RA.

Results of the first GWAS from the IMSGC showed a highly significant association ($P = 3 \times 10^{-7}$) of MS with a SNP in the *IL7R* gene,[28] although not reaching genome-wide significance. Strikingly, the same SNP was almost simultaneously identified and replicated in a different study (in total 2,853 cases, 3,204 unrelated controls and parents of 1,338 MS patients, combined *P*-value $= 2.9 \times 10^{-7}$) which used genomic convergence (multifactorial, multistep approach that combines gene expression with genomic linkage analysis) to select candidate genes that were differentially expressed between MS patients and controls.[54] Since then the association of MS with *IL7R* has been confirmed in several populations.[38, 55-59] In two studies,[38,58] including our meta-analysis (described in **chapter 4.2**) a genome-wide significant association with MS was found.

Apart from the *HLA* region to date *IL7R* clearly represents the gene with most evidence for an association with MS susceptibility. However, a recent study of African Americans did not provide evidence for an association in that population.[60]

IL2RA

IL2RA encodes for the alpha chain of the interleukin-2 receptor (IL2R, also known as CD25). The involvement of the IL2R chain in the pathogenesis of MS might be related to the important role that the IL2–IL2R pathway plays in adaptive immune functions. For MS patients carrying the *IL2RA* variant a change in the rate of membrane-bound and soluble forms of the IL2R was found.[61,62]

In the GWAS of the IMSGC, besides the *HLA* region, only one SNP in the *IL2RA* gene achieved genome-wide significant association with MS ($P = 3 \times 10^{-8}$).[28] Yet, several groups have reported an association of SNPs within *IL2RA* with MS,[37,38,56,59,63,64] although the associated variations varied across

populations. In our meta-analysis (**chapter 4.2**), we also showed a genome-wide significant association of *IL2RA* with susceptibility for MS ($P = 7 \times 10^{-11}$).[38] It has not yet been possible to pin-point the exact causative variation(s) in this gene or to detect functional relevance. In a large Australian population a fine-mapping study suggested allelic heterogeneity at the *IL2RA* locus and the existence of at least two independent susceptibility alleles.[63]

CD58

CD58 encodes a ligand for the T-cell specific CD2 membrane molecule, an adhesion molecule that transduces important signals for T-cell proliferation and differentiation. The *CD58* genetic variant that is most strongly associated with MS has a protective effect on MS development. Functional analyses revealed that this variant was associated with increased mRNA expression in cell lines as well as in mononuclear cells from MS patients. Furthermore, *CD58* mRNA levels are elevated in MS subjects during clinical remission, again supporting the hypothesis of a protective effect.[65]

In the first GWAS of the IMSGC, evidence was already found for an association of *CD58* with susceptibility for MS ($P = 2 \times 10^{-5}$). The Australian and New Zealand Multiple Sclerosis Genetics Consortium showed a genome-wide significant association of *CD58* with MS in the combined analysis of their GWAS and their replication study (in total 3,874 cases and 5,723 controls, P combined$=10 \times 10^{-8}$).[40] Furthermore a meta-analysis of GWAS (2,624 MS cases and 7,220 controls) and subsequent replication in an independent set of 2,215 MS patients and 2,116 controls from the United States (US) and United Kingdom (UK) showed a genome-wide significant association ($P = 3 \times 10^{-10}$)[39], which is in line with our meta-analysis ($P = 4 \times 10^{-9}$)[38] (**chapter 4.2**).

CLEC16A

For the CLEC16A protein still little is known regarding its function in humans. However, it is very likely that it does have a function in immune function. It is a member of the C-type lectin receptor family, of which members have described to provide signals for a decision between tolerance and immunity. They can bind bacterial products as well as endogenous ligands and their signal can counteract the signal of Toll-like receptors, therewith influencing T-helper cell function. Our group previously implied a role for C-type lectin receptors in MS pathogenesis, and discussed a link with infections.[66]

In the GWAS performed by the IMSGC, a SNP in intron 22 of the *CLEC16A* gene showed evidence for an association with MS ($P = 4 \times 10^{-6}$).[28] In the recently published meta-analysis and replication study in MS patients from the UK and the US, the SNP rs11865121 in the *CLEC16A* gene was found to be associated with susceptibility to MS ($P = 2 \times 10^{-7}$) in the joint analysis.[39] In our meta-analysis (**chapter 4.2**), we showed genome-wide significant association for the same SNP that showed evidence for association with MS in the GWAS performed by the IMSGC ($P = 3 \times 10^{-12}$).[28,38] In the Sardinian population an association with MS was revealed (1,498 MS cases and 1,706 matched controls) of a SNP in intron 19 ($P = 7 \times 10^{-5}$)[67] and in a large Caucasian replication cohort (2,369 trio families, 5,737 cases and 10,296 unrelated controls) of another SNP in intron 19 ($P = 2 \times 10^{-16}$).[36] Both SNPs have already been associated with Type 1 diabetes[68,69] and are in perfect LD with each other ($r^2=1$). The *CLEC16A* SNP in intron 22, that was identified in our meta-analysis to be genome-wide significantly associated with MS, is in a different

haplotype block (R^2=0.2). This suggests that in a single gene different SNPs are involved in different autoimmune disorders.

High evidence for an association

CD226

CD226, expressed on natural killer cells, T cells, monocytes and subsets of B cells and platelets, is a transmembrane receptor of the immunoglobulin receptor family. It is implicated in T cell and NK cell mediated cytotoxicity and platelet activation. CD226 recognizes on most cell types, including neurons, endothelial cells and fibroblast, expressed poliovirus receptor and nectin-2.[70] Interaction of CD226 with poliovirus receptor on endothelial cells is involved in endothelial transmigration of leukocytes.[71] In experimental autoimmune encephalomyelitis (EAE), a rodent model for MS, anti-CD226 treatment results in reduced severity and delayed onset of the disease. This could be the result of altered T-cell activation, altered monocyte extravasation or altered natural killer cell responsiveness towards target cells.[72]

An association with a coding SNP in the *CD226* gene (Gly307Ser) was first reported for Type 1 diabetes.[69] Subsequent studies also showed an association of this gene with susceptibility for MS.[36,73-75] In the study by the IMSGC (2,369 trio families, 5,737 cases and 10,296 unrelated controls) a genome-wide significant association was found ($P = 5 \times 10^{-8}$).[36]

CD6

CD6 is a lymphocyte receptor that belongs to the scavenger receptor cysteine-rich superfamily. Its expression has also been reported in certain regions of the brain. CD6 contributes to either positive or negative modulation of the activation and differentiation signals for lymphocytes.[87,88]

A genome-wide significant association of *CD6* with MS was identified by pooling together data from three separate GWAS exploring susceptibility alleles for MS (total of 2,624 cases and 7,220 controls) and replication in an independent US/UK cohort (2,215 cases and 2,116 controls, combined $P = 4 \times 10^{-09}$).[39] In one other recent replication study in Spanish MS patients (2,515 cases and 2,942 controls) an association of *CD6* with MS was also identified (P=0.004).[89]

TNFRSF1A

TNFRSF1A encodes the p55 receptor for tumor necrosis factor alpha (TNF-α). Functional studies in human suggest that dysregulation of the TNF-α pathway has a role in the onset of MS. Diminished TNF-α activity seems to be associated with onset of CNS inflammatory lesions in clinical studies.[90]

The meta-analysis and replication study from de Jager et al.[39] showed (besides *CD6*) a genome-wide significant association of *TNFRSF1A* with susceptibility for MS (combined $P = 2 \times 10^{-11}$). In the Spanish replication study[89] a significant association of MS with *TNFRSF1A* was also identified (P=0.001). In both studies evidence was found for the existence of two independent risk effects at this locus.

TYK2

TYK2 is a proximal tyrosine kinase in the STAT signalling pathway that is important for signalling by type I interferons and induction of Th1 cell differentiation upon antigen stimulation of dendritic cells.[91] The

protective effect of the rare variant (MAF=0.04) in the *TYK2* gene is predicted to give a less active variant of TYK2, which could lead to a shift to a protective Th2 response, and/or reduced Th1 or Th17 activation in MS.[91]

Alleles with a MAF below 0.05 are seldom included in GWAS because of the difficulties in detecting moderate genetic effects at this frequency level.[7] In a combined analysis of results of a GWAS using non-synonymous coding SNPs (ns-SNPs) for typing patients with autoimmune inflammatory thyroiditis, ankylosing spondylitis and MS, a rare variant of the *TYK2* gene was identified as a possible susceptibility factor for MS.[41] By adding 4,234 MS patients and 2,983 controls and 2,053 trio families to the analysis substantial evidence was found for an association of *TYK2* with MS (combined $P = 3 \times 10^{-06}$).[42] Because of the low frequency of the allele a genome-wide significant association of this gene with MS is very hard to establish. Therefore, an additional independent set of 5,429 MS patients and 6,167 controls was genotyped. By combining the data from all the three studies (10,642 cases, 10,620 controls and 2,110 MS trios) a convincing level of genome-wide significance was reached ($P = 5 \times 10^{-09}$).[43]

KIF21B

KIF21B is a kinesin-like protein involved in axonal transport.[92] *KIF21B* is expressed in a variety of immune cells. Although *KIF21B* has not been functionally associated with neurodegeneration or inflammation, given the nature and role of its protein in neurons, there is a plausible biological role for this gene in MS.

In a follow-up study of the IMSGC genome-wide significant association of *KIF21B* with susceptibility for MS was found. In this study, approximately 30,000 SNPs were genotyped, that demonstrated mild-to-moderate levels of significance ($P \leq 0.10$) in their initial GWAS,[28] in an independent set of 1,343 MS cases and 1,379 controls. Several of the most significant findings were replicated in another independent data set of 2,164 MS cases and 2,016 controls. In a combined analysis, using data from the original screen (in total 931 trios, 3,507 cases and 8,024 controls) a combined *P*-value for *KIF21B* of 7×10^{-10} was found.[45] Furthermore, in a recent Belgian study (791 cases 1,098 controls) an association of this gene with MS was found ($P = 0.01$).[46]

Modest evidence for an association

KIF1B

KIF1B encodes a kinesin superfamily member believed to be responsible for axonal transport of mitochondria and synaptic vesicle precursors.[76,77] It has an ATP-ase binding domain and is enriched in motor neurons. Recently, dysregulation of ATP-ases and mitochondrial mislocalization has been shown to play a role in several neurodegenerative diseases.[76] There is now increasing evidence that neurodegenerative processes, besides immunologic processes, are important in MS pathology. Irreversible axonal loss is an important mechanism in the development of permanent neurological symptoms.[78,79] Though primary demyelination may underlie early axonal loss, further progression of neurodegeneration occurs when the compensatory capacity of the CNS is exceeded and the threshold of axonal loss is reached.[78] Mechanisms proposed for this loss of nerve fibers include mitochondrial dysfunction, reduced ATP production and altered expression of sodium channels.[80-82] *KIF1B* knockout

mice have CNS abnormalities such as atrophy.[83]

We conducted a GWAS in 45 MS patients and 195 controls from the GRIP population and described the results in **chapter 5**. In the screening phase no SNP revealed genome-wide significant association with MS. However, we did detect a signal in the *KIF1B* locus, which was characterized by multiple associated SNPs. After multivariate logistic regression analysis, only SNP rs10492972 remained significantly, although not genome-wide significantly, associated with MS. We replicated that SNP in three replication sets of MS patients: (1) a cohort from the Dutch general population, (2) a cohort from the Swedish general population and (3) trio families from the Canadian Collaborative Project on the Genetic Susceptibility to MS. We studied 2,679 patients and 3,125 controls in the replication phase. A meta-analysis of the replication results of the four independent populations resulted in genome-wide significant association of the rs10492972[C] SNP with MS ($P = 3 \times 10^{-10}$). The a priori power to detect *KIF1B* was low, even in a genetically isolated population, regarding the small numbers of patients used for the GWAS. The ability to locate this gene is most probably due to the extended LD around the *KIF1B* SNP rs10492972 in the isolate.[84]

However, in another study (by the IMSGC) this association of *KIF1B* with MS could not be replicated.[85] In this study the rs10492972[C] variant of the *KIF1B* gene was genotyped in eight case-control and three trio family collections (in total 22,854 individuals; 8,391 cases, 8,052 unrelated controls and 2,137 trio families). None of these populations showed any evidence of statistically significant association (Figure 1). In an Italian cohort rs10492972 has been genotyped in an outbred sample of 222 primary progressive and progressive relapsing MS patients and 221 healthy controls of unique northern Italian origin. Carrying the rs10492972[C] variant had no effect on risk of disease or rate of disability progression.[85]

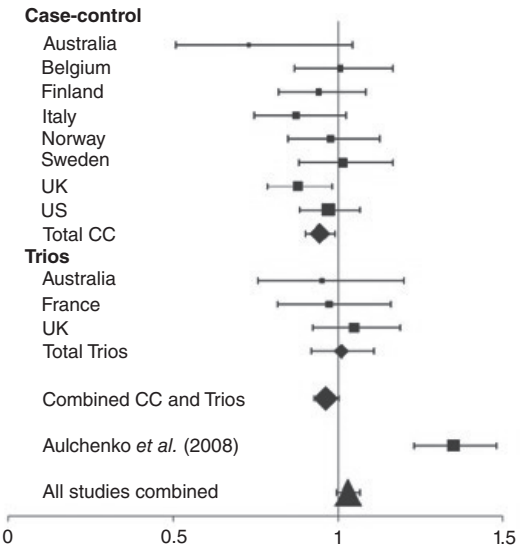


Figure 1 Odds ratio for the rs10492972[C] allele. The area of the symbol is proportional to the number of cases included in the respective analysis. The error bars indicate the 95% CI. Data derived from Booth et al.[85]

Although the association found between the *KIF1B* gene and MS has not been replicated, there is certainly not a lack of biological plausibility for an association of *KIF1B* and MS, given the increasingly recognized role of kinesins in neurological disease and the recent report on KIF1B and myelination.[86]

As described another recent study from the IMSGC demonstrated quite strong evidence for an association of MS with *KIF21B,*[45] a molecule sharing properties with *KIF1B*. A recent Australian-New Zealand screen found marginal association with an area around the *KIF1A* gene.[40] Furthermore, more evidence that there does indeed exist a real association of *KIF1B* with MS comes from an unpublished study in an independent group of Canadian multiplex families (Ebers), in which an association of MS with rs10492972[C] was shown.

We were surprised that our findings could not be replicated by the study of the IMSGC. There might be several explanations for this. A striking difference between the Italian study and the study done by the IMSGC versus our studies is the difference of allele frequencies in controls of rs10492972[C] (0.27 in our studies versus 0.33 in the Italian study and 0.34 in the IMSGC consortium). Especially the controls from the genetic isolate in our population appear to have a low allele frequency, which may suggest that our control group was biased in some way. This may be attributed to founder effects, random effects that may occur in some isolates. Confounding due to population stratification is not a very likely explanation as it should have occurred in the four different populations that we analysed. Furthermore, in our paper we used EIGENSTRAT analysis to adjust for this type of confounding and this type of confounding cannot affect the analysis of the multiplex MS families from Canada and the UK.

Genotyping errors, which are often revealed as deviations from Hardy-Weinberg Equilibrium (HWE), could also be an explanation for the different study results. Genotype- and-allele frequencies are not in HWE for the UK cohort (P =0.04) and are borderline for the Italian cohort (P =0.08) of the IMSGC consortium. This could play a role in terms of the genotype distributions seen in cases and controls in the various cohorts of the IMSGC.

Further, the different results could be the consequence of prevalence-incidence bias. A gene that is associated with survival after onset of the disease may result in a wide range of odds ratios in different studies, because of subtle differences in inclusion procedures. The Dutch population cohort was ascertained relatively early in the disease phase and a borderline significant association between *KIF1B* and age at diagnosis (P =0.09) was seen with age at diagnosis being 2.8 years later in patients with the *KIF1B CC* genotype. Adjusting for these age differences, a nominal significant association (P =0.04) still remained between *KIF1B* and the time between age at diagnosis and age at study inclusion, the *KIF1BCC* genotype being associated to a shorter period between diagnosis and study inclusion.

The differences in allele frequencies in controls across populations of our study versus the IMSGC, the deviations from HWE seen in controls from the IMSGC and the possibility of incidence-prevalence bias together, calls for further studies. Finally, deep sequencing of *KIF1B* may also help in resolving the issue.

IRF8

IRF8 is a member of the interferon regulatory factor (IRF) family, which is specifically expressed in immune

cells. The *IRF8* locus contains an important transcription factor involved in responses to type I interferons (IFNα and β). Upregulation of interferon responses has been noted in peripheral blood of a subset of untreated MS patients.[93,94] However, the role of interferons in the onset of MS is still unclear. Results of a recently performed expression study show that the susceptibility allele near *IRF8* is associated with higher mRNA expression of interferon-response pathway genes in subjects with MS.[39]

A genome-wide significant association of *IRF8* with susceptibility for MS was identified in the recently published meta-analysis (combined $P = 4 \times 10^{-09}$).[39] So far, only one replication study for *IRF8* has been performed. In this study no association was found, which could also be due to lack of power (1,271 cases and 1,525 controls).[89]

STAT3

STAT3 encodes a transcription factor that is involved in multiple pathways and functions, including the Jak-STAT pathway, neuron axonal guidance, apoptosis, activation of immune responses and Th17 cell differentiation.[95] Mouse studies have shown that the targeted deletion of *STAT3* in CD4+ T-cells prevents the development of experimental autoimmune encephalomyelitis.[96] Moreover, increased phosphorylated STAT3 was reported in T cells of patients evolving from a clinically isolated syndrome to defined MS and in relapsing patients.[97]

Evidence for an association of *STAT3* with MS came from a GWAS, conducted in 68 distantly related MS patients and 136 controls from the high-risk isolated population of Southern Ostrobothnia in Finland. The loci with the lowest *P*-values (all $< 10 \times 10^{-04}$) were replicated in 711 MS patients and 1,029 controls from Finland (83 MS patients and 365 controls were from the isolate in Finland). Subsequently, the top two SNPs were validated in 3,859 MS patients and 9,110 controls from more heterogeneous populations. In the combined analysis a genome-wide significant association of a common variant in *STAT3* with MS was identified (combined $P = 3 \times 10^{-10}$). This finding demonstrates the power of using a genetic isolate to complement large-scale GWAS in identifying common genes and not only in identifying rare high-impact alleles.[44] However, in a replication study in Spanish MS patients no association of *STAT3* with MS was detected (1,540 cases and 1,720 controls).[98]

TMEM39A

Almost nothing is known about *TMEM39A* (mRNA-transmembrane protein 39A) and what biological role it might play with regard to MS susceptibility.

In the follow-up study of the IMSGC, *TMEM39A* was found to be genome-wide significantly associated with MS (combined $P = 3 \times 10^{-08}$).[45] Further studies have to be performed to strengthen the evidence for an association of this gene with MS.

CYP27B1

CYP27B1 encodes the enzyme 25-hydroxyvitamin D-1 alpha hydroxylase, which hydroxylates 25-hydroxyvitamin D into the bioactive form (1,25-dihydroxyvitamin D (1,25(OH)2D). This bioactive form regulates through the vitamin D receptor (VDR) the calcium metabolism. It also has important immune functions modulating innate as well as adaptive immunity and tolerance[99] and B-cell homeostasis.[100] Furthermore it can direct activated T cells toward a T helper type 2 anti-inflammatory phenotype and

induce dendritic cells with tolerogenic properties.[101] Epidemiological data suggest a link between vitamin D deficiency and increased incidence of MS and other autoimmune diseases.[102-104]

The Australia and New Zealand Multiple Sclerosis Genetics Consortium (ANZgene) performed a GWAS in 1,618 MS patients and 3,413 controls. Replication was performed in an independent set of 2,256 cases and 2,310 controls. Two SNPs at chromosome 12 were associated at the genome-wide level in the combined cohort ($P = 5 \times 10^{-11}$ and $P = 3 \times 10^{-10}$). The strongest candidate gene given available genetic, immunological and epidemiological evidence was considered to be *CYP27B1*.[40]

Systems biology

Many more markers tested in GWAS have shown association with MS, although not genome-wide significant. Because of the large multiple testing involved in these studies, very few exceed the genome-wide significant threshold and those that do not are generally neglected. In many cases where loci with small but measurable genetic effects are involved, it is likely that, accepting the null hypothesis of no association, these loci represent false negatives. It is possible that meaningful combinations of genes, that showed only modest evidence of association in GWAS, can be identified if they belong to the same biological pathway or mechanism.

In systems biology an integrative approach is used of available molecular, physiological, clinical, genetic, gene expression, proteomic and/ or biological information with the ultimate goal to identify pathways involved in the pathogenesis of diseases. Thereby real-world systems are represented as networks that interconnect different entities. In this way certain properties can emerge that cannot be derived from the individual analysis of each of their components.

Baranzini and colleagues have recently put current information on modest genetic effects in MS into a model linked to a human protein network. The aim of their study was to identify sub-networks containing a higher proportion of genes associated with MS than expected by chance. Indeed interesting sub-networks of genes from several immunologic pathways involved in MS susceptibility were identified. Part of them were already known, but also novel neural pathways were uncovered.[105]

MS and other autoimmune diseases

Intriguing is the fact that most of the above identified immune related SNPs that showed genome-wide significant association with MS, have overlapping associations with autoimmune conditions such as rheumatoid arthritis, inflammatory bowel disease, SLE, Type 1 diabetes and thyroid autoimmunity (Table 1 of the introduction of this thesis).

This suggests that in MS and other autoimmune diseases, common pathways are involved in pathogenesis. When different autoimmune disorders share susceptibility genes, one would expect relatives to be not only at greater risk for MS than the general population, but also at greater risk of other autoimmune disorders. Several studies have investigated the rate of autoimmune diseases in MS patients and their biological relatives.[106-109] The results have been conflicting, which may relate to differences in methodologies and populations studied.

It is striking to note that some immunomodulatory treatments in MS can trigger other autoimmune diseases. The results of a phase II clinical trial with the lymphocyte-depleting humanized monoclonal antibody alemtuzumab (Campath-1H; targeting CD52 a protein present on the surface of mature lymphocytes) show that it is highly effective in the treatment of early relapsing-remitting MS.[110] A single pulse of treatment results in rapid, profound and prolonged lymphopenia. Especially CD4+ T cells recover slowly, remaining depleted for at least 5 years.[111] Thirty % of patients developed autoimmunity months to years after pulsed treatment with alemtuzumab. Most people develop thyroid autoimmunity, mainly Graves' disease, and a few idiopathic thrombocytopenic purpura (ITP) and in a very few other blood components are targeted. In addition 5.5 % of patients develop sustained non-thyroid autoantibodies without clinical disease.[111] While the association between lymphopenia and autoimmunity is well recognized, most lymphopenic subjects do not develop autoimmunity. It has been shown that autoimmunity arose in those patients with greater T-cell apoptosis and cell cycling in response to treatment. This phenomenon is driven by higher levels of IL-21. Patients who went on to develop autoimmunity had more than 2-fold higher levels of serum IL-21 than the patients who did not. The production of IL-21 seems to be genetically determined.[110]

These findings may suggest that by driving cycles of T cell expansion and apoptosis to excess, IL-21 increases the stochastic opportunities for T cells to encounter self antigen and so for autoimmunity.[110]

Near future

The identified risk variants only explain part of the total genetic risk of MS. With the exception of *TYK2*, the identified risk variants are all common. Larger GWAS will probably result in identifying more common risk variants with modest effects on disease risk. As alleles with a MAF below 0.05 are seldom included in GWAS because of the difficulties in detecting moderate genetic effects at this frequency level,[7] rare variants are in general not identified with these GWAS. Whole exon sequencing or whole genome sequencing will be necessary to identify risk alleles with a frequency lower than 5% with effects in between common SNPs and very rare alleles (mutations).

The newly identified genetic associations support the view that MS is an autoimmune disease in which T-cells and inflammatory responses have a dominant role. However, the GWAS described in **chapter 5** suggests that the neurodegenerative phase is also influenced by neuronal genes. The next step in MS genetics will be to identify the real disease-mediating variants in the recently discovered genetic risk loci. One of the major tools will be DNA sequencing around the current SNPs, or even at the whole genome level.

The identification of the exact disease susceptibility genes does not necessarily define the pathway involved in disease development or the molecular pathway the gene product is involved in. To determine this, functional studies are necessary. Potential effects of disease-associated variants could first be analysed in cells and tissues from healthy controls that may or may not have the predisposing genotype. Next, functional and expression analyses can be done, as the newly discovered genetic variants may be located in either known or putative regulatory elements, or coding regions. The

molecular pathways where the genetic variants are involved in should of course be analysed in-depth. The simultaneous quantitative analyses of gene expression will be required as more risk genes become known. And of course the genetic variation on a genome-wide basis in a large number of individuals will provide a more thorough understanding of the molecular networks that are perturbed by the disease associated genetic variants.[112]

Obtaining immune cells for such functional studies is much easier than accessing human CNS cells or tissue. Recent developments in the field of induced pluripotent stem cells (iPS cells) may change this situation. Those cells can now be generated from fibroblasts,[113-115] keratinocytes[116] and blood progenitor cells[117] and are likely to revolutionize the generation of *in vitro* disease models.

However, variation in gene expression is also affected by factors as epigenetic modification, factors that are not readily identified by SNP or DNA sequencing studies. Epigenetic modifications are markers of how environmental factors may influence gene expression and may be very important in linking genetics to environment. The identification of epigenetic processes in demyelination[118] indicates that investigating dysregulated posttranslational modifications in MS could result in a better understanding of its pathogenesis.

Furthermore, the network based approach in system biology can result in better knowledge of pathogenesis of MS and finally to better therapies. The genomic portrait of an individual may allow a predictive and personalized approach to therapy. In the near future this process will probably be aided by powerful computers, with information technologies that will manage the available information from the patient's tests, the patient's medical history and genetic history and ever evolving scientific databases.

Interaction between identified risk genes

Between the *EVI5-RPL5* locus and *HLA-DR15* a modest statistical interaction was found, manifesting as an association in the *HLA-DR15* non-carriers, but not in the *HLA-DR15* carriers.[40] Further studies will be required to further analyse these possible interactions and also possible interactions between other identified risk alleles.

Gene-environment interactions

A challenging element of future studies will be the characterization of the environmental factors that interact with genetic factors to finally cause the disease.

As I already described in the introduction, an interaction between vitamin D and *HLA-DRB1*15* and between EBV and *HLA-DRB1*15* have been shown before. Besides EBV and vitamin D numerous bacterial and viral infections have been described as potential candidates, as they are often associated with relapses of MS. However, no single infection has consistently been associated with disease.[119] It is unclear how so many different infections could have a role in the pathogenesis of the disease and how they might interact with genetic factors, unless their role is non-specific, such as in creating inflammation or damaging the blood-brain barrier.

A better understanding of the function of identified risk genes and their role in pathogenesis

of MS may not only explain how suspected environmental factors interact with such genetic factors to cause disease, but could also result in the identification of additional environmental risk factors.

Pharmacogenomics

Current therapies can prevent the appearance of new lesions in the central nerve system (CNS) of MS patients and moderate the relapsing-remitting phase, but they have much less effect on the progression of the disease and long-term disability. Clinical response to MS medication varies substantially among individual patients. A better understanding of the etiology of MS and the pathogenesis should provide rational bases for developing new drugs for MS. Pharmacogenomic approaches aim at identifying genetic variations that affect the response to certain drugs and thus may help the physician to decide which patients may benefit from a certain therapy and which patients are prone to develop adverse side effects.[120]

IFNß is the most widely used disease modifying therapy for the treatment of MS. However its main benefits, fewer and less severe relapses as well as delayed disease progression, are seen in approximately 50% of patients. Thus, effectively selecting potential responders and non-responders *before* initiating this costly therapy would be preferred by patients and clinicians.[120] Recently, the first two GWAS were done in an attempt to identify response-predictive SNPs. In these GWAS evidence was found for a potential correlation between genes that code for neurotransmitter-gated channels and a response to IFNb therapy.[121,122] However, the results of these studies need to be replicated before they can have clinical consequences, but it brings the concept of personalized medicine somewhat closer to the MS patients.

Novel genes in animal models

Work in animal models can also result in a more comprehensive understanding of the effects of individual genetic variations on disease development. Through the EAE model, relevant molecular pathways in MS have already been dissected.[123] This has led to the development of new therapies, most notably natalizumab (Tysabri), an α4 integrin-specific antibody that blocks T cell entry into the CNS.[124]

The roles of predisposing genetic variants in the pathogenesis of MS can be investigated further by the development of a humanized form of the EAE model in which the animals express transgenes derived from MS patients. Humanized mice have already been used to analyse the varying effects of different HLA genes in MS and their conclusions are supported by complementary genetic data from human populations. However, so far none of the newly identified non-MHC genetic risk variants has been analysed in animal models for MS, but this will become possible in the near future.[125]

References

1. Barcellos LF, Sawcer S, Ramsay PP, et al. Heterogeneity at the HLA-DRB1 locus and risk for multiple sclerosis. Hum Mol Genet 2006;15(18):2813-24.
2. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996;273(5281):1516-7.
3. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461(7265):747-53.
4. Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet Epidemiol 2008;32(4):381-5.
5. Cardon LR, Palmer LJ. Population stratification and spurious allelic association. Lancet 2003;361(9357):598-604.
6. Hattersley AT, McCarthy MI. What makes a good genetic association study? Lancet 2005;366(9493):1315-23.
7. Wang WY, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. Nat Rev Genet 2005;6(2):109-18.
8. Kuokkanen S, Gschwend M, Rioux JD, et al. Genomewide scan of multiple sclerosis in Finnish multiplex families. Am J Hum Genet 1997;61(6):1379-87.
9. Wright AF, Carothers AD, Pirastu M. Population choice in mapping genes for complex diseases. Nat Genet 1999;23(4):397-404.
10. Aulchenko YS, Vaessen N, Heutink P, et al. A genome-wide search for genes involved in type 2 diabetes in a recently genetically isolated population from the Netherlands. Diabetes 2003;52(12):3001-4.
11. Varilo T, Peltonen L. Isolates and their potential use in complex gene mapping efforts. Curr Opin Genet Dev 2004;14(3):316-23.
12. Kallio SP, Jakkula E, Purcell S, et al. Use of a genetic isolate to identify rare disease variants: C7 on 5p associated with MS. Human molecular genetics 2009;18(9):1670-83.
13. Sajantila A, Salem AH, Savolainen P, Bauer K, Gierig C, Paabo S. Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. Proc Natl Acad Sci U S A 1996;93(21):12035-9.
14. Kittles RA, Perola M, Peltonen L, et al. Dual origins of Finns revealed by Y chromosome haplotype variation. Am J Hum Genet 1998;62(5):1171-9.
15. Peltonen L, Palotie A, Lange K. Use of population isolates for mapping complex traits. Nat Rev Genet 2000;1(3):182-90.
16. Dyment DA, Ebers GC, Sadovnick AD. Genetics of multiple sclerosis. Lancet Neurol 2004;3(2):104-10.
17. Compston A, Coles A. Multiple sclerosis. Lancet 2002;359(9313):1221-31.
18. Dyment DA, Cader MZ, Willer CJ, Risch N, Sadovnick AD, Ebers GC. A multigenerational family with multiple sclerosis. Brain 2002;125(Pt 7):1474-82.
19. Modin H, Masterman T, Thorlacius T, et al. Genome-wide linkage screen of a consanguineous multiple sclerosis kinship. Mult Scler 2003;9(2):128-34.
20. Sadovnick AD, Risch NJ, Ebers GC. Canadian collaborative project on genetic susceptibility to MS, phase 2: rationale and method. Canadian Collaborative Study Group. Can J Neurol Sci 1998;25(3):216-21.
21. Fernald GH, Yeh RF, Hauser SL, Oksenberg JR, Baranzini SE. Mapping gene activity in complex disorders: Integration of expression and genomic scans for multiple sclerosis. J Neuroimmunol 2005;167(1-2):157-69.
22. Sawcer S, Ban M, Maranian M, et al. A high-density screen for linkage in multiple sclerosis. Am J Hum Genet 2005;77(3):454-67.
23. Risch N. Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 1990;46(2):222-8.
24. Goring HH, Terwilliger JD, Blangero J. Large upward bias in estimation of locus-specific effects from genomewide scans. Am J Hum Genet 2001;69(6):1357-69.
25. Ebers GC, Koopman WJ, Hader W, et al. The natural history of multiple sclerosis: a geographically based study: 8: familial multiple sclerosis. Brain 2000;123 Pt 3:641-9.
26. Willer CJ, Dyment DA, Cherny S, et al. A genome-wide scan in forty large pedigrees with multiple sclerosis. J Hum Genet 2007;52(12):955-62.

27. Dyment DA, Cader MZ, Herrera BM, et al. A genome scan in a single pedigree with a high prevalence of multiple sclerosis. J Neurol Neurosurg Psychiatry 2008;79(2):158-62.
28. Hafler DA, Compston A, Sawcer S, et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 2007;357(9):851-62.
29. D'Netto MJ, Ward H, Morrison KM, et al. Risk alleles for multiple sclerosis in multiplex families. Neurology 2009;72(23):1984-8.
30. Binzer M, Forsgren L, Holmgren G, Drugge U, Fredrikson S. Familial clustering of multiple sclerosis in a northern Swedish rural district. J Neurol Neurosurg Psychiatry 1994;57(4):497-9.
31. Callander M, Landtblom AM. A cluster of multiple sclerosis cases in Lysvik in the Swedish county of Varmland. Acta Neurol Scand 2004;110(1):14-22.
32. Marrosu MG, Lai M, Cocco E, et al. Genetic factors and the founder effect explain familial MS in Sardinia. Neurology 2002;58(2):283-8.
33. Tienari PJ, Sumelahti ML, Rantamaki T, Wikstrom J. Multiple sclerosis in western Finland: evidence for a founder effect. Clin Neurol Neurosurg 2004;106(3):175-9.
34. Hoppenbrouwers IA, Cortes LM, Aulchenko YS, et al. Familial clustering of multiple sclerosis in a Dutch genetic isolate. Mult Scler 2007;13(1):17-24.
35. Kuokkanen S, Sundvall M, Terwilliger JD, et al. A putative vulnerability locus to multiple sclerosis maps to 5p14-p12 in a region syntenic to the murine locus Eae2. Nat Genet 1996;13(4):477-80.
36. The expanding genetic overlap between multiple sclerosis and type I diabetes. Genes Immun 2009;10(1):11-4.
37. Rubio JP, Stankovich J, Field J, et al. Replication of KIAA0350, IL2RA, RPL5 and CD58 as multiple sclerosis susceptibility genes in Australians. Genes Immun 2008;9(7):624-30.
38. Hoppenbrouwers IA, Aulchenko YS, Janssens AC, et al. Replication of CD58 and CLEC16A as genome-wide significant risk genes for multiple sclerosis. J Hum Genet 2009;54(11):676-80.
39. 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(7):776-82.
40. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. Nat Genet 2009;41(7):824-8.
41. Burton PR, Clayton DG, Cardon LR, et al. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. Nat Genet 2007;39(11):1329-37.
42. Ban M, Goris A, Lorentzen AR, et al. Replication analysis identifies TYK2 as a multiple sclerosis susceptibility factor. Eur J Hum Genet 2009;17(10):1309-13.
43. Mero IL, Lorentzen AR, Ban M, et al. A rare variant of the TYK2 gene is confirmed to be associated with multiple sclerosis. Eur J Hum Genet 2009.
44. Jakkula E, Leppa V, Sulonen AM, et al. Genome-wide association study in a high-risk isolate for multiple sclerosis reveals associated variants in STAT3 gene. Am J Hum Genet;86(2):285-91.
45. Comprehensive follow-up of the first genome-wide association study of multiple sclerosis identifies KIF21B and TMEM39A as susceptibility loci. Hum Mol Genet;19(5):953-62.
46. Goris A, Boonen S, D'Hooghe M B, Dubois B. Replication of KIF21B as a susceptibility locus for multiple sclerosis. J Med Genet 2010, in press.
47. Hoppenbrouwers IA, Aulchenko YS, Ebers GC, et al. EVI5 is a risk gene for multiple sclerosis. Genes Immun 2008;9(4):334-7.
48. Watts C. The exogenous pathway for antigen presentation on major histocompatibility complex class II and CD1 molecules. Nat Immunol 2004;5(7):685-92.
49. Jones EY, Fugger L, Strominger JL, Siebold C. MHC class II proteins and disease: a structural perspective. Nat Rev Immunol 2006;6(4):271-82.
50. Ramagopalan SV, McMahon R, Dyment DA, Sadovnick AD, Ebers GC, Wittkowski KM. An extension to a statistical approach for family based association studies provides insights into genetic risk factors for multiple sclerosis in the HLA-DRB1 gene. BMC Med Genet 2009;10:10.
51. Hiremath MM, Chen VS, Suzuki K, Ting JP, Matsushima GK. MHC class II exacerbates demyelination in vivo independently of T cells. J Neuroimmunol 2008;203(1):23-32.
52. Jersild C, Svejgaard A, Fog T. HL-A antigens and multiple sclerosis. Lancet 1972;1(7762):1240-1.

53. Compston A CC, Lassmann et al. McAlpine's Multiple Sclerosis. Edinburgh: Churchill Livingstone; 2005.
54. 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(9):1083-91.
55. Lundmark F, Duvefelt K, Iacobaeus E, et al. Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. Nature Genet 2007;39(9):1108-13.
56. Weber F, Fontaine B, Cournu-Rebeix I, et al. IL2RA and IL7RA genes confer susceptibility for multiple sclerosis in two independent European populations. Genes Immun 2008;9(3):259-63.
57. O'Doherty C, Kantarci O, Vandenbroeck K. IL7RA polymorphisms and susceptibility to multiple sclerosis. N Engl J Med 2008;358(7):753-4.
58. Alcina A, Fedetz M, Ndagire D, et al. The T244I variant of the interleukin-7 receptor-alpha gene and multiple sclerosis. Tissue antigens 2008;72(2):158-61.
59. Akkad DA, Hoffjan S, Petrasch-Parwez E, Beygo J, Gold R, Epplen JT. Variation in the IL7RA and IL2RA genes in German multiple sclerosis patients. J Autoimmun 2009;32(2):110-5.
60. Johnson BA, Wang J, Taylor EM, et al. Multiple sclerosis susceptibility alleles in African Americans. Genes Immun 2010;11(4):343-50.
61. Maier LM, Anderson DE, Severson CA, et al. Soluble IL-2RA levels in multiple sclerosis subjects and the effect of soluble IL-2RA on immune responses. J Immunol 2009;182(3):1541-7.
62. Maier LM, Lowe CE, Cooper J, et al. IL2RA genetic heterogeneity in multiple sclerosis and type 1 diabetes susceptibility and soluble interleukin-2 receptor production. PLoS Genet 2009;5(1):e1000322.
63. Perera D, Stankovich J, Butzkueven H, et al. Fine mapping of multiple sclerosis susceptibility genes provides evidence of allelic heterogeneity at the IL2RA locus. J Neuroimmunol 2009;211(1-2):105-9.
64. Alcina A, Fedetz M, Ndagire D, et al. IL2RA/CD25 gene polymorphisms: uneven association with multiple sclerosis (MS) and type 1 diabetes (T1D). PloS one 2009;4(1):e4137.
65. De Jager PL, Baecher-Allan C, Maier LM, et al. The role of the CD58 locus in multiple sclerosis. Proc Natl Acad Sci U S A 2009;106(13):5264-9.
66. t Hart BA, Laman JD, Bauer J, Blezer E, van Kooyk Y, Hintzen RQ. Modelling of multiple sclerosis: lessons learned in a non-human primate. Lancet Neurol 2004;3(10):588-97.
67. Zoledziewska M, Costa G, Pitzalis M, et al. Variation within the CLEC16A gene shows consistent disease association with both multiple sclerosis and type 1 diabetes in Sardinia. Genes Immun 2009;10(1):15-7.
68. Hakonarson H, Grant SF, Bradfield JP, et al. A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. Nature 2007;448(7153):591-4.
69. Todd JA, Walker NM, Cooper JD, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet 2007;39(7):857-64.
70. Bottino C, Castriconi R, Moretta L, Moretta A. Cellular ligands of activating NK receptors. Trends Immunol 2005;26(4):221-6.
71. Reymond N, Imbert AM, Devilard E, et al. DNAM-1 and PVR regulate monocyte migration through endothelial junctions. J Exp Med 2004;199(10):1331-41.
72. Dardalhon V, Schubart AS, Reddy J, et al. CD226 is specifically expressed on the surface of Th1 cells and regulates their expansion and effector functions. J Immunol 2005;175(3):1558-65.
73. Hafler JP, Maier LM, Cooper JD, et al. CD226 Gly307Ser association with multiple autoimmune diseases. Genes Immun 2009;10(1):5-10.
74. Wieczorek S, Hoffjan S, Chan A, et al. Novel association of the CD226 (DNAM-1) Gly307Ser polymorphism in Wegener's granulomatosis and confirmation for multiple sclerosis in German patients. Genes Immun 2009;10(6):591-5.
75. Alcina A, Vandenbroeck K, Otaegui D, et al. The autoimmune disease-associated KIF5A, CD226 and SH2B3 gene variants confer susceptibility for multiple sclerosis. Genes Immun 2010;11(5):439-45.
76. Boldogh IR, Pon LA. Mitochondria on the move. Trends Cell Biol 2007;17(10):502-10.
77. Nangaku M, Sato-Yoshitake R, Okada Y, et al. KIF1B, a novel microtubule plus end-directed monomeric motor protein for transport of mitochondria. Cell 1994;79(7):1209-20.
78. Dutta R, Trapp BD. Pathogenesis of axonal and neuronal damage in multiple sclerosis. Neurology 2007;68(22 Suppl 3):S22-31; discussion S43-54.

79. DeLuca GC, Williams K, Evangelou N, Ebers GC, Esiri MM. The contribution of demyelination to axonal loss in multiple sclerosis. Brain 2006;129(Pt 6):1507-16.
80. Dutta R, McDonough J, Yin X, et al. Mitochondrial dysfunction as a cause of axonal degeneration in multiple sclerosis patients. Ann Neurol 2006;59(3):478-89.
81. Craner MJ, Newcombe J, Black JA, Hartle C, Cuzner ML, Waxman SG. Molecular changes in neurons in multiple sclerosis: altered axonal expression of Nav1.2 and Nav1.6 sodium channels and Na+/Ca2+ exchanger. Proc Natl Acad Sci U S A 2004;101(21):8168-73.
82. Waxman SG. Axonal dysfunction in chronic multiple sclerosis: meltdown in the membrane. Ann Neurol 2008;63(4):411-3.
83. Zhao C, Takita J, Tanaka Y, et al. Charcot-Marie-Tooth disease type 2A caused by mutation in a microtubule motor KIF1Bbeta. Cell 2001;105(5):587-97.
84. Aulchenko YS, Hoppenbrouwers IA, Ramagopalan SV, et al. Genetic variation in the KIF1B locus influences susceptibility to multiple sclerosis. Nat Genet 2008;40(12):1402-3.
85. Booth DR, Heard RN, Stewart GJ, et al. Lack of support for association between the KIF1B rs10492972[C] variant and multiple sclerosis. Nat Genet 2010;42(6):469-71.
86. Lyons DA, Naylor SG, Scholze A, Talbot WS. Kif1b is essential for mRNA localization in oligodendrocytes and development of myelinated axons. Nat Genet 2009;41(7):854-8.
87. Hassan NJ, Simmonds SJ, Clarkson NG, et al. CD6 regulates T-cell responses through activation-dependent recruitment of the positive regulator SLP-76. Mol Cell Biol 2006;26(17):6727-38.
88. Castro MA, Oliveira MI, Nunes RJ, et al. Extracellular isoforms of CD6 generated by alternative splicing regulate targeting of CD6 to the immunological synapse. J Immunol 2007;178(7):4351-61.
89. Swaminathan B, Matesanz F, Cavanillas ML, et al. Validation of the CD6 and TNFRSF1A loci as risk factors for multiple sclerosis in Spain. J Neuroimmunol 2010;223(1-2):100-3.
90. van Oosten BW, Barkhof F, Truyen L, et al. Increased MRI activity and immune activation in two multiple sclerosis patients treated with the monoclonal anti-tumor necrosis factor antibody cA2. Neurology 1996;47(6):1531-4.
91. Tokumasa N, Suto A, Kagami S, et al. Expression of Tyk2 in dendritic cells is required for IL-12, IL-23, and IFN-gamma production and the induction of Th1 cell differentiation. Blood 2007;110(2):553-60.
92. Marszalek JR, Weiner JA, Farlow SJ, Chun J, Goldstein LS. Novel dendritic kinesin sorting identified by different process targeting of two related kinesins: KIF21A and KIF21B. Journal Cell Biol 1999;145(3):469-79.
93. van Baarsen LG, van der Pouw Kraan TC, Kragt JJ, et al. A subtype of multiple sclerosis defined by an activated immune defense program. Genes Immun 2006;7(6):522-31.
94. Degre M, Dahl H, Vandvik B. Interferon in the serum and cerebrospinal fluid in patients with multiple sclerosis and other neurological disorders. Acta Neurol Scand 1976;53(2):152-60.
95. Egwuagu CE. STAT3 in CD4+ T helper cell differentiation and inflammatory diseases. Cytokine 2009;47(3):149-56.
96. Liu X, Lee YS, Yu CR, Egwuagu CE. Loss of STAT3 in CD4+ T cells prevents development of experimental autoimmune diseases. J Immunol 2008;180(9):6070-6.
97. Frisullo G, Nociti V, Iorio R, et al. The persistency of high levels of pSTAT3 expression in circulating CD4+ T cells from CIS patients favors the early conversion to clinically defined multiple sclerosis. J Neuroimmunol 2008;205(1-2):126-34.
98. Cenit MC, Alcina A, Marquez A, et al. STAT3 locus in inflammatory bowel disease and multiple sclerosis susceptibility. Genes Immun;11(3):264-8.
99. Adorini L, Penna G. Control of autoimmune diseases by the vitamin D endocrine system. Nature Clin Pract Rheumatol 2008;4(8):404-12.
100. Chen S, Sims GP, Chen XX, Gu YY, Chen S, Lipsky PE. Modulatory effects of 1,25-dihydroxyvitamin D3 on human B cell differentiation. J Immunol 2007;179(3):1634-47.
101. Boonstra A, Barrat FJ, Crain C, Heath VL, Savelkoul HF, O'Garra A. 1alpha,25-Dihydroxyvitamin d3 has a direct effect on naive CD4(+) T cells to enhance the development of Th2 cells. J Immunol 2001;167(9):4974-80.
102. Munger KL, Levin LI, Hollis BW, Howard NS, Ascherio A. Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. JAMA 2006;296(23):2832-8.
103. Holick MF. Vitamin D deficiency. N Engl J Med 2007;357(3):266-81.

104. van der Mei IA, Ponsonby AL, Dwyer T, et al. Vitamin D levels in people with multiple sclerosis and community controls in Tasmania, Australia. J Neurol 2007;254(5):581-90.

105. Baranzini SE, Galwey NW, Wang J, et al. Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. Hum Mol Genet 2009;18(11):2078-90.

106. Heinzlef O, Alamowitch S, Sazdovitch V, et al. Autoimmune diseases in families of French patients with multiple sclerosis. Acta Neurol Scand 2000;101(1):36-40.

107. Barcellos LF, Kamdar BB, Ramsay PP, et al. Clustering of autoimmune diseases in families with a high-risk for multiple sclerosis: a descriptive study. Lancet Neurol 2006;5(11):924-31.

108. Henderson RD, Bain CJ, Pender MP. The occurrence of autoimmune diseases in patients with multiple sclerosis and their families. J Clin Neurosci 2000;7(5):434-7.

109. Broadley SA, Deans J, Sawcer SJ, Clayton D, Compston DA. Autoimmune disease in first-degree relatives of patients with multiple sclerosis. A UK survey. Brain 2000;123(Pt 6):1102-11.

110. Jones JL, Phuah CL, Cox AL, et al. IL-21 drives secondary autoimmunity in patients with multiple sclerosis, following therapeutic lymphocyte depletion with alemtuzumab (Campath-1H). J Clin Invest 2009;119(7):2052-61.

111. Coles AJ, Cox A, Le Page E, et al. The window of therapeutic opportunity in multiple sclerosis: evidence from monoclonal antibody therapy. J Neurol 2006;253(1):98-108.

112. Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M. Mapping complex disease traits with global gene expression. Nat Rev Genet 2009;10(3):184-94.

113. Park IH, Zhao R, West JA, et al. Reprogramming of human somatic cells to pluripotency with defined factors. Nature 2008;451(7175):141-6.

114. Takahashi K, Tanabe K, Ohnuki M, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007;131(5):861-72.

115. Yu J, Vodyanik MA, Smuga-Otto K, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science 2007;318(5858):1917-20.

116. Aasen T, Raya A, Barrero MJ, et al. Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. Nat Biotechnol 2008;26(11):1276-84.

117. Loh YH, Agarwal S, Park IH, et al. Generation of induced pluripotent stem cells from human blood. Blood 2009;113(22):5476-9.

118. Shen S, Sandoval J, Swiss VA, et al. Age-dependent epigenetic control of differentiation inhibitors is critical for remyelination efficiency. Nat Neurosci 2008;11(9):1024-34.

119. Correale J, Fiol M, Gilmore W. The risk of relapses in multiple sclerosis during systemic infections. Neurology 2006;67(4):652-9.

120. Miller A, Avidan N, Tzunz-Henig N, et al. Translation towards personalized medicine in multiple sclerosis. J Neurol Sci 2008;274:68-75.

121. Byun E, Caillier SJ, Montalban X, et al. Genome-wide pharmacogenomic analysis of the response to interferon beta therapy in MS. Arch Neurol 2008;65:337-44.

122. Cenit MD, Blanco-Kelly F, de las Heras V, et al. Glypican 5 is an interferon-beta response gene: a replication study. Mult Scler 2009;15:913-7

123. Friese MA, Montalban X, Willcox N, Bell JI, Martin R, Fugger L. The value of animal models for drug development in multiple sclerosis. Brain 2006;129(Pt 8):1940-52.

124. Steinman L. Blocking adhesion molecules as therapy for multiple sclerosis: natalizumab. Nat Rev Drug Discov 2005;4(6):510-8.

125. Fugger L, Friese MA, Bell JI. From genes to function: the next challenge to understanding multiple sclerosis. Nat Rev Immunol 2009;9(6):408-17.

Summary

Multiple sclerosis (MS), the most common neurological disease affecting young adults, is a chronic disorder. It has at least a partial immune component. The disease is characterized by inflammation, demyelination, and primary or secondary axonal degeneration. Clinical features include various neurological dysfunctions, such as visual and sensory problems, limb weakness or gait disturbance. The clinical course can be relapsing-remitting from onset, characterized by relapses with full recovery, or be primary progressive from onset, characterized by chronic disease progression without clinical remission. Like most other (auto)immune disorders, MS is believed to be a complex disorder, arising from complex interactions of both environmental and genetic factors. Typical features for complex disorders are the modest heritability without a classic Mendelian mode of transmission and heterogeneity, which means that a large number of genes contribute to the overall susceptibility. However, for long the major histocompatibility complex (MHC) has been the only genetic association with susceptibility for MS. A recent powerful genome-wide association study (GWAS) performed by the International Multiple Sclerosis Genetics Consortium (IMSGC) revealed the existence of non-MHC susceptibility loci of modest effect. Currently, genome-wide significant association with MS has been shown for different non-MHC variants by further GWAS and replication studies. Identification of genes influencing susceptibility for MS may result in a better understanding of its pathophysiology and may reveal pathways and targets for therapeutical interventions.

In **chapter 1**, a general introduction is given on the epidemiology and genetics of MS.

In **chapter 2**, the pattern of familial aggregation of MS in a recently founded genetically isolated population in the southwest of the Netherlands was studied as part of the research program Genetic Research in Isolated Populations (GRIP). Around 1750 this GRIP population was founded by about 150 individuals. Until recently there was almost no inward and outward migration in this population. During the last two centuries this population underwent an exponential growth. Currently, this population consists of > 20,000 individuals. The availability of extensive genealogical records for this community allows accurate assessment of familial aggregation in this population. For our study we ascertained 48 MS patients from the isolate. Clinical characteristics of MS patients from the isolate did not differ from those of Dutch outbred MS patients. Of the 48 MS patients, 24 could be linked to a common ancestor in 14 generations. Using genealogical information, we found that MS patients from the isolate were significantly more often related to each other than a non-MS control group from the same isolate.

Chapter 3 presents our findings regarding the parental relationship of MS patients that could be linked to a common ancestor. This study was performed in the same community as described in chapter 2. We showed that mothers of MS patients were more closely related than their fathers. This skewed relationship shows evidence for a maternal effect in MS.

In **chapter 4** the results of two studies are reported in which we replicated the 17 risk single nucleotide polymorphisms (SNPs), located in 14 loci, identified in the recent GWAS from the IMSGC.

The first study (**chapter 4.1**) was performed in the GRIP population. Apart from the *HLA* locus, the *EVI5* gene on chromosome 1 was confirmed as a novel risk gene. The fact that despite the small sample size of this study, evidence for *EVI5* as a risk allele for MS was found, underlines the strength of

studying genetically isolated populations in complex diseases. Subsequently, this finding was validated for the general MS population in an independent set of MS patients from the Canadian Collaborative Project on the Genetic Susceptibility to MS (CCGPSMS). Again, a weak but significant risk effect was observed, which strengthens the evidence that *EVI5* is a risk gene for MS.

The second replication study (**chapter 4.2**) was performed in three independent cohorts: in MS patients from the GRIP population, in MS patients from the Dutch general population, and in MS patients from the CCPGSMS. Results obtained in this study were also pooled with those obtained in the original IMSGC study and with those obtained in a recent Australian replication study. After meta-analysis, genome-wide significant association with MS was found for six out of the 17 SNPs in five different loci: *HLA, CD58, CLEC16A, IL2R* and *IL7R*. The overall ORs of the non-HLA SNPs were modest, between 1.16 and 1.23. We calculated the predictive power of the 14 by the IMSGC identified risk loci. The predictive power was too low to have clinical significance, for example for diagnostic or prognostic purposes. The discriminative accuracy of *HLA* alone was better than the discriminative value of the other 13 genes together.

In **chapter 5**, we show the results of a GWAS we did in the GRIP population. We reasoned that the relative homogeneity of the population would be advantageous. No SNP in this GWAS showed genome-wide significant association with MS. However, a signal in the *KIF1B* locus on chromosome 1 was detected, that was characterized by multiple associated SNPs. After a logistic regression analysis only one SNP (rs10492972), located in intron 5 of *KIF1B* remained significantly associated with MS. The association with the other SNPs was most likely due to linkage disequilibrium (LD) with the SNP rs10492972. Since *KIF1B* is known to encode a protein involved in axonal transport, we followed-up that SNP in three replication cohorts of MS patients: a cohort from the Dutch general population, a cohort from the Swedish general population and a cohort from the CCPGSMS. In each replication cohort, rs10492972 was significantly associated with MS. A meta-analysis of the results even showed a genome-wide significant association with MS. Most earlier identified risk genes do have a function in the immune system and have overlapping associations with other autoimmune diseases. Our study for the first time identified a risk gene with specificity for the central nervous system, which may explain part of the progressive neurodegeneration seen in MS patients.

Finally, **chapter 6** provides a discussion of our main findings and how these fit into current knowledge of MS genetics and pathogenesis. Furthermore, future prospectives in MS genetics and the consequences for the understanding of the pathogenesis of MS and the development of new therapies for MS are discussed.

Samenvatting

Multiple Sclerose (MS), de meest voorkomende neurologische ziekte onder jong volwassenen, is een chronische aandoening. Het betreft meest waarschijnlijk een auto-immuun aandoening en wordt gekenmerkt door ontsteking, demyelinisatie en primaire of secundaire axonale degeneratie. De ziekte presenteert zich door neurologische uitvalsverschijnselen zoals visusverlies, gevoelsstoornissen, krachtsverlies of problemen met de coördinatie. Het beloop kan vanaf het begin 'relapsing-remitting' zijn en wordt dan gekenmerkt door exacerbaties van de ziekte en remissies. Ook kan het beloop direct progressief zijn zonder enige remissies. Zoals van de meeste (auto-)immuun aandoeningen wordt ook van MS gedacht dat het een complexe ziekte is, die tot uiting komt door complexe interacties tussen genetische en omgevingsfactoren. Bij complexe aandoeningen speelt erfelijkheid een bescheiden rol, zijn meerdere genen betrokken bij het tot uiting komen en ontbreekt een typische Mendeliaanse wijze van overerving. Tot voor kort vormde het 'major histocompatibility complex' (MHC) de enige genetische associatie met MS. In een recent door de 'International Multiple Sclerosis Genetics Consortium' (IMSGC) verrichte genoomwijde associatie studie (GWAS) werden voor het eerst ook non-MHC genen geïdentificeerd met een bescheiden effect op het risico op MS. Middels GWAS en replicatie studies is nu van meerdere non-MHC genen een genoomwijd significante associatie met MS aangetoond. Het identificeren van genen die de kans op MS beïnvloeden zal leiden tot een beter begrip van de onderliggende pathofysiologische mechanismen die een rol spelen bij het tot uiting komen van MS. Dit zal uiteindelijk kunnen resulteren in de ontwikkeling van betere behandelingsmogelijkheden voor MS.

In **hoofdstuk 1** wordt een algemene introductie gegeven over de epidemiologie en genetica van MS.

In het kader van het onderzoeksprogramma 'Genetic Research in Isolated Populations' (Genetisch onderzoek in geïsoleerde populaties; GRIP) bestudeerden we in **hoofdstuk 2** de familiaire aggregatie van MS in een recent ontstane, genetisch geïsoleerde gemeenschap in het zuidwesten van Nederland. Deze geïsoleerde populatie werd rond 1750 door zo'n 150 individuen gesticht. Tot voor kort was er nauwelijks migratie vanuit en naar het isolaat. Gedurende de laatste 200 jaar onderging deze populatie een exponentiële groei. Op dit moment bestaat de populatie uit meer dan 20.000 individuen. De beschikbaarheid van uitvoerige genealogie maakt deze gemeenschap uitermate geschikt voor het betrouwbaar onderzoeken van familiaire aggregatie. Aan onze studie namen 48 MS-patiënten afkomstig uit het isolaat deel. Klinische karakteristieken van deze MS-patiënten verschilden niet van die van MS-patiënten afkomstig uit de algehele Nederlandse populatie. Van de 48 MS-patiënten konden 24 patiënten gekoppeld worden aan één gemeenschappelijke voorouder 14 generaties terug. Met behulp van de genealogische gegevens toonden we aan dat MS-patiënten afkomstig uit het isolaat significant vaker aan elkaar verwant waren dan een willekeurige controle groep van niet-MS-patiënten afkomstig uit het isolaat.

In **hoofdstuk 3** onderzochten we of de moeders van MS-patiënten uit de GRIP populatie, die te koppelen waren aan een gemeenschappelijke voorouder, nauwer aan elkaar verwant waren dan de vaders. Dit bleek inderdaad het geval te zijn, wat sterk suggereert dat een maternaal effect een rol speelt bij het beïnvloeden van de kans op MS.

In **hoofdstuk 4** presenteren we de resultaten van twee replicatie studies van de 17 single nucleotide polymorphisms (SNPs) gelokaliseerd in 14 verschillende gen regio's, waarvan de IMSGC een associatie aantoonde met MS in hun recente GWAS.

De eerste replicatie studie (**hoofdstuk 4.1**) werd uitgevoerd in de GRIP populatie. We bevestigden de reeds bekende relatie tussen *HLA* en MS. Tevens toonden we aan dat *EVI5*, gelokaliseerd op chromosoom 1, een risicogen is voor MS. Het feit dat een relatie kon worden aangetoond tussen *EVI5* en MS in deze relatief kleine populatie bevestigt de kracht van het gebruik van een genetisch isolaat voor het bestuderen van de genetica van complexe aandoeningen. Vervolgens hebben we onze bevinding gevalideerd in een onafhankelijke populatie bestaande uit MS-patiënten die deelnamen aan het ´Canadian Collaborative Project on the Genetic Susceptibility to MS´ (CCPGSMS). Opnieuw kon een zwakke, maar significante associatie van *EVI5* met MS worden aangetoond.

De tweede replicatie studie (**hoofdstuk 4.2**) werd uitgevoerd in drie onafhankelijke cohorten: MS-patiënten afkomstig uit GRIP, MS-patiënten afkomstig uit de algehele Nederlandse populatie en MS-patiënten die deelnamen aan het CCPGSMS. Uiteindelijk werd een meta-analyse verricht waarbij de resultaten werden samengevoegd met die van de IMSGC en met die van een Australische groep. Er werd genoomwijd significante associatie voor 6 van de 17 SNPs, gelokaliseerd in 5 verschillende loci aangetoond: *HLA, CD58, CLEC16A, IL2RA* en *IL7R*. De odds ratio's van deze non-HLA SNPs waren laag (1,16-1,23). We toonden aan dat de voorspellende waarde van de 14 door de IMSGC geïdentificeerde risico varianten laag was, en geen consequenties zal hebben voor de klinische praktijk zoals voor diagnostische of prognostische doeleinden. De discriminatieve waarde van enkel *HLA* bleek beter te zijn dan die van de andere 13 gen loci samen (0,63 versus 0,60).

In **hoofdstuk 5** tonen we de resultaten van een GWAS die we uitvoerden in de GRIP populatie. Geen enkele SNP in deze GWAS toonde een genoomwijd significante associatie met MS, maar in het *KIF1B* locus op chromosoom 1 bevonden zich wel meerdere SNPs die een associatie toonden. Na een logistische regressie analyse bleef slechts één SNP (rs10492972), gelokaliseerd in intron 5 van het *KIF1B* gen, significant geassocieerd met MS. Meest waarschijnlijk was de associatie met de andere SNPs het gevolg van 'linkage disequilibrium' (LD) met de SNP rs10492972. Vanwege het feit dat *KIF1B* ook een eiwit codeert wat betrokken is bij axonaal transport, hebben wij de desbetreffende SNP gerepliceerd in 3 cohorten MS-patiënten: een cohort afkomstig uit de algehele Nederlands populatie, een cohort afkomstig uit Zweden en een cohort dat deelnam aan het CCPGSMS. In elk cohort bleek rs10492972 significant geassocieerd te zijn met MS. Een meta-analyse van de resultaten leverde zelfs een genoomwijd significante associatie op van rs10492972 met MS. Onze studie is de eerste waarin een risicogen wordt geïdentificeerd met specificiteit voor het centraal zenuwstelsel. Mogelijk kan het deels de neurodegeneratieve component van MS verklaren.

Tenslotte worden in **hoofdstuk 6** de belangrijkste bevindingen van onze studies besproken, waarbij we ze in een breder perspectief plaatsen.

Dankwoord

Dit proefschrift had nooit tot stand kunnen komen zonder de bijdrage van vele mensen, die ik bij deze graag zou willen bedanken.

Mijn promotoren Prof.dr.ir. C.M. van Duijn, Prof.dr. R.Q. Hintzen en Prof.dr. B.A. Oostra zijn de drijfveren achter dit proefschrift.

Beste Cock, ik heb veel geleerd van de kritische manier waarop jij naar mijn artikelen keek. In een mum van tijd wist jij er altijd weer de zwakke punten uit te halen en de artikelen te verbeteren.
Beste Rogier, ik heb je leren kennen als een zeer gedreven onderzoeker. Je ongelofelijke kennis en de daaruit voortvloeiende ideeën en gedachtegangen hebben mij weten te stimuleren en motiveren.
Beste Ben, met name het afgelopen jaar heb ik veel geleerd van je opbouwende kritiek, wat zeker heeft bijgedragen aan een goede afloop van mijn promotietraject. Ook wist je een aantal zaken in de complexe genetica voor mij te verduidelijken.

Veel dank ben ik verschuldigd aan de overige leden van de leescommisie, te weten Prof.dr. C. Wijmenga, Prof.dr. P.A. van Doorn en Prof.dr. J.D. Laman.
Prof. Wijmenga, ik waardeer het zeer dat u als hoogleraar uit Groningen en deskundige op het gebied van de genetica wilt plaatsnemen in de promotiecommissie, waarvoor mijn dank.
Beste Pieter, hartelijk dank voor je steeds aanwezige interesse in mijn onderzoek en je enthousiasme. Dankzij jou konden de formulieren voor mijn promotie toch nog op tijd worden ingevuld.
Beste Jon, ik wil je hartelijk danken voor je kritische blik, je nuttige adviezen en prikkelende vragen gedurende mijn promotietraject.

Prof.dr. P.A.E. Sillevis Smitt en Prof.dr. B.A. 't Hart wil ik danken voor het plaatsnemen in de Grote Commissie.

Alle patiënten, hun partners en familieleden wil ik danken voor deelname aan het onderzoek. Bedankt ook voor jullie gastvrijheid gedurende mijn bezoeken bij jullie thuis.

Ook alle huisartsen, onder enthousiaste vertegenwoordiging van Pieter Snijders, uit de regio Rucphen en neurologen uit het Amphia ziekenhuis in Breda en het Franciscus Ziekenhuis in Roosendaal wil ik bedanken voor hun bereidheid mee te werken aan ons onderzoek.

Petra Veraert, dank voor je onmisbare gedrevenheid bij het verzamelen van de genealogische gegevens van de deelnemende MS-patiënten uit GRIP.

Yurii and Fan thanks for the genetic analyses you did. Without you I would not have been able to write any of my papers.

Dank ook aan de mensen werkzaam in het laboratorium van de genetische epidemiologie, voor het verwerken en opslaan van alle samples die wij steeds weer aanleverden en uiteraard ook voor alle typeringen die werden uitgevoerd. Ik heb de samenwerking altijd als prettig ervaren!

Co-auteurs van mijn artikelen wil ik allen danken voor hun bijdrage aan mijn proefschrift.

Max Loomans en Mirha Halilovic wil ik bedanken voor hun hulp bij de dataverzameling.

Alle mede-onderzoekers van de 22e en collega's van de neurologie, bedankt voor de collegialiteit en gezelligheid, zowel op de werkvloer als daarbuiten.

Heleen, als ervaringsdeskundige, zorgde jij ervoor dat ik door de bomen het bos bleef zien, bedankt daarvoor en ook voor je interesse en vriendschap.

Beste Dorine, ontzettend bedankt voor je luisterend oor en je goede adviezen.

Jeannette Lokker en Jacqueline Habermehl, bedankt voor de administratieve ondersteuning rondom mijn promotie.

Lieve Laura, bedankt voor al je hulp bij het verzamelen van materiaal voor ons onderzoek. Wat was het altijd gezellig om samen met jou en Josje de aan ons onderzoek deelnemende patiënten te bezoeken. Dankzij mijn uitstekende topografische kennis kwamen jullie ook eens op plekken waar jullie anders nooit zouden komen, met of zonder lekke band. Wat vond ik het erg toen je vertelde dat je ons ging verlaten omdat je elders een baan had gevonden. Gelukkig blijven we onze gezellige avondjes, meestal onder genot van een heerlijke stamppot bereid door Josje, voortzetten.

Mijn paranimfen Josje en Naghmeh, ik ben ontzettend blij dat jullie mijn paranimfen zijn.
Lieve Josje, eigenlijk promoveren we vandaag samen. Zonder jouw hulp, steun en luisterend oor zou ik nooit zo ver hebben kunnen komen. Naast de bergen werk die je voor het onderzoek verzette, zorgde jij door je gezelligheid en humor altijd weer voor veel werkplezier, bedankt daarvoor!
Lieve Naghmeh, wat heb ik veel aan je gehad, zowel bij het verzamelen van data als bij het bespreken van allerlei zowel leuke als minder leuke onderzoeksperikelen. Ik waardeer je betrokkenheid en nuchtere adviezen erg.

Lisette, wat bijzonder samen op één dag promoveren! Leuk om met jou samen een feest te kunnen geven!

Het schrijven van dit proefschrift leidde er wel eens toe dat de tijd voor mijn gezin, familie en vrienden onder druk kwam te staan, iets wat na 12 januari zeker zal veranderen.

Lieve vrienden, ik wil jullie ontzettend bedanken voor de gezelligheid naast het werk en jullie begrip en interesse. Door voor- en tegenspoed de laatste jaren is me extra duidelijk geworden hoeveel jullie voor mij betekenen!

Mijn oom, Theo Bemelmans, bedankt voor het ontwerpen van mijn kaft en voor het opvrolijken van ons huis.

Mijn familie en schoonfamilie wil ik danken voor hun steun. In het bijzonder wil ik mijn broer en zus, Stijn en Anne-Marie, noemen omdat ze heel belangrijk voor mij zijn.
Lieve papa en mama, ik had me geen fijnere ouders kunnen wensen dan jullie. Altijd staan jullie voor ons klaar en niets is jullie te veel, ontzettend bedankt daarvoor!

Mijn lieve kleine schatten Sophie en Casper, wat voel ik me gelukkig met jullie! Dankzij jullie is thuis komen telkens weer een feest.

Lieve Hok Hay, mijn onmisbare maatje, bedankt voor al je geduld, je steun en je rotsvaste vertrouwen in mij, ook op momenten dat ik het zelf even kwijt was.

About the author

Ilse Anne Hoppenbrouwers was born on august 1st, 1976 in Breda (the Netherlands). In June 1994, she graduated from the 'Katholieke Scholengemeenschap Etten-Leur' in Etten-Leur.
In september 1994, she started medical school at the 'Erasmus MC' in Rotterdam. During this study, she participated in research on the role of IGFBP-1 (Insulin-like Growth Factor Binding Protein 1) and GH (Growth Hormone) in diabetes associated kidney disease at the Department of Molecular Endocrinology of the Erasmus MC (head: Prof.dr. S.L.S. Drop). She obtained her medical degree in December 2000.
For one year she started working in clinical neurology at the Department of Neurology of the 'Maasstad Ziekenhuis' in Rotterdam.
In June 2002, she started working on the studies described in this thesis at the Department of Neurology (Prof.dr. R.Q. Hintzen) and the Genetic Epidemiology Unit of the Department of Epidemiology & Biostatistics (Prof.dr.ir. C.M. van Duijn) in close collaboration with the Department of Clinical Genetics (Prof.dr. B.A. Oostra) of the Erasmus MC in Rotterdam.
From June 2005 onwards, she combined this work with her residency in Neurology at the Erasmus MC in Rotterdam (head: Prof.dr. P.A.E. Sillevis Smitt).
She is married with Hok Hay Oei. They have two beautiful children: Sophie en Casper.

List of publications

Van Neck JW, Dits NF, Cingel V, **Hoppenbrouwers IA**, Drop SL, Flyvbjerg A. Dose-response effects of a new growth hormone receptor antagonist (B2036-PEG) on circulating, hepatic and renal expression of the growth hormone/insulin-like growth factor system in adult mice. J Endocrinol 2000;167:295-303

Hoppenbrouwers IA, Cortes LM, Aulchenko YS, Sintnicolaas K, Njajou O, Snijders PJ, Oostra BA, van Duijn CM, Hintzen RQ. Familial clustering of multiple sclerosis in a Dutch genetic isolate. Mult Scler 2007;13:17-24

Hoppenbrouwers IA, Liu F, Aulchenko YS, Ebers GC, Oostra BA, van Duijn CM, Hintzen RQ. Maternal transmission of multiple sclerosis in a dutch population. Arch Neurol 2008;65:345-348

Hoppenbrouwers IA, Aulchenko YS, Ebers GC, Ramagopalan SV, Oostra BA, van Duijn CM, Hintzen RQ. EVI5 is a risk gene for multiple sclerosis. Genes Immun 2008;9:334-337

Aulchenko YS,* **Hoppenbrouwers IA**,* Ramagopalan SV, Broer L, Jafari N, Hillert J, Link J, Lundström W, Greiner E, Dessa Sadovnick A, Goossens D, Van Broeckhoven C, Del-Favero J, Ebers GC, Oostra BA, van Duijn CM, Hintzen RQ. Genetic variation in the KIF1B locus influences susceptibility to multiple sclerosis. Nat Genet 2008;40:1402-1403

Hoppenbrouwers IA, Aulchenko YS, Janssens AC, Ramagopalan SV, Broer L, Kayser M, Ebers GC, Oostra BA, van duijn CM, Hintzen RQ. Replication of CD58 and CLEC16A as genome-wide significant risk genes for multiple sclerosis. J Hum Genet 2009;54:676-680

Jafari N, **Hoppenbrouwers IA**, Hop WC, Breteler MM, Hintzen RQ. Cigarette smoking and risk of MS in multiplex families. Mult Scler 2009;15:1363-1367

Jafari N, Broer L, **Hoppenbrouwers IA**, van Duijn CM, Hintzen RQ. Infectious mononucleosis-linked HLA class I single nucleotide polymorphism is associated with multiple sclerosis. Mult Scler 2010;16(11):1303-1307

Hoppenbrouwers IA, Hintzen RQ. Genetics of multiple sclerosis. Accepted for publication in Biochim Biophys Acta

* shared first authors

PhD portfolio

	Year	(ECTS)
Specific courses (e.g. Research School, Medical Training)		
Summer course 'Genetic Epidemiology', Erasmus MC, Rotterdam	2002	3
ESNI course, Tampere, Finland	2002	1
Immunity in the central nervous system: MS as a model (MolMed), Erasmus MC, Rotterdam	2005	1
Seminars and workshops		
Dept. of Neurology	2002-2011	2
Presentations		
MS research days, Groningen, NL; poster presentation 'Co-occurrence of autoimmune diseases with MS in a large pedigree'	2004	1
MS research days, Amsterdam, NL; poster presentation 'Familial clustering of MS in a Dutch genetic isolate'	2005	1
Dept. of Neurology, Erasmus MC, Rotterdam, NL; oral presentation 'Genetics of MS'	2006	1
MS research days, Rotterdam, NL; poster presentation 'Maternal transmission of MS in a genetic isolate in the Netherlands'	2006	1
MS research days, Hasselt, Belgian; oral presentation 'Risk alleles for multiple sclerosis in the Netherlands'	2007	1
Dept. of Neurology, Erasmus MC, Rotterdam, NL; oral presentation 'Risk alleles for MS'	2008	1
MS audit, Rotterdam, NL; oral presentation 'First neuronally expressed gene associated with MS'	2008	1
International conferences		
ECTRIMS, Milan, Italy; poster presentation 'Co-occurrence of autoimmune diseases in families of patients with MS'	2003	1
ENS, Barcelona, Spain; poster presentation 'Co-occurence of autoimmune diseases with MS in a large pedigree'	2004	1
AAN, Miami Beach, Florida, USA; poster presentation 'Familial clustering of MS in a Dutch genetic isolate'	2005	1
ECTRIMS, Madrid, Spain; poster presentation 'Maternal transmission of MS in a genetic isolate in the Netherlands'	2006	1
AAN congress, Chicago, NL; poster presentation 'Risk alleles for multiple sclerosis in a Dutch genetic isolate'	2008	1
ACTRIMS, Montreal; poster presentation: 'Risk alleles for multiple sclerosis in a Dutch genetic isolate'	2008	1
Total		20